James A. Kent *Editor*

Handbook of Industrial Chemistry and Biotechnology

Volume 1 · 12th Edition

Springer

Handbook of Industrial Chemistry and Biotechnology

James A. Kent, PhD
Editor

Handbook of Industrial Chemistry and Biotechnology

Volume 1 and 2

12th Edition

Springer

Editor
James A. Kent
Jupiter, FL, USA

ISBN 978-1-4614-4258-5 ISBN 978-1-4614-4259-2 (eBook)
DOI 10.1007/978-1-4614-4259-2
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012952590

© Springer Science+Business Media New York 2012
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To
My Wife ANITA
and our
FIVE SONS

Foreword

Since the time I was a freshman in college, in 1935, I depended greatly on a fantastic source of chemical information, Riegel's *Handbook of Industrial Chemistry*. Later this invaluable source became Kent and Riegel's *Handbook of Industrial Chemistry and Biotechnology*. I am so privileged to have this opportunity to write a Foreword for the 12th edition of this work.

I have used it regularly throughout all of my professional life as a Professor of Chemical Engineering for 54 years and author of 94 technical books. This book is not meant to be an encyclopedia, and it is not. Rather, the *Handbook of Industrial Chemistry and Biotechnology* is a rich source of industry-wide information and insights. Designed and edited by Dr. Jim Kent, the 12th edition of this widely respected and frequently consulted reference work provides a wealth of information and guidance on industrial chemistry and biotechnology. Industries covered span the spectrum from salt and soda ash to advanced dyes chemistry, the nuclear industry, the wonders of biotechnology, and, most recently, electrochemical energy storage devices.

Dr. Emil Riegel started the handbook in the 1930s, and Dr Kent assumed responsibility beginning with the sixth edition in 1965. Over a period of roughly 50 years and seven editions, under his expert and enthusiastic guidance, it has grown from a handful of chapters, covering mostly the heavy chemicals which characterized that era of the chemical industry, to reflect the ever-changing nature of the industrial scene as well as the needs of scientists and engineers who work in the chemical industry. The present edition draws upon the expertise of 70 contributors, virtually all of whom are embedded in the industries they cover, and who wholeheartedly and even passionately share their intimate knowledge and insights.

Several editions ago, reflecting his early industry experience with Dow and Monsanto, Jim decided to introduce new chapters he chooses to call chemical industry infrastructure. Thus, the Handbook contains chapters on Chemical Process Safety, Emergency Preparedness, and the statistical concepts which underlie process design and control, and quality control. Also, reflecting the increasingly important area of energy, the Handbook includes several nicely integrated chapters on energy producing materials with the broad view of their usefulness in serving as raw material for a variety of industrial processes. Chapters on Fertilizers and Agricultural Chemicals provide crucial information needed to feed the growing population of the world. Adding a chapter on Sustainability in recent editions was timely and inspired.

Some years ago, I understand, a reviewer observed that he had referred to the handbook for information on a particular topic, and left it some hours later, having been captured by several of the interesting and informative chapters on topics in which he had no professional interest. The editor has retained that flavor through succeeding editions, and the contributing authors have uniformly presented their information in well-organized and very interesting chapters.

I have on a great number of occasions over the years consulted successive editions of the Handbook on a variety of topics and never left without having found exactly what I was looking for. The Handbook deserves a place in the library of every chemistry and chemical engineering department, as well as on the desk or in the library of everyone who has a responsibility in promoting, managing, regulating, or writing about any portion of the chemical industry. I have always considered it the most helpful book in my personal library.

John J. McKetta, Jr. Ph.D.
The J.C.Walter Chair Emeritus
Department of Chemical Engineering
The University of Texas, Austin, Texas, USA

Preface

The central aim of this book is to present an up-to-date account of the science, engineering, and industrial practices which underlie major areas of the chemical process industry. It contains 38 chapters covering various areas of that industry. It should be noted that the products and processes associated with a particular area are discussed within the context of the corresponding chapter, rather than in the isolated manner characteristic of an encyclopedia.

This edition of the *Handbook of Industrial Chemistry and Biotechnology*, the twelfth, includes the topics included in the eleventh (2007), all of which have been carefully reviewed, updated, and revised as necessary. Indeed, some of the chapters have been completely rewritten by new authors. This edition also includes two entirely new chapters, *Industrial Applications of Nanostructured Materials* and *Electrochemical Energy Storage*. The Handbook follows the essential arrangement of earlier versions, i.e., several chapters devoted to general or “infrastructure” topics, with most of the book being given over to the various areas of the chemical process industry. This includes a wealth of very useful infrastructure material and greatly enhanced process industry content. In keeping with past practice, the chapters are written by individuals with demonstrated expertise and real world experience in their respective fields. All told, the work may in many respects be regarded as a sourcebook for practice in the chemical process industries.

The infrastructure or contextual material mentioned above is contained in three chapters which lie in the area often referred to as “green chemistry.” The first of these, titled *Green Engineering: Integration of Green Chemistry, Pollution Prevention and Risk Based Considerations*, provides an excellent guide for applying the methods of green chemistry and engineering to process and product development activities, whether for new products and processes, or for upgrading older ones. Written by a team of experts in the field, the chapter can be of considerable help to all practicing chemists and chemical engineers, as well as to students studying in either discipline. Another new chapter, *Industrial Catalysis; A Practical*

Guide, is a valuable adjunct to the "Green" chapter, since catalysis is an important aid in the practice of green chemistry. The third chapter in what might be termed the "green" group is entitled *Environmental Chemical Determinations*.

There is an important distinction to be made here. *Green chemistry*, or *sustainable chemistry*, is defined by that chapter's authors as "the use of chemistry to reduce pollution at the source, through the design of chemical products and processes that reduce or eliminate the use or generation of unwanted or hazardous substances." *Green engineering* is defined as "the design, commercialization, and use of processes and products that are feasible and economical, yet at the same time minimize (1) generation of pollution at the source, and (2) risk to human health and the environment." Risk assessment methods used in pollution prevention can help quantify the degree of impact of individual chemicals, and thus are a valuable tool for intelligent design of products and processes.

Biotechnology first appeared in the Handbook several editions back as a chapter titled *Industrial Fermentation*. In later editions that chapter has evolved into four, reflecting the rapidly increasing importance of biotechnology. To paraphrase the authors of one chapter, integrating gene discovery, functional genomics, molecular evolution and design, metabolic pathway engineering, and production processes including formulation of delivery systems, the chapter will help readers learn how to design and produce biotechnology products. Persons interested in the biotechnology area will be pleased by the breadth of coverage found in this handbook. Finally, the chapter entitled *Biomass Conversion*, introduced in the previous edition, addresses an area of great interest in connection with world energy needs. Written by a team whose primary work lies in that area, it provides comprehensive coverage of the subject, from biomass structure and composition to thermochemical and biological routes for conversion to energy and a host of chemicals and products including liquid transportation fuels. This chapter defines the opportunity for using sustainable sources of biomass as feedstock for new refineries that will produce fermentable sugars and chemical intermediates, from which several forms of fuels can be made.

As mentioned earlier, the core of the Handbook comprises 30 chapters which are devoted to various areas of the chemical process industry. This information, together with the supporting infrastructure material described above, viz., process safety, emergency preparedness, statistical methods, green engineering and chemistry, provides in toto many sophisticated and useful tools to aid in the design of new products and processes and for the evaluation of older ones. The handbook should also prove useful to individuals who possess a background in chemistry or chemical engineering and work in related areas such as regulatory agencies and environmental organizations. Among other benefits, it will help ensure that the work of such individuals reflects knowledge of relevant contemporary science and engineering and industry practices. Reflecting realities in the world energy situation, the Handbook, as in past editions, includes a chapter devoted to the nuclear industry.

Individuals who have responsibilities in the chemical process industries are usually engaged, consciously or otherwise, in continually reviewing their operations to ensure that they are safe, efficient, and in compliance with current environmental regulations. They are also, or should be, anticipating future needs. It is hoped that the information contained herein will provide the wherewithal by which chemists, chemical engineers, and others who have an interest in the process industries, for whatever reason, can ensure that they have touched every base, dotted every *i*, and crossed every *t* in their quest to make the processes and products for which they are responsible as environmentally sound, safe, and efficient as possible.

Because of the scope of the book and the large number of products and processes it covers, some redundancy is inevitable. For example, more than one chapter includes discussions of gasification and hydrogen production. However, there are significant differences in emphasis in the various discussions. Thus, rather than distract readers by referring them to information in locations outside that of their primary interest, such topics have been left intact in the context in which they are discussed.

As in all the earlier versions of this work for which I have been privileged to serve as designer and editor, I am happy to acknowledge again the unselfish and enthusiastic manner in which the contributing authors have shared their knowledge and insights so that many others may learn and still others may benefit. The picture of a bit of knowledge, like a pebble tossed into a quiet pond, its ripples spreading ever more widely, is, I think, apt. And if knowledge is power, the authors who have contributed their expertise to this work are pleased to have had this opportunity to empower others. All have been unstinting in their efforts to make their contributions as complete and informative as possible, within the boundaries of the space available, and I am indeed humbled and honored to have had a part in bringing it about. Needless to say, errors of omission and shortcomings in organization are mine.

Finally, grateful acknowledgement is made to the publishing houses and technical/scientific organizations for permission to reproduce copyrighted illustrations, tables, and other materials, and to the many industrial concerns which contributed drawings, photographs, and text material. I wish also to express my thanks to the Springer editor, Dr. Kenneth Howell, for his many helpful suggestions and support along the way. And finally, I wish to recognize with thanks the hard work and patience of my well-organized and capable daughter-in-law, Pamela Kent, for her unstinting efforts in helping to prepare the final manuscript.

Jupiter, FL, USA James A. Kent

Contents

Volume 1

1 Recent History of the Chemical Industry 1973 to the Millennium: The New Facts of World Chemicals Since 1973 1
Fred Aftalion

2 Safety Considerations in the Chemical Process Industries 47
John F. Murphy

3 Managing an Emergency Preparedness Program 107
Thaddeus H. Spencer and James W. Bowman

4 Applied Statistical Methods and the Chemical Industry 131
Stephen Vardeman and Robert Kasprzyk

5 Green Engineering: Integration of Green Chemistry, Pollution Prevention, and Risk-Based Considerations 155
David Shonnard, Angela Lindner, Nhan Nguyen, Palghat A. Ramachandran, Daniel Fichana, Robert Hesketh, C. Stewart Slater, and Richard Engler

6 Industrial Catalysis: A Practical Guide 201
Robert J. Farrauto

7 Environmental Chemical Determinations 231
William L. Budde

8 Nanotechnology: Fundamental Principles and Applications 249
Ranjit T. Koodali and Kenneth J. Klabunde

9 Nanostructured Materials: Industrial Applications 265
Kristin Clement, Angela Iseli, Dennis Karote, Jessica Cremer, and Shyamala Rajagopalan

10 Synthetic Organic Chemicals 307
Scott D. Barnicki

11 Chemistry in the Pharmaceutical Industry 391
John F. Kadow, Nicholas A. Meanwell, Kyle J. Eastman, Kap-Sun Yeung, and Joseph Payack

12 Manufactured Textile Fibers 419
Bhupender S. Gupta

13 Dye Application, Manufacture of Dye Intermediates and Dyes 475
H.S. Freeman and G.N. Mock

14 The Chemistry of Structural Adhesives: Epoxy, Urethane, and Acrylic Adhesives 549
Kirk J. Abbey and Dennis J. Zalucha

15 Synthetic Resins and Plastics 573
Rudolph D. Deanin, Joey L. Mead, Ming Wei, and Bridgette M. Bludhlall

16 Rubber 621
D.F. Graves

17 The Agrochemical Industry 643
K.N. Singh and Kavita Merchant

18 Petroleum and Its Products 699
Paul R. Robinson, S. Romanow-Garcia, and H.L. Hoffman

19 Coal Technology for Power, Liquid Fuels, and Chemicals 749
Burtron H. Davis and James Hower

Volume 2

20 Natural Gas 807
James D. Idol, Mahmood Moshfeghian, and Arland H. Johannes

21 The Nuclear Industry 829
Tom Congedo, Edward Lahoda, Regis Matzie, and Keith Task

22 Synthetic Nitrogen Products 875
Gary R. Maxwell

23 Phosphorus and Phosphates 939
Glenn A. Gruber

24 Fertilizers and Food Production 959
Amit H. Roy

25 Sulfur and Sulfuric Acid 997
Gerard E. d'Aquin and Robert C. Fell

26 Salt, Chlor-Alkali, and Related Heavy Chemicals 1017
Thomas F. O'Brien

27 Industrial Gases 1039
Steven J. Cooke

28 Wood and Wood Products 1057
Yuan-Zong Lai

29 Pigments, Paints, Polymer Coatings, Lacquers, and Printing Inks 1117
Dean C. Webster and Rose A. Ryntz

30 Industrial Biotechnology: Discovery to Delivery 1131
Gopal K. Chotani, Timothy C. Dodge, and Michael V. Arbige

31 Industrial Enzymes and Biocatalysis 1183
Joseph C. McAuliffe

32 Industrial Production of Therapeutic Proteins: Cell Lines, Cell Culture, and Purification 1229
Marie M. Zhu, Michael Mollet, and Rene S. Hubert

33 Biomass Conversion 1249
Stephen R. Decker, John Sheehan, David C. Dayton, Joseph J. Bozell, William S. Adney, Bonnie Hames, Steven R. Thomas, Richard L. Bain, Stefan Czernik, Min Zhang, and Michael E. Himmel

34 **Animal and Vegetable Fats, Oils, and Waxes** 1323
Edmund W. Lusas, Muhammad S. Alam, Richard C. Clough, and Mian N. Riaz

35 **Sugar and Other Sweeteners** 1403
Mary An Godshall

36 **Soap, Fatty Acids, and Synthetic Detergents** 1431
Janine Chupa, Steve Misner, Amit Sachdev, Peter Wisniewski, and George A. Smith

37 **Chemical Explosives** 1473
Don H. Cranney and Walter B. Sudweeks

38 **Electrochemical Energy Storage: Applications, Processes, and Trends** 1497
Gerardine G. Botte and Madhivanan Muthuvel

Index 1541

Recent History of the Chemical Industry 1973 to the Millennium: The New Facts of World Chemicals Since 1973

1

Fred Aftalion

Overcapacities and the Search for Remedies

The first oil shock that occurred at the end of 1973 with the Yom Kippur war served to pinpoint the crisis which world chemicals were already undergoing.

The chemical industry's soaring development after the war was due to the extraordinary burst of innovations occurring between 1935 and 1955 and coinciding with sectors served by chemicals. Production units multiplied in Europe as well as in the United States and Japan.

Two other factors contributed to this rapid growth. The use of oil as a substitute for coal provided the chemical industry with abundant, cheap raw material that was easy to transport. With interest rates lagging behind the rate of monetary erosion over a number of years, industry leaders were tempted to carry out investments that they would not have made had currencies remained stable and interest rates higher. The fear of these leaders that competition would get the better of them if they slowed down their investments, the race for market shares advocated by a number of consultant firms like the *Boston Consulting Group,* the belief—quite widespread among world chemicals leaders—that they had to keep building new units to keep up with forecast needs, all had a share in building up production overcapacities which were already becoming apparent before 1973 in certain sectors of heavy chemicals (petrochemicals, synthetic fibers, thermoplastics, and fertilizers).

The establishment of an OPEC cartel that led to a rise in the price of a barrel of crude oil from $3 to $12, then the 1979 Iranian Revolution which made it soar to $40, and finally the publication of the gloomy forecasts of the Club of Rome experts which mistakenly saw oil shortages ahead when, in fact, these had been artificially engineered by the Cartel members—all these facts upset chemical leaders in industrialized countries. And yet some of them still continued to invest in new plants during the stock-building lulls that occurred in 1974 and 1979 through consumers' speculating on new price rises.

This only made the necessary adjustments much harder when they had to be carried out at the beginning of the 1980s. Companies were to suffer greatly from an error of judgment, building new plants at great expense at the same time that economic growth rates tumbled from over 10% to a mere 2–3%. Caught between the increasing cost of their hydrocarbon raw materials and the ever-lower prices they had to use to sell their products in markets where offer exceeded demand, leading chemical companies in industrialized countries were forced to go through agonizing reappraisals.

This led them to act in a number of different directions. First and foremost, they had to lower their operating costs by cutting down on excess personnel and taking the measures needed to increase the productivity of each company. At the same time, they had to reduce, in a concerted way if possible, the overcapacities affecting the hardest-hit sectors. Finally, it seemed advisable to redirect production into areas that were less sensitive to economic change. This meant increasing the share of specialties in relation to commodities in overall turnover.

A new generation of leaders was called upon to carry out the socially painful and politically delicate job of rationalizing and restructuring the chemical industry through layoffs and plant closures. These same leaders were also given the more exalting, but just as difficult, task of defining the redeployment strategy that needed to be followed and of determining on a case-by-case basis the sectors that should be abandoned and those that, on the contrary, had to be invested in force.

By 1973, it was obvious that the chemical industry had reached a degree of maturity to the extent that all the companies involved in that area in industrialized countries were long established and that no discovery likely to affect

---

F. Aftalion (✉)

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_1,

its development had been made over the last two decades. While new areas of research like composite materials and biotechnologies had emerged, no immediate fallout was expected for a number of years. Thus failing any rapid internal growth brought about by major scientific breakthrough, the strategy of leaders anxious to refocus or diversify their portfolio of activities very often consisted of a kind of monopoly game, as a range of production was shifted from one enterprise to another without anything new being created.

The Restructuring of Sectors in Distress

Priority action was required in petrochemicals, in the large thermoplastics, in fertilizers, and in synthetic fibers where the most serious investment mistakes had been made. The hardest cases were those of *petrochemicals* and *thermoplastics*. For one thing, a steam cracker cannot technically operate under 60% of its capacity. For another, the products that emerge are linked to one another in almost invariable proportions. Finally, a polymerization unit cannot have its pace slowed down without this affecting the upstream monomer unit to the same extent.

In addition to such rigidities, there was the need to reduce not only the quantities produced but also the number of production units. The problem then arose of sharing the sacrifices among the different producers within an economic area.

The problem was most easily solved in Japan because of the discipline which MITI managed to establish within the country's petrochemical industry. Making the most of a new law that allowed competing producers to act in concert, a cartel was set up with the object of cutting down ethylene production. Four groups of petrochemical producers were formed within which the necessary arbitrations took place. This led Sumitomo to close its Niihama units, Mitsubishi a number of its Mizushima plants, and Showa Denko two of its Ohita installations.

At the same time, producers reached agreements on cutting down competing PVC and polyolefin sales networks, while MITI authorized the import of naphtha through an organization consisting of Japan's petrochemical producers. Its price served as a marker for naphtha produced in Japan.

In Europe, of course, it was difficult to show such disregard for market laws. The views of the European Economic Community Commission in Brussels had to be taken into account, and they upheld the principle of free competition as set down in article 85 of the Rome Treaty. Moreover, in Western Europe there were a number of petrochemical industries that operated according to the rules of private capitalism while there were others, as in France, Italy, Austria, Norway, and Finland, that were state controlled and more concerned about retaining market share than ensuring profitability.

Despite such obstacles, unilateral decisions were taken and bilateral arrangements carried out among firms, leading to some measure of production rationalization. Between 1980 and 1984, 25 ethylene and 8 polyethylene units were scrapped in Western Europe while ethylene oxide capacities were reduced by 10%.

The 1983 agreement between ENI and Montedison put some order in Italy's chemical industry, as ENI took over the PVC and polyethylene operations of Montedison. Previously in France, Rhône-Poulenc had sold its petrochemicals division and its thermoplastics to the Elf Aquitaine group. At the same time, steam crackers were being shut down in Feyzin and Lavera, and a vinyl chloride unit in Jarrie. The association between BP Chimie and Atochem in polypropylene and the exchange of Atochem's Chocques unit for ICI's Rozenburg polyethylene unit were other instances of rationalization.

The Brussels Commission also gave its approval to three large-scale operations: the ICI and BP Chemicals exchange of polyethylene/PVC, the vertical integration of vinyl chloride involving AKZO and Shell Chemicals, and the recent Enichem and ICI association, which produced European Vinyls Corporation and was intended to lead to major capacity cuts in PVC.

In West Germany, rationalization measures were less spectacular because the heads of Germany's leading chemical companies had not waited for the crisis to delineate their respective fields of operation and to establish close links with international oil companies, either through long-term supply contracts or through parity associations.

A number of American companies became involved in restructuring. Union Carbide sold its Antwerp site to BP Chemicals; Monsanto, its Seal Sands acrylonitrile unit to BASF; Esso, its Stenungsund steam cracker to Statoil; while Hercules joined up with Montedison to set up the Himont company, which accounted for 20% of the world polypropylene market.

In the *United States* the petrochemical industry set its house in order along purely capitalistic lines. Each company involved acted alone for fear of infringing antitrust legislation and the main concern was to restore profitability. Unlike Du Pont, which acquired Conoco, other chemical companies tried to get rid of their petrochemicals. Hercules sold its DMT units to Petrofina, and subsequently its 40% stake in Himont to Montedison, while Monsanto was shedding its Texas City petrochemical site.

Major divestments took place, particularly in the major thermoplastics, which were taken over by individual entrepreneurs who bought up the units the chemical giants wished to get rid of. As Hoechst, Union Carbide, Du Pont, Monsanto, ICI, and USS Chemicals withdrew from a number of the major oil-based intermediates as well as from polystyrene, polyethylene, and PVC, a number of large, hitherto unknown companies emerged: Huntsman Chemical, El Paso Products, Aristech, Vista Chemical, Sterling Chemicals, and Cain Chemical.

At the same time, oil companies were integrating downstream petrochemicals and polymers. Such was the case of Occidental Petroleum, which through its chemical subsidiary Hooker (later Oxychem) bought up Tenneco and Diamond Shamrock's PVC in 1986, becoming the largest American producer in this area. Likewise, BP Chemicals fully acquired its subsidiary Sohio. The long-standing petrochemical divisions of the large oil groups returned to profits in 1986 after some painful tidying up but no agonizing reappraisal, helped along by falling oil prices and dollar rates.

Most of them had cut down on operating costs and diversified to the point at which they were able to face up to the economic ups and downs without too much apprehension. Productivity improvements and a better utilization of existing capacities because of higher demand put Exxon, Mobil, and Texaco on the way to prosperity in petrochemicals in 1986.

Standard Oil of California added the petrochemicals of Gulf Oil, purchased in 1984, to its subsidiary Chevron Chemical. Other United States petrochemical producers took advantage of special circumstances. Amoco was served by a strong terephthalic (TPA) base and its good performance in polypropylene; Arco, by its Lyondell subsidiary in Channelview, Texas, and by its development of the Oxirane process through which propylene oxide could be produced by direct oxidation with styrene as a coproduct. The process also led to MTBE (methyl tertiary-butyl ether), the antiknock agent used as a substitute for tetraethyl lead.

Even Phillips Petroleum, badly affected by Boone Pickens' takeover attempt, managed to make substantial profits from its petrochemicals because of drastic restructuring. New prospects were also opening up for the United States chemicals industry as needs grew for butene and hexene comonomers used to produce linear low-density polyethylene (LLDPE), also as consumption of higher olefins to prepare detergent alcohols increased and as demand for MTBE used as a gasoline additive soared.

The problem of overcapacities in *chemical fibers* in each economic region was both easier to overcome because of the small number of producers and more complicated because of outside factors. In *Europe,* producers suffered heavy losses from 1973 onward. For one thing, the Europeans were not particularly suited to manufacture chemical fibers at satisfactory cost, a fact that was proved by growing imports from Southeast Asia. For another, the capacity increases decided upon did not tally with any comparable increase in demand in the foreseeable future.

In view of such imbalance, one might have thought that a number of producers would withdraw from the market. But this did not happen because some of them had to heed government instructions to maintain employment. Also textiles accounted for only a share of the business of the companies involved and could be kept up through the profits generated in other areas. From 1978 to 1985 two agreements were implemented with the blessing of the European Economic Community Commission. The first aimed for a linear reduction of existing capacities; the second and more important one allowed each producer to specialize in those areas where it held the best cards, giving up what amounted to marginal productions.

Thus Courtaulds withdrew from polyester and from Nylon to concentrate on its acrylics and cellulose fibers; ICI focused on Nylon and Bayer on acrylics; Rhône-Poulenc withdrew from acrylics but revamped its Nylon and polyester units well integrated in upstream intermediates; Montedison decided in favor of polyester and acrylics; AKZO focused on polyesters and on aramide fibers while keeping up its profitable rayon sector. Such efforts, which aimed to reduce European chemical fiber capacities by 900,000 tons and to increase productivity through specialization, undoubtedly corrected the situation.

Nonetheless, European producers are still faced with two kinds of competition: first imports of synthetic fibers from Turkey, Taiwan, South Korea, and Mexico, against which it is hopeless to expect that the multi-fibers agreements—which contravene GATT rules—will constitute a permanent obstacle; and second, imports of natural fibers such as cotton, for which prices have fallen spectacularly in recent times.

The Japanese solution to chemical fiber overcapacities naturally involved MITI which pushed through a 17% cut in existing polyester, Nylon filament, and acrylic fiber capacities between 1978 and 1982. These were linear cuts, however, and did not restrict the range of synthetic fibers developed by each producer, contrary to the specializations that marked the second stage of Europe's approach.

The United States was faced with an additional problem because its market remained wide open to textile imports from developing countries. These imports constituted an indirect threat to American producers of chemical fibers. Their first reaction was to reduce their bases in Europe. Du Pont closed its acrylic units in Holland in 1978 and in Northern Ireland in 1980; the following year it ceased production of polyester thread in its Uentrop unit in Germany. Monsanto did likewise in 1979, shutting down its Nylon units in Luxembourg and Scotland and selling its acrylic fiber installations in Germany and Ireland to Montedison.

In the United States itself, capacity cuts were not so substantial and the 1983 upturn boosted utilization of remaining units to 80% of their capacities. Major American producers such as Du Pont, Celanese, and Monsanto returned to satisfactory profit margins. Other companies for which fibers were not an essential sector withdrew from this area. Chevron Chemical, for instance, shut down its Puerto Rico Nylon and polypropylene fiber units between 1980 and 1982 as well as the polypropylene fiber unit in Maryland.

The fertilizer market was in no better shape than the petrochemicals and chemical fibers markets, for world producers had largely allowed supply to exceed demand.

The situation in this area was further complicated by the unequal distribution worldwide of the raw materials required to produce fertilizers and the special attention which governments bestowed on agriculture. Such attention had led to a surfeit of production units and their increasing control by governments, either directly through taking a stake in the companies concerned, or indirectly through establishing ceiling prices for home sales or export subsidies. The emergence of new producers in Eastern countries and in developing areas increased the share of state-controlled companies in world production from 30 to 64% for ammonia, from 40 to 65% for potash, and from 10 to 46% for phosphoric acid between 1967 and 1986.

In Western Europe, nitrate fertilizer producers had deemed it expedient to set up a cartel arrangement for exporters called Nitrex. But the collapse of demand in countries outside its area had prevented it from functioning properly, sparking a fight for market shares even within the community.

As a country like Morocco switched from its long-established role as phosphate exporter to downstream ammonium phosphate and superphosphate integration, traditional fertilizer producers were forced to reappraise their strategy and take severe rationalization measures.

Japan which had none of the required raw materials and, accordingly, had high production costs, began, as early as the 1970s, gradually to cut down capacities along the lines jointly agreed upon by the authorities and the five main Japanese producers of nitrate and phosphate fertilizers.

In *Europe,* the pressure of events disrupted the whole market as the number of producers was drastically reduced. Because of market proximity, production both from Eastern Europe of nitrates and from Africa of superphosphates were becoming dangerously competitive. Supply conditions for natural gas varied according to each country's policies. France, for instance, agreed to pay extra for Algeria's gas, while Holland's Groningen gas, which Dutch ammonia producers were getting at a very favorable price, was linked to the price of petroleum products. On the other hand, a number of Scandinavian state-controlled companies like Norsk Hydro and Kemira were pushing ahead with ambitious fertilizer programs, taking advantage of their interests in North Sea oil or of the conditions under which they were being supplied with oil and gas from the Soviet Union.

Between mergers and acquisitions, the structure of the fertilizer industry in Western Europe was spectacularly pared down. A few giants emerged to dominate the market. In France, there was CdF Chimie, later to be known as ORKEM, which had just taken a 70% stake in Air Liquide's subsidiary La Grande Paroisse, and Cofaz, which was taken over by Norsk Hydro; in Western Germany, there was BASF and Ruhr Stickstoff; in Britain, ICI and Norsk Hydro, which bought up Fisons; in Italy, ANIC and Montedison's subsidiary Fertimont; in Holland, DSM's UKF and Norsk Hydro NSM; in Finland, Kemira, which took over both Britain's Lindsay and its Kesteven facilities.

But the scene has not yet become sufficiently clear, since the competing companies do not all enjoy, within the community, the same raw materials supply conditions, and Europe is still open to imports from other countries that do not apply the rules of market economy.

In the *United States,* the situation was in many ways different. With its large sulfur, natural gas, phosphate, and even potash resources, America's fertilizer industry rested on a sound base. It was an exporter of minerals and fertilizers, and did not have to worry to the same extent as Europe's industry about competing imports from Socialist countries. But reserves of sulfur extracted by the Frasch process have been depleted in Louisiana and Texas, and President Ronald Reagan's "payment in kind" (PIK) farm-acreage cuts reduced the fertilizer requirement of American farmers. These farmers are also much in debt and are having trouble selling their products on saturated markets.

Consequently, very little money has been sunk into extracting phosphate rock in Florida or in increasing nitrogen fertilizer capacities, for a new ammonia and urea unit can cost as much as $250–500 million to build in the US, depending on the state of the existing infrastructure.

With such dim market prospects, it is understandable that W.R. Grace has decided to shut down its Trinidad ammonia unit, or that a company as large as International Mineral Chemicals has tried to diversify through purchase of Mallinckrodt and has put half its fertilizer assets up for sale.

The Nationalization of France's Chemical Industry

When a left-wing government came to power in 1981, France's chemical industry was in dire straits judging from the losses of the major groups: CdF Chimie was losing 1,200 million francs; Péchiney Ugine Kuhlmann 800 million francs; Rhône-Poulenc 330 million francs; Chloé Chimie 370 million francs; Atochimie 130 million francs; and EMC 100 million francs. Admittedly world chemicals were in poor shape. But while French leaders were posting losses amounting to 7–10% of their turnover, Hoechst and BASF were still making consolidated profits that year of 426 million DM and 1,290 million DM, respectively, even though they had noticeably slumped.

There were many reasons, some of them old, for the difficulties of France's chemical industry as illustrated by losses of 7 billion francs in 7 years—4 billion francs in 1981 alone. Caught between increasingly heavy charges and price controls on the home market, France's chemical entrepreneurs never managed after the war to achieve sufficiently profitable margins. They ran up high debts to make up for their lack of funds, building up ever heavier financial costs.

A further disadvantage of France's chemical industry was its scattered production sites, originally due to the need during the two World Wars to keep plants far from the battlefields. For both social and political reasons, it was inconceivable in France to have a site like BASF's Ludwigshafen where 52,000 people are concentrated on six square kilometers with three thermal power plants and countless production sites. The first concentrations which President Georges Pompidou sought to carry out had not changed things much, neither had they cut down increased operating costs. Indeed, the leaders of merged companies had not cared at the time to close sites down and reduce personnel, two moves that might have improved the performance of the new groups.

Although the state spent considerable sums for chemical research, particularly through CNRS and the universities, the fallout for industry was scarce because of the persistent lack of communication between industry and those doing research.

The research and development sectors of the companies themselves made few breakthroughs, so that the chemical industry had to rely for a large part on foreign technologies, a fact that left little room for maneuver.

In addition to the difficulties inherent in their environment, France's companies also suffered the effects of bad management decisions in specific areas. Rhône-Poulenc had been badly prepared for the chemical fibers slump and had sunk too much money in heavy chemicals. These did not fit in with the group's original calling, as its leaders demonstrated when they withdrew, at the height of the crisis, from petrochemicals and the base thermoplastics, concentrating on specialties. The purchase of GESA from PUK in 1978, of Sopag the following year from the Gardinier brothers, and the sale of Lautier were hardly fortunate decisions for a group that could draw no advantage from getting further into fertilizers and that could have diversified to good purpose on perfumes through Lautier.

PCUK had never managed to strengthen Francolor's international base to good purpose and had finally sold it to ICI. Also, it wasted a lot of money in belatedly trying to develop a PVC chain. In 1981, PCUK was negotiating with Occidental Petroleum the sale of its chemical division, which had long since ceased to be of interest to the group's leaders.

At no time since it was set up was CdF Chimie master of its destiny, subject as it was to political pressures rather than economic rationality. Constantly in the red despite a number of worthwhile activities, it received the final blow when the untoward decision was taken in 1978 to build, on borrowed money, a one-billion-franc petrochemical site in Dunkirk in the framework of Société Copenor set up in joint venture with the Emirate of Qatar.

Elf Aquitaine had established under Sanofi a small conglomerate with profit-making subsidiaries involved in pharmaceuticals and perfumes. But Atochem, set up on a joint basis by Total and Elf, was a loss-making concern, as was Chloé Chimie, a cast-off of Rhône-Poulenc, which retained only 19.50% of its capital, while Elf and Total each acquired a 40.25% stake in the new chemical entity.

EMC was more a mining than a chemical company. It focused on potash, having restricted its diversification to the purchase of the animal food company Sanders and to a subsidiary in Tessenderlo, Belgium.

It was in this environment that the nationalization measures decided upon by the new Socialist government took place. The state took control of 40% in value of production of commodity chemicals and 70% of petrochemicals in France, an event that had no precedent in the free world's industrial countries.

Société L'Air Liquide, which figured as one of the companies to be nationalized on the initial Socialist list, escaped this fate, no doubt because the disadvantages of taking over this star multinational had been pointed out to the President of the Republic by one of his brothers, who was adviser to the group. On the other hand, Roussel-Uclaf, which had never needed state funds, found the government partly in control of its capital in addition to the main shareholder Hoechst.

Short of the extreme solutions advocated by some socialists in favor of a single French chemical entity, the nationalized part was cut up along the lines announced by the Ministry of Industry on November 8, 1982. The restructuring signaled the death of PCUK as an industrial enterprise. Its various sectors were shared out among the other state-controlled groups. Most favored was Rhône-Poulenc, which received the agrochemicals and pharmaceuticals sectors with Sedagri and Pharmuka as well as the Wattrelos and La Madeleine sites in the north of France, together with a plant in Rieme, Belgium. At the same time, its fluorine division was boosted. The lion's share went to Elf Aquitaine with what amounted to two-thirds of PCUK's turnover, including, in particular, the halogen and peroxide products.

Complex negotiations with Total (Compagnie Française des Pétroles) ended with the group's withdrawing from Atochimie and Chloé Chimie, after which Elf Aquitaine set up its Atochem subsidiary to encompass all its chemical activities. After a long and brilliant independent career, Rousselot was split between Atochem and Sanofi.

Already sorely tried, CdF Chimie came out the worst from the restructuring. It inherited the Oxo alcohols and organic acids of the Harnes unit and had to call upon Esso Chimie to ensure their survival; it also got an ABS unit that was too small, which it exchanged with Borg Warner for a 30% share in their European subsidiary company—the Villers Saint-Paul site, which could become profitable only with the help of the industries to be set up there; the polyester resins division of the Chauny unit, and the downstream activities of the Stratinor subsidiary, both open to stiff competition. Among the lot there were some profitable sectors, however, such as Norsolor's acrylics, well integrated on the Carling site, and Société Lorilleux, a small ink multinational of PCUK's. But CdF Chimie was left to manage the difficult fertilizer sector swollen by Rhône-Poulenc's and EMC's divestments (GESA and APC), as well as a petrochemical branch set off balance by the unfinished Dunkirk site. As for EMC, all it got from PCUK was the historic site of Loos, which nevertheless served to boost its chlorine and potash divisions.

This enormous restructuring job, no doubt, did produce chemical groups with sounder bases and a more promising future. But the financial cost to the country was considerable, for not only were the shareholders refunded with public money to compensate for nationalization, but the companies that were now state controlled had to be bailed out: their losses in 1982 were even higher than those registered the previous year. Just as high was the social cost. Manpower cuts which the former company leaders had been loath to carry out had become not only absolutely necessary but also easier to implement by a leftwing administration.

Restructuring in Italy and Spain

As was to be expected, the path to overcapacities aided by state subsidies had brought *Italy's chemical industry* to the edge of the precipice. In 1981, SIR and Liquichima, on the brink of bankruptcy, had been taken over by ENI, the state-controlled oil group whose own chemical subsidiary ANIC was also losing considerable sums of money. Montedison had been able to show balanced books only once in 10 years, in 1979. Its debts had soared to $2 billion in 1984.

The rather belated restructuring measures consisted, in their first stage, in the sale of the state's 17% share in Montedison to private interests. Then Italy's petrochemicals and plastics companies were shared out between Montedison and ENI's chemical subsidiary Enichem.

These two groups then set out to concentrate their efforts on polyesters and acrylics in the fibers area. At the same time, Montedison gave up control of SNIA Viscosa, specializing in polyamides, to Bombrini-Parodi-Delfino (BPD). The restructuring, carried out together with manpower cuts and unit shutdowns, made it possible for Montedison in 1985 and Enichem in 1986 to post operating profits after long years in the red. Enichem received a further boost from association with ICI in PVC and with BP and Hoechst in polyethylene, for it had emerged from the restructuring in a less favorable position than Montedison because it was still saddled with commodity chemicals.

Montedison, now 45% owned by the Ferruzzi sugar group, reinforced its strategic sectors by purchasing Allied-Signal's fluorine polymers through its stake in Ausimont, by fully acquiring the Farmitalia and Carlo Erba pharmaceutical subsidiaries, and by buying from Hercules its 50% share in Himont, the joint subsidiary set up in 1983 in polypropylene.

The two Italian giants were still very much in debt, a fact that could lead to further divestments. But their leaders could nevertheless contemplate the future with some equanimity. Their heavy chemical sectors were finally merged under Enimont in 1988.

The *Spanish chemical industry* was also faced with considerable difficulties. Short of innovations, it had developed through foreign technologies and had lived a sheltered life behind customs barriers and import licenses not conducive to cost cuts. Neither Spain's petrochemicals industry, which was in the hands of the Enpetrol state group and the private company CEPSA, nor the main national companies Explosivos de Rio Tinto (ERT) and Cros, were in a position to face without transition the pressure of competition felt when Spain joined the Common Market. This was particularly true of ERT, which had missed bankruptcy by a hair, and Cros, which had remained in the red for a long time. Neither would be able to avoid severe restructuring.

Their total merger project failed through lack of financial means, and it was Kuwait in the end which, through the Kuwait Investment Office, took a 47% share in ERT and 24% in Cros in 1987 and promised to provide the necessary cash for the two groups to form a joint fertilizer subsidiary.

Arab Countries Gain a Foothold

As soon as OPEC was set up, Middle Eastern countries had sought to find ways to invest their oil revenues in downstream industries. *Kuwait*'s approach was, preferably, to acquire shares in existing companies. It thus bought up Gulf Oil's interests in Europe, took a share in Germany's Hoechst, and injected considerable capital into ERT and Cros in Spain.

Qatar had chosen to associate with CdF Chimie to set up a petrochemical base in the Emirate and to build the Dunkirk site through Copenor.

Saudi Arabia's policy has been to develop a national petrochemical industry that would sell its products worldwide. More than Qatar and Kuwait, it had abundant supplies of ethane and methane extracted from gases that were being flared. The ethane separation capacities of its refineries alone accounted for a potential of 3.5 million tons a year of ethylene.

Sabic, the body in charge of the project, had cleverly involved itself with major international groups such as Mobil, Exxon, Shell, and Mitsubishi. Production would then be easier to place in Europe, North America, and Southeast Asia without wounding national feelings. The first giant methanol unit came on stream in 1983, while the other Saudi productions located in Al Jubail and Yanbu have gradually begun supplying low- and high-density polyethylene, ethylene glycol, ethanol, dichloroethane, vinyl chloride (monomer and PVC), and styrene as the relevant units came on stream.

Since 1970, Saudi Arabian Fertilizer has been producing urea and melamine in Dammam, in association with Sabic; the two companies have scheduled construction of a 1,500-tons-a-day ammonia unit in Al Jubail.

Because of the obviously low cost of the principal local methane and ethane raw materials, and because the fixed costs of the installations are high with regard to variable costs, European petrochemical producers were afraid that Saudi Arabia with its low home consumption would flood outside markets with its ethylene derivatives and methanol at cut prices. So far, however, Saudi exports have not shaken up the market because they have been carefully channeled through the distribution networks of Sabic's international partners.

Taking a different course than Algeria with its liquefied natural gas, the Gulf States have thus upgraded their natural resources and already account for 10, 5, and 4% of world production of methanol, ethylene, and polyethylene, respectively.

The American Chemical Industry Caught off Balance

The difficulties resulting from world overcapacities were enhanced in the United States by the behavior of financial circles and the reaction to this behavior of the US chemical industry leaders. America's chemical giants had reached their advanced stage of development because of the long patience of their shareholders and the acumen of their leaders based on 30 years of product and process innovation. Just like their German and Swiss counterparts, the US chemical industry leaders had upheld the notion of long-term interest over the more immediate concern of the various types of shareholders.

The shock waves sent out by the two oil crises, which had not spared the United States, the growing influence of financial analysts on the behavior of shares quoted on the Stock Exchange, and the arrival at the head of the large industrial groups of graduates from glamorous business schools trained more in finance than in technology gave the scene a new twist. Shareholders were more interested in the instant profits they could draw from breaking up a group than with the added value that could be patiently built up through its development.

Drawn along by their own convictions or under pressure from bankers and "raiders," US chemical leaders were constantly redeploying their activities. The *leveraged buy-out* (LBO) system had already been applied by the leaders of FMC's American Viscose division when they sought to buy, with the help of the banking world in the early 1970s, the Avtex rayon and polyester producer, which thereby became a successful company. Despite the risk to buyers in borrowing from financial organizations as much as 90% of the amounts needed for the purchase, the system was eagerly seized upon by individuals wishing to set up their own business and taking advantage of the disenchanted mood of potential sellers. This is how *Huntsman* became the world's leading producer of styrene and polystyrene after buying up the relevant sectors from companies like Shell and Hoechst, which wanted to pull out of them.

Likewise, it is because Du Pont, having spent $7.4 billion to acquire Conoco, sought to reduce its debts by selling part of Conoco's chemicals and also because Monsanto, ICI, and PPG were withdrawing from petrochemicals, that firms like *Sterling Chemicals, Vista Chemical,* and *Cain Chemical* have emerged since 1984. Cain Chemicals was itself to be taken over by Oxychem (Occidental Petroleum) in 1988. Various acquisitions made at the right moment turned Vista within 3 years into one of the leading PVC and detergent alcohol producers in the United States. Through purchases made in its behalf by Sterling Chemicals, Cain Chemicals became a major petrochemical company with assets worth $1 billion in 1987, including ethylene, ethylene oxide, glycol, and polyethylene units, all strategically located in the Gulf of Mexico area. A further newcomer on the American scene was *Aristech,* which emerged through the takeover by its management of the heavy chemicals division of USX (US Steel).

All these companies were acquired under very favorable conditions, as more often than not they were sold by the large groups at 25% of their replacement value. Contrary to assumed notions, individual entrepreneurs were thus able to acquire installations which until then only the most powerful groups could afford to run. These groups gave up whole sections of their traditional chemicals to redeploy in specialties for which they had no particular disposition and, at times, in areas even further removed from their original areas of competence. Thus *Diamond Shamrock* gave up its chemicals to Occidental Petroleum at the worst possible time, to devote itself exclusively to the energy sector, which in fact failed to live up to expectations.

One of the most powerful of America's chemical companies, *Allied Chemical,* became a high-technology conglomerate under the leadership of Edward L. Hennessy, Jr., who was formerly with United Technology. After acquiring Bendix and Signal, it took on the name of *Allied-Signal* and is now focusing on electronics and space, having entrusted a large part of its chemicals to the portfolio subsidiary Henley, which will sell them to the highest bidder. As for *Monsanto,* it shed a number of fibers, plastics, and petrochemical units both in Europe and in the United States and decided to hinge its further development on biotechnologies, a new area for the group. It bought up in particular the aspartame producer *Searle* for $2.6 billion.

At the same time as these changes were being wrought by the protagonist themselves, other major changes were taking place under outside pressure. Wily businessmen acting as "raiders," with the help of financial concerns that issued high-risk and high-interest "junk bonds" to finance a large share of the targeted acquisitions, set their sights on large companies quoted on the Stock Exchange: they acted in the belief that the company's parts would be worth more sold separately than as a whole.

The raiders' takeover bids had instant attraction for shareholders, and their criticism of the way the firms they were after were being managed was often not without truth. But it stood to reason that once the raiders had bought the company, they would break it up to reduce financial charges and to refund the money borrowed for the raid. The more interesting assets were often the first to be sold off, for they found ready buyers. To counter the raiders, the managers of the targeted firms were likely to raise the ante. But this only aggravated the financial problem, and the group's dismantling was unavoidable.

The instant advantage which both shareholders and raiders drew from these operations was obvious. But their consequence was, sooner or later, to destabilize the enterprises concerned, when these did not disappear altogether. The most spectacular case was *Union Carbide,* coveted in 1985 by the real estate developer S. Hayman, who had already taken over GAF Corporation.

To fight off the raid, Union Carbide had to borrow $3 billion. To reduce such an unbearable debt, the group's management was forced to sell its best sectors (batteries, consumer products, engineering plastics, agrochemicals) and even its headquarters in Danbury, Connecticut. This was how one of the best chemical concerns in the United States, with sales amounting to $10 billion, was left with only three areas of business after divesting to the tune of $5.3 billion. Even these areas—industrial gases, petrochemicals and plastics, and graphite electrodes—were faced with stiff competition. And with debts that still remain three times as high as the industry's average, Union Carbide is in no position to invest in the short term in anything likely to push it back to its former major rank in chemicals.

Other US companies involved in chemicals were also the victims of raiders in 1985. To fight off C. Icahn, *UniRoyal* was taken over by its management and was forced to sell off its chemicals to Avery, which in turn placed them on the block, before accepting a LBO by the management. *Phillips Petroleum* had to buy back its shares from C. Icahn and B. Pickens and was forced to sell $2 billion worth of assets to refund part of its debt. And what about *Gulf Oil,* which sold itself to Standard Oil of California to escape the clutches of Boone Pickens, or *Stauffer Chemical,* which changed hands three times within a single year from Chesebrough-Pond to Unilever and finally to ICI, when it was broken up among ICI, AKZO, and Rhône-Poulenc?

Attracted to the US market, European investors had also joined the raiders' ranks. This is how the Britain-based Hanson Trust managed to acquire *SCM*. This was a company that had just completed its restructuring; but after Sylvachem was sold off by the new owners, it retained only chemical production of titanium dioxide.

Anglo-French tycoon J. Goldsmith, unable to take control of Goodyear, nevertheless made substantial profits from his raid on the company. Goodyear was left with the sole alternative of withdrawing from all the sectors except chemicals in which it had diversified outside of tires.

In a number of cases, transactions led to an agreement between the heads of companies that had stock options and were eager to make a profit, and the potential buyers. This was how *Celanese,* an able and well-diversified company that had the means to retain its independence and competitiveness with regard to any major company, was acquired by Hoechst following a transaction that was satisfactory both to the German buyer and to the shareholders of the American group, at least for the time being.

The fear that their company might be the target of an "unfriendly" takeover bid induced the boards of directors of some of the well-managed chemical companies to guard against such attacks either through deceptively appealing offers—"poison pills"—or through purchase of their own shares. This was certainly not the best for industrial firms to make use of their funds.

Coping with Safety and Environmental Problems

Handling chemicals has never been without danger, if only because of the unstable and harmful nature of a number of substances when they are placed in certain conditions of temperature, pressure, or concentration.

Chemists have always been haunted by the risks of explosion. The explosion which occurred on September 3, 1864, in the Heleneborg laboratory near Stockholm, where Alfred Nobel was handling nitroglycerin, caused the death of five persons, including Emile Nobel, his younger brother. The ammonia synthesis unit set up by BASF within the Oppau plant was totally destroyed in 1921 by an explosion causing the death of over 600 people. In 1946, the French cargo ship *Le Grand Camp,* carrying 2,500 tons of ammonia nitrate, exploded in Texas City, killing 512 people. Other disasters, such as that of Flixborough in England, which took place through rupture of a Nypro caprolactam pipe within the plant in 1974, or again the one caused in a holiday camp in Los Alfaques in Spain when a tank wagon carrying propylene exploded in 1978, are reminders of the explosive nature of certain chemical products and of the need to handle them strictly according to the prescribed security rules.

A number of chemicals, fortunately a limited number, become dangerous either when they are used wrongly, or when they are accidentally set free. *Thalidomide,* put on the market in 1957 by the German company Chemie Gruenenthal, was indeed a powerful sedative. But it took 3 years to perceive that when prescribed to pregnant women, it dramatically crippled the newborn children. The synthetic intermediate for insecticides, *methyl isocyanate,* which Union Carbide has used for years without incident in its West Virginia Institute plant, caused over 2,000 deaths when it escaped in 1984 from a storage tank in Union Carbide's Bhopal plant in India.

Other products act insidiously, so that it is harder to establish their effects on human and animal health and more generally on the environment. Indeed, progress in understanding the safe dosage of minute quantities of impurities has enabled governments to fix with greater care the maximum allowed content of *vinyl chloride monomer, formaldehyde,* and *benzene* beyond which these products could become dangerous for workers to handle.

Lessons have been drawn from accidents caused by faulty handling of certain substances. Through the work carried on by Alfred Nobel, we know how to stabilize nitroglycerin in the form of dynamite, and since 1946 methods have been devised to avoid the spontaneous explosion of ammonium nitrate. Ammonia units with capacities of 1,500 tons a day have been operating for decades without incident.

Because of the painful thalidomide episode, long and costly tests are now carried out to study the possible secondary effects of pharmaceutically active substances. A great number of drugs that today save many lives would not have been available had they needed to go through the long periods of tests that are now required by legislation.

Likewise, in industrial countries, increasingly stringent regulations limit noxious vapor discharge from chemical plants, which are required to treat their effluents effectively. The transport of dangerous substances is also closely monitored by the authorities. Such precautions stem not only from the publicity which the media now gives to any catastrophe worldwide, but also from the public's instinctive distrust of chemistry, which it still regards as a mysterious science.

But just as an air crash does not mean the end of commercial aviation, neither does the damage caused by improper use of certain substances mean the end of the chemical industry. The image of chemicals is tarnished, however. Citizens who deliberately risk their own death, when they are not actually killing others, because of speeding on the roads or because they are addicted to alcohol, tobacco, or drugs, are less and less inclined, for all that, to accept accidental security breaches when these are not caused by themselves.

Politicians in our parliamentary democracies who wish to please public opinion feel the urge to take into account demands that are more emotional than scientific, and advocate restrictions even when these go against the best interests of the citizens. The *Three Mile Island* nuclear power plant accident in the United States which resulted in no fatalities, the more recent *Chernobyl* explosion which, as of 1988 had directly caused two deaths, have, with no good reason, prevented any resumption of the US nuclear program and have aroused fears in European countries in people least likely to give way to mass hysteria.

The *Seveso* leak, which occurred in Italy on July 10, 1976, in the trichlorophenol unit belonging to Hoffmann-La Roche's subsidiary Givaudan, did have an impact on the immediate environment and a number of people were temporarily affected by the dioxin vapors. But the accident caused no lasting harm. It was the publicity which the media gave to it that forced Hoffmann-La Roche to close down the unit, turning Seveso into a dead city.

The litigation over residues left in the ground by Occidental Petroleum's affiliate Hooker, in *Love Canal,* in the state of New York, led to the evacuation of all the area's residents, beginning in 1978. But no clear explanation has yet been given of the ailments some of the inhabitants have been complaining about.

The lack of universally accepted scientific explanations for certain phenomena has often meant that the precautionary measures taken by one country do not necessarily apply in another. Where sweeteners are concerned, for instance, some governments have banned *saccharin* and other governments allow its use. The same is true of *cyclamates* and *aspartame.*

DDT was banned as an insecticide as early as 1974 by most industrial nations. But it is still widely used in many developing countries. The risks of *eutrophication* are perceived differently by governments, so that legislation applying to products for the production of detergents, like *alkylbenzene sulfonate*, *tripolyphosphate*, or *nitriloacetic acid* (NTA), differs from country to country.

The agreement which a number of nations reached in 1987 to ban the use of *chlorofluorocarbons* in aerosols is so far the only instance of harmonized legislation, even though no one has so far managed to prove scientifically that the chlorofluorocarbons (CFCs) really destroy the atmosphere's ozone layer.

Thus while it is understandable that authorities must be careful to soothe the fears of a public that is insufficiently informed of the dangers that threaten it, it must also be aware of the economic and social costs of refusing to accept the risks inherent in any human activity, and also conscious of the uncertainties surrounding the rules and regulations taken to satisfy its demands.

Some companies are turning the necessity of cleaning up the environment into new opportunities to improve their profitability. Thus Du Pont has found a useful application as a building material for the calcium sulfate that was piling up as a by-product in one of its Texas plants.

Scientific and Technological Breakthroughs

Short of fundamental discoveries over the past 15 years, the chemical industry has gone forward by systematically developing its store of knowledge in processes and products.

Process Improvement

Higher crude oil prices had revived studies in the use of coal as a chemical feedstock. But while the Fischer–Tropsch synthesis was still used in South Africa by Sasol, the only other industrial gasification unit was the one Eastman Kodak brought on stream in Kingsport, Tennessee, in 1983, to produce *coal-based acetic anhydride.* The coal came from the Appalachian Mountains and was cheap enough relative to oil prices at the time to warrant such an installation, and the plant is now to be expanded.

Together with these studies on synthetic gas, some progress has been achieved in the use of a group of aluminosilicates, the *zeolites*, as selective catalysts to boost certain reactions. Half the world production of *p*-xylene and a quarter of the production of ethylbenzene, an intermediate required to prepare styrene, are carried out using the zeolite-based ZSM-5 catalysts developed by Mobil Oil, which played a pioneer role in this area.

Applications of the olefin *metathesis* reversible chemical reaction, discovered by Phillips Petroleum in the 1960s, were also developed in the subsequent years. By this reaction, Arco produces propylene from ethylene and butene-2; Hercules prepares its plastic, Metton, from dicyclopentadiene; and Shell synthesizes its C_{12}-C_{14} SHOP (Shell Higher Olefin Process) alcohols used for detergents.

The application of *electrochemistry in organic synthesis* had already served to bring on stream in the United States in 1965 Monsanto's first industrial adiponitrile process from acrylonitrile. This was followed in 1977 by a similar installation in Seal Sands, England, which was later bought up by BASF.

The former *Reppe chemistry,* still practiced in Germany by BASF and in the United States by GAF, also led to new developments as demand for certain intermediates such as the 1,4-butane-diol increased. This diol, now also obtained from maleic anhydride, is used to produce PBT polyesters through reaction with TPA acid and leads to other major derivates (tetrahydrofuran, butyrolactone, *N*-vinylpyrrolidone).

New synthetic processes for the preparation of established products were also industrially developed: in Japan the manufacture of methyl methacrylate from C_4 olefins, by Sumitomo and Nippon Shokubai; in France, the simultaneous production of hydroquinone and pyrocatechin through hydrogen peroxide oxidation of phenol by Rhône-Poulenc; in the United States the production of propylene oxide through direct oxidation of propylene operating jointly with styrene production, developed by Ralph Landau and used in the Oxirane subsidiary with Arco, which the latter fully took over in 1980; in Germany and Switzerland, the synthesis of vitamin A from terpenes, used by BASF and Hoffmann-La Roche.

Processes apparently well established were still further improved, such as the *electrolysis of sodium chloride,* dating back to the last century: diaphragm and then membrane cells were substituted for mercury cells, which were a possible source of pollution.

Important progress was also made in *chemical engineering,* such as use of rotary compressors in ammonia synthesis or ICI's fermentation reactors in Billingham to produce the Pruteen protein from methanol reactors, having no mobile parts.

Product Development

Although research was not as fruitful after 1960, new materials put on the market in the 1970s were the outcome of research in high polymers essentially conducted within industry.

It was through such research that ICI's PEEK (polyether ether ketone), one of the first high-performance aromatic polymers, was put on sale, as well as Du Pont's aramide fibers Nomex and Kevlar, more resistant than steel in like volume.

To the range of engineering plastics were added polyethylene and polybutylene terephthalates (PET and PBT), as well as General Electric's poly ethers, the PPO (polyphenylene oxide) produced through polymerization of 2,6-xylenol and the Noryl plastic produced by blending PPO with polystyrene. Other special polymers, derived like the polycarbonates from bisphenol A, were added to this range: polyarylates, polysulfones, polyetherimides.

A major step forward was taken in the area of base thermoplastics with the application of Union Carbide's Unipol process. Variations of this were subsequently offered by other LDPE producers such as Dow and CdF Chimie (now ORKEM).

Under a process that consisted in copolymerizing in the existing high-pressure installation ethylene with 5–10% of an *α-olejin* (butene-1, hexene-1), a stronger linear LLDPE was produced with a higher melting point than LDPE. Thinner films could thus be produced that were just as strong but required less material.

The new polymers opened up an unexpected market for producers of C_4, C_6, and C_8 α-olefins like Shell, Ethyl, and Chevron. Their higher linear α-olefins were also used either for polyalphaolefins (PAO) intended for synthetic lubricants or to prepare detergent alcohols.

While no great new plastic has emerged over the last 15 years, researchers in major chemical companies did their utmost to improve both the features and the performance of known polymers.

As we have just seen, they improved LLDPE by adding comonomers in the carbon chain. But also through additives they managed to render polymers more resistant to fire, to oxidation, and to alteration through ultraviolet rays.

This slowly gave rise to a new industry that consisted in supplying polymer producers and plastic processors, not only pigments and charges, but also antioxidants, light stabilizers, and fireproofing agents. Added in small doses to the polymer, they added to its value by extending its life span. Such an activity, in which the Swiss firm Ciba-Geigy plays a noteworthy role, was boosted by the spectacular development of polypropylene, a particularly sensitive polymer that has to be stabilized with appropriate additives.

Another way of improving the performance of polymers consisted in blending them either with other polymers or with inert materials such as glass fibers, carbon fibers, or various mineral fillers. Thus were produced a series of *alloys and composite* materials. Glass fiber-reinforced polyester has long been in common use. But the possibility of introducing carbon fiber obtained through pyrolysis of polyacrylonitrile (PAN) fibers already developed in aeronautics, opened up fresh prospects, particularly in the area of sports articles. The need, in turn, to link organic polymers and mineral fillers led to coupling agents such as the silanes which Union Carbide and Dynamit Nobel have put on the market.

This is how, little by little, spurred on by the demands of the processing industries which are also under pressure from major clients like the automobile industry, a number of companies have brought a large number of improvements to plastics. While not very spectacular, these improvements have appreciably added value to existing materials.

More generally, the requirements of many downstream industrial sectors have hastened the development of derivatives that otherwise might have remained laboratory curiosities. Discoveries of new molecules have been particularly inspired by the needs of plant protection. This was because agriculture, before it became a crisis sector, offered worldwide markets for crop protection agents, and also because product approval was easier to obtain, and therefore less costly, than in the case of pharmaceuticals. The success of glyphosate, which Monsanto put on the market in 1971 under the trade name Round Up, has made it the world's leading selective herbicide, for it can be used throughout the year and becomes harmless when absorbed into the ground. A new range of synthetic pyrethroids, developed in the United Kingdom by Elliott of the National Research and Development Corporation (NRDC), a government agency, was marketed from 1972 onward under the trademarks of Permethrin, Cypermethrin, and Decis. These wide-spectrum insecticides owe their success to the fact that they are exceptionally active in small doses and are not toxic to humans. With increasingly strict legislation and stiff competition among pesticide producers at a time of slumping agricultural markets, the golden days could well be over for crop protection products, so that the years ahead are likely to be more favorable for restructuring than for new discoveries.

Over the last 15 years, the *pharmaceutical sector* also made great demands on the ingenuity of chemists. But from the time of the thalidomide drama, the testing times required by health authorities have increased, to the point that since 1980 10–12 years are needed instead of the three to four previously required to bring a drug on the market from the time of its discovery. Research and development costs, accordingly, have grown fourfold over the last 10 years, dangerously reducing the number of new specialties provided for patients each year. Because of such delays, a patent protecting a new substance may be left with but a few years of validity when final approval is granted to the laboratory that made the discovery.

Such difficulties have apparently not affected the zeal of researchers. Nor have they diminished the sums devoted each year to research and development, which on the contrary have been constantly on the increase. This is because any major discovery may have worldwide portent. And in most developed countries there is a system of refunding to patients the cost of ethical drugs, so that a new active principle may provide the laboratory that has exclusive rights over it with a considerable source of profits even if such refunds are coupled with tight price controls.

And while it is also true that 30 pharmaceutical companies alone account for 60% of worldwide ethical drug sales, the sums of money invested in research do not always get their full return. Thus it is that a small company like Janssen's laboratory, Janssen Pharmaceuticals, in Belgium, which was acquired in 1979 by Johnson and Johnson and which has among its discoveries diphenoxylate (1963) and loperamide (1975), has proved more innovative over the last 15 years than the Rhône-Poulenc group, which has produced no major new molecule during the same time, although it devotes far more money to its research.

Indeed, success depends at least as much on chance, the ability of researchers, and the strategy of management in that area as on the sums expended. Valium and Librium, which have been providing Hoffmann-La Roche with its largest profits since the end of the 1960s, were the outcome of Leo Sternbach's acumen. Instead of merely modifying the meprobamate molecule as management had requested, he began studying the sedative properties of benzodiazepines used as dyestuff intermediates and on which he had worked for 20 years previously at Cracow University.

One of the most prolific inventors of the 1960s was most certainly Sir James Black, a Nobel laureate in 1988. While working for ICI, he discovered the first β-blocking agent Propranolol in the early 1960s. He also discovered Cimetidine, sold under the trade name of Tagamet as an antiulcer agent by SmithKline & French from 1974 onward, and which has become the world's largest selling specialty. After working successively for ICI, SmithKline & French, and for Wellcome in Britain, Sir James now has his own business, and he is convinced that small competent teams are, by nature, more innovative than the large armies of researchers which many of the big companies have set up.

Likewise, the successful ventures of Merck Sharp & Dohme cannot be dissociated from the work of its president, Roy Vagelos. This biochemist, a latecomer to research, supervised the whole process of work to bring Mevacor, the new cholesterol miracle drug, onto the market. It has just been approved by the US Food and Drug Administration. Mevacor was but the crowning touch to Merck's scientific tradition with its long series of discoveries: α-methyldopa against hypertension, indomethacin and sulindac to fight arthritis, and cefoxitine, an antibiotic.

At a time when pharmaceutical research is becoming increasingly costly and the likelihood of a great discovery remains hazardous, success will come to laboratories which not only sink large sums of money into research but also rely on teams where competence does not necessarily rhyme with size, and whose management has reached a sufficient level of scientific maturity.

The Craze for Biotechnology

The catalytic action of living organisms, or rather of the proteins they contain, had received the beginnings of an explanation with the experiments of Payen and Persoz on malt amylase separation in 1833 and with JJ Berzelius's catalyst theory in 1835. In 1897 Eduard Büchner demonstrated that a yeast extract could turn sucrose into ethyl alcohol. Fermentation took place without the presence of living organisms through enzymes. In this case zymase was the catalyst.

Ethyl alcohol, already known to alchemists, was used by industry towards the middle of the last century when continuous distillation in columns was devised by Ireland's Aeneas Coffey in 1830 and when it became exempt from excise duties on alcohol if methanol was added to it.

After alcohol, *lactic acid* was the second product obtained industrially from sugar fermentation, starting in 1880. The levo-isomer is still made this way to the tune of 20,000 tons a year.

In 1890, the Japanese chemist Jokichi Takamine had introduced a fermentation process in the United States by which an enzyme blend was produced. This takadiastase catalyzed starch and protein hydrolysis. Some years later in 1913, Boidin and Effront discovered the *Bacillus subtilis* that produced an *α-amylase* stable under heat. This enzyme was used to desize cloth and later in the sugar fermentation process.

During World War I, Chaim Weizmann had succeeded in producing for the British Admiralty acetone and butanol on a large scale through anaerobic fermentation of starch. The Germans were then producing as much as 1,000 tons a month of glycerin from sugar. These war productions proved no longer competitive in peacetime. But *citric acid,* which Pfizer began producing in 1923 from sucrose, is still biochemically made today from *Aspergillus niger,* which Currie advocated in 1917.

The discovery of *penicillin* and its industrial development during World War II have led the pharmaceuticals industry increasingly to resort to *biosynthesis* for the preparation of its active principles. Through rigorous selection of the microorganisms extracted from the soil or from various molds, the cost of an antibiotic like penicillin has been brought down to $30 per kilogram, compared with $25,000 per gram initially—an impossible target if the exclusively synthetic process had been used. Moreover, it became possible to extend the range of antibiotics that could be used. The antianemia *vitamin* B_{12} and most of the amino acids were prepared in the same way through culture of microorganisms in selected environments containing precursors.

In the case of *steroids,* biosynthesis permitted reactions that could not be achieved through direct synthesis. In 1952, this was how Upjohn researchers in the United States managed to introduce on carbon atom 11 of the steroid nucleus, a hydroxyl group –OH, using the *Rhizopus arrhizus* fungus, making the switch from the pregnancy hormone progesterone to cortisone and its derivatives.

Microorganisms are also capable of separating optical isomers. In the case of sodium glutamate, where it is necessary to start from levo-glutamic acid to obtain the desired flavor, and where synthesis produces only a racemic blend, it was a particular yeast called *Micrococcus glutamicus* that led to the required isomer through carbohydrate fermentation.

Considering that sodium glutamate, like other amino acids, is contained in soy sauce, which is a traditional Japanese food, it is not surprising that Japan should have become interested very early in this type of fermentation. Firms like Ajinomoto and Kyowa Hakko dominate the world market for amino acids and particularly for *glutamic acid* and *L-lysine*. It is also through enzymes that the resolution of *DL-methionine* into its optical isomers is achieved since its laboratory synthesis yields the racemic form.

Heat-stable amylases are frequently used in both the United States and Japan to produce *syrups with a high fructose content* from corn starch.

Single-cell proteins such as ICI's Pruteen were produced through culturing microorganisms on a bed of organic material.

Interest in biosynthesis grew still further with the discovery in 1953 of the structure of DNA, then in the 1960s of the genetic code of proteins. It then became possible to clone microbe or plant cells, through *genetic engineering,* by recombination of fragments of genetic material from different species. Thus, towards the end of the 1970s, the biotechnology firm Genentech succeeded in isolating the human insulin gene and to insert it into the DNA of the *Escherichia coli* bacteria: through reproduction, these bacteria produced the *first human insulin,* which Eli Lilly and Company has been marketing since 1982.

The *human growth hormone* (HGH), which can only be extracted in minute quantities from the pituitary glands, can now be isolated in larger quantities through genetic engineering.

Monoclonal antibodies (mabs), which replicate the antibodies in the organism with the added advantage of being "immortal," were discovered in 1975 by scientists working at the Cambridge Medical Research Council in the United Kingdom. They serve more particularly as reactive agents for medical diagnostic purposes.

Through *plant genetics,* it has also been possible to render plants resistant to chemical agents (Calgene, Monsanto) as well as to improve crop yields (Pfizer) with new seeds.

With the prospects which *biogenetics* was opening up for medicine and agriculture, a number of private laboratories sprang up in the United States between 1971 and 1978—*Genentech, Cetus, Genetic Institute, Biogen, Amgen*, and *Agrigenetics* to mention but the principal ones. These laboratories managed to finance their work with the help of venture capital, research contracts with the major chemical firms like Du Pont, Monsanto, Eastman Kodak, W.R. Grace, or shares purchased on the stock exchange.

Vast sums of money have been spent over the last 10 years but with small tangible results, prompting the definition of biogenetics as a business likely to bring in a small fortune as long as a large one is invested! Thus far the only commercial fallout of biogenetic research involved human insulin (Eli Lilly), the human growth hormone (Genentech, KabiVitrum), the hepatitis B vaccine (Merck, Smith, Klein-RIT), interferon (Boehringer, Ingelheim), the amylase enzyme (Novo), a number of veterinary vaccines (AKZO Pharma), and monoclonal antibodies for diagnostic reactive agents. Hopes raised by interferon and interleukin-2 as cancer cures have not materialized, but the tissue plasmogen activator (TPA) as a blood clot dissolver in heart attacks was approved by the US Food and Drug Administration (USFDA).

Plant genetic research is encountering opposition from the US Department of Agriculture and the Environmental Protection Agency. Pressured by environmentalists, the US administration is loath to approve developments which could affect the environment in unknown ways. In addition to these administrative obstacles, there is uncertainty over patent rights, for there are no legal precedents. Finally, the biocompanies recently set up will need to associate with large pharmaceutical groups to develop and market the products born of their research.

Generally speaking, although *biotechnology* has acquired credibility in many areas, its development is being slowed by scientific, economic, and administrative obstacles. First and foremost, proteins are complex substances that cannot be handled as easily as the simple molecules involved in traditional organic syntheses.

It is true that Japan's Ajinomoto and Kyowa Hakko, in particular, have become masters of the art of producing amino acids. Likewise, enzymes have remained the specialty of Novo (now Novo Nordisk) in Denmark, Gist Brocades in Holland, and Bayer's subsidiary Miles in the United States, which together account for 60% of the world needs in the area.

Even when they are technologically sound, however, bioproducts may turn out to be economically uncompetitive. The profitability of L-lysine from 1 year to the next, for instance, depends on soy market prices. In the same way, the single cell proteins which BP produced in 1963 in Lavera from a petroleum base, using a process developed by France's Champagnat, never managed to compete with soy cakes for animal food. ICI has also just been forced to close down its 50,000-ton Pruteen unit in Billingham.

At current crude oil prices, the production of ethanol from biomass is not profitable, either. Whether produced from beets, sugar cane, or corn, it can become competitive only if it is subsidized. And these subsidies would only be forthcoming for political reasons: to please their farmer voters, the French, Brazilian, and United States governments would adopt such a policy to absorb excess agricultural products. From cereals, corn in particular, starch is produced and hydrolyzed to form glucose which ferments to ethanol.

Powerful groups like American Corn Products and France's *Roquette Frères* produce starchy matters in this way. The former is also the leading producer of *isoglucose* (a blend of glucose and fructose) in the United States, while the latter is the largest producer of *sorbitol.* Starch can, therefore, compete directly with saccharose both for foodstuffs and for industrial uses as a fermentation or enzyme-reaction base.

This gives rise to a permanent conflict in Europe between the starch manufacturers on the one hand and the sugar and beet refiners on the other, a conflict that the EEC Commission with its *Common Agricultural Policy* of quotas and subsidies has been unable to settle. The only point of agreement between the two parties is the price which they demand for their production from downstream community industries, a price that is far higher than world rates.

Spurred on by the Italian sugar group Ferruzzi-Eridiana, Montedison's and now Enimont's main shareholder and an associate of France's Béghin-Say sugar group, there is a campaign under way to introduce ethanol into gasoline. Farmers, of course, support the move because incorporating 7% of ethanol in gasoline would mean for a country like France the use of two million tons of sugar or four million tons of cereals. But ethanol happens to be in competition with methanol and the new MTBE antiknock agent as a gasoline additive. More important, a tax rebate would be needed at current gasoline prices to induce the oil industry to incorporate ethanol in prime rate gasoline. So the "farm" lobby can receive satisfaction only at the expense of the taxpayer, whether American, Brazilian, or European.

The rules that have always governed the use of ethanol, government policy favoring one agricultural raw material over another, the new constraints that limit the marketing of genetically engineered products—all these factors serve to remind those interested in the development of biotechnology how narrow is their room for maneuvering.

The Fine Chemicals Approach

In their search for products that could provide better margins than those achieved from commodity chemicals, the industry had hit upon *fine chemicals.* These typically involved derivatives from organic synthesis obtained in multipurpose units and sold in relatively small quantities at high prices.

The German and Swiss *dye manufacturers* (Hoechst, BASF and Bayer, as well as Ciba-Geigy and Sandoz) were in the most favorable position to develop such advanced chemicals. They had a long tradition behind them of multiple-stage syntheses involving intermediate derivatives that could also serve to prepare pharmaceutically active principles or pesticides. Starting from a number of major raw materials and working according to the chemical-tree concept, these producers can work down the line to well-defined molecules which they use in their own downstream production or sell as synthetic intermediates to outside clients.

In Europe, the giant ICI group, which had retained a strong position in dyes, also became involved in this kind of chemicals.

France, with PCUK having closed down in 1980 its Société des Matières Colorantes in Mulhouse and then having sold Société Francolor to ICI, had restricted its ambitions in this area. It retained only a few products of Rhône-Poulenc and of its 51% subsidiary *Société Anonyme pour l'Industrie Chimique* (SAIC), located in Saint-Fons and in Mulhouse-Dornach, respectively.

As was to be expected, the US chemical leaders, Du Pont, Allied Chemical, American Cyanamid, GAF, and Tenneco Chemicals, had all withdrawn between 1976 and 1979 from the dyes sector. Only three medium-sized companies were still active in this area: *Crompton & Knowles, American Color,* and *Atlantic Chemical.*

Yet at the end of World War II, America's dye production had been the leading one worldwide. For over 30 years it had enjoyed high customs tariffs protection through the American Selling Price clause. But dyes were produced by giant companies used to large-scale continuous productions. Their engineers were not trained to run month-long syntheses campaigns involving many stages. Moreover, American marketing executives were little attracted to the German methods for "motivating" their clients. There was also the fact that during the 1960s, the US dye manufacturers had come to rely on imported intermediates. With rising prices and the textile slump, they found themselves caught between rising purchase costs and falling selling prices. Finally, unlike their European counterparts, the US manufacturers had never given international scope to their dye business. It remained restricted to the home market.

For all these reasons and also because they were not tied down like the Germans by any prestigious tradition, they unhesitatingly gave up dyes, losing at the same time the know-how needed to succeed in fine chemicals.

With more modest means, other firms were more successful. They either developed their own "chemical tree" or put to good use the know-how acquired through development of certain processes.

Ethyl became a bromine and derivatives specialist and an expert in orthoalkylation (orthoalkyl phenols and anilines). Its acquisition of Dow's bromine activities has given Ethyl a leading role in this field. *DSM* developed its fine chemicals from the benzoic acid produced during manufacture of synthetic phenol by toluene oxidation. *Atochem* took advantage of the sulfur resources of its parent company Elf Aquitaine to build up successfully a thioorganic chemicals industry (thioglycol, mercaptans, DMSO). Its position will be further strengthened by the takeover of Pennwalt. *PPG* in the United States and *Société Nationale des Poudres et Explosifs (SNPE)* in France are producing a wide range of phosgene-based derivatives to be used in the most varied manner (carbonates, chloroformates). More than any other company, *Lonza* has extended its range of fine chemicals (diketenes, HCN derivatives, pyrazoles, pyrimidines). *Reilly Tar* has become a world leader in pyridine and derivatives. *Dottikon* in Switzerland and *Kema Nobel* in Sweden have put to use their nitration experience to extend their range of nitrated intermediates. Among others, *Rhône-Poulenc* and *Montedison* are involved in organic fluorine derivatives while *Hüls*' fine chemicals division has specialized in alkylation, hydrochlorination, and catalytic hydrogenation.

Thus a number of firms with special know-how in a family of products or in processes that were not among the biggest have succeeded in taking a more than honorable place as suppliers of fine chemical derivatives alongside the organic synthesis specialists originating from the dye business.

The Attraction of Specialty Chemicals

Besides fine chemicals sold according to specifications but accounting for only a small part of the sales of major companies, *specialty chemicals* held attractions for companies wishing to diversify. These chemicals involved substances or mixtures whose composition mattered less than the function for which they were intended: the test of success lay in performance. Thus old family businesses or more recent companies born of a leader's entrepreneurial spirit had been successful in performance products, whether these were paints, inks, or glues; or in specialties, cosmetics, detergent, or electronics industries.

Indeed, not much capital is needed to manufacture specialty chemicals compared with what is required for commodity chemicals. The development of new products is both quicker and less costly than it would be to find new processes for large-volume products or to bring to the market an original active principle for an ethical drug.

This largely explains why specialty chemicals managed to remain until the early 1970s products for medium-sized private companies. In the long run, however, the internationalization of trade, the size of advertising budgets for consumer products, and the necessary adaptation to new technologies requiring highly qualified personnel all called for funds that were not always available to family businesses. Many small owners were forced to sell out, and their need coincided with the attraction they held for large chemical groups trying to diversify away from heavy chemicals. They hoped to find in specialty chemicals the profit margins which their traditional branch of chemicals no longer supplied.

Barring a few exceptions such as *Gulf Oil* or *Diamond Shamrock,* which withdrew from downstream chemicals, all the major companies, both in Europe and in the United States, decided to make specialty chemicals a priority in their development strategy. In truth, some of them had not waited for the energy crisis for them to take a firm foothold in the specialty market.

In the United States, *Du Pont* and *PPG* had a long-established reputation in industrial and consumer paints. *W. R. Grace* since buying Dewey & Almy, and *Rohm & Haas* because of its age-old tradition in acrylics, drew substantial profits from their specialties. This was also true of *American Cyanamid* (additives for plastics, cosmetics) and of *Monsanto* (products for rubber, special polymers). Since its withdrawal from the tire business, *BF Goodrich,* aside from its PVC lines, is concentrating now on specialties.

In Europe, ICI had already acquired a large paints sector (Duco, Dulux). The three major German leaders—*Bayer, BASF,* and *Hoechst*—had not yet made great inroads into the specialties market, but the Swiss *Ciba-Geigy* could be said to be particularly well established in certain areas like additives for polymers, in which it was a world leader. *Rhône-Poulenc* had assembled some of its activities within a "chemical specialties" division. But on the whole, they could be said to be offshoots of fine chemicals rather than actual specialties, with the exception of the performance products brought out by subsidiaries such as Orogil, SFOS, Soprosoie, and Vulnax. Orogil is now fully owned by Chevron, however, and Vulnax has been acquired by AKZO. Failing to develop through internal growth, *AKZO* had very early developed its specialties by buying up companies involved in peroxides, paints, oleochemicals, and now rubber additives.

To increase their specialty sectors as fast as possible, the leaders of large companies found it more expedient to do so through acquisitions. The prices paid for the most interesting purchases can be considered high because, very often, they amounted to 15–20 times the profits. But the financial sacrifices made by the buyers seemed worthwhile, for they gained a foothold in the market without the long preliminary work that would otherwise have been needed.

There were, of course, many companies that were sufficiently important or prosperous to escape being bought up. Even then their independence was often at stake. Thus *Nestlé* took a share in the cosmetics group *l'Oréal*; and in the United States, the raider Perelman managed to buy *Revlon.*

Considering that the grass always looks greener on the other side of the fence, for many leaders of the chemical giants diversification into new areas might seem more attractive than mere concentration in well-known sectors; and it was in this sense that specialty chemicals seemed a good proposition. In 1983, *Olin* began to get involved in electronic chemicals by buying up 64% of *Philip Hunt Chemicals* and took a firm foothold in the sector through successive acquisitions. Other groups became interested in enhanced oil recovery and exploration, for the future of oil seemed assured at the time. In both cases, however, the electronics and oil exploration slowdown did not confirm established forecasts. The investments made in these areas have yet to prove their profitability.

Moreover, many firms were unable to contribute anything except capital to the development of sectors far removed from their traditional areas of business. They became discouraged and ended by selling out, not without suffering heavy losses. *Hercules* was seen to back out of its water treatment sector and *Rhône-Poulenc* from its very recently acquired media business.

Even when the businesses acquired are not too different, trouble can arise through disagreement between the new owners and the former boss of the purchased firm over how to manage it. The former tries to impose his own personnel and procedures, while the latter, used to making his own decisions, is unable to fit into a large unwieldy concern. As a large part of the worth of an acquisition in specialty chemicals lies in the competence of the personnel involved, some purchasers have understood that it is to their benefit to leave the day-to-day running of the business to those who have already shown their worth and to centralize only those activities related to the financing of new investments. This was how *Witco* proceeded in the United States, most likely because the father of the current president had founded and built up the business to the point of making it one of America's leading specialty concerns.

ICI followed the same policy when it bought *Beatrice Chemicals* for $750 million in 1985. But in this case, it was important to delegate power, because Beatrice Chemicals consisted of ten distinct companies established in 18 countries and involved in different businesses (composite materials, vinyl resins for paints, leather auxiliaries). Keeping in mind that cultural differences may produce problems that are not always easy to solve, the strong involvement of the big chemical groups in the specialties area over the past years had drastically changed the structures of the sector.

The Paint Industry

Few industries have been as affected by the restructurings of the past 10 years as the paint industry. The extension of markets worldwide owing to the multiplication abroad of client factories of this industry, the technological revolutions brought about by the introduction of electrophoresis, of water-based lacquers, and of powder coatings had the twin effect of pushing the chemical leaders to expand worldwide in this area and to lead those paint companies that were still independent to sell out for want of the funds needed to develop their research base. *ICI*, which was strongly established only in Britain and in the Commonwealth, became the world's leading paint producer with 750 million liters after buying *Valentine* in France and, especially, the *Glidden* division of the US-based SCM for $580 million in 1986.

PPG has been pushed back to second world position with 450 million liters. But with its 100% stake in France's *Corona* and its controlling share in Italy's IVI and in Germany's Wülfing, the US-based PPG has maintained a comfortable technological lead in the application of cataphoresis in automobile bodies, accounting for 60% of the world market in this specialty.

Through its costly $1 billion purchase of America's *Inmont, BASF* has become the world's third-ranking paint producer, leaving behind its German Rival *Hoechst,* which was too busy bailing out its British subsidiary Berger Paints to get a foothold in the US market.

AKZO, which holds an honorable place among the leaders, has not been able to penetrate the United States market, either. Most of its recent acquisitions (Blundell, Permoglaze, Sandtex, Levis) were European.

Other companies with comparable 250-million-liter paint capacities are Japan's *Nippon Paint* and *Kansai Paint,* as well as America's *Du Pont.* These three firms, however, have restricted their ambitions to filling the needs of their home markets. With a broader international base, Courtaulds' subsidiary, *International Paints,* ranks among the top ten, although it is mainly involved in the very special sector of marine paints.

Ranking fifth in the world with its 300-million-liter capacity, *Sherwin Williams* is the only large paint company that has retained its independence. It remains focused on the United States, essentially in the decoration market.

Although France is the world's third largest market for paints after the United States and Germany, none of its national manufacturers has thought of striking out beyond its frontiers. Indeed, most of the French companies involved in the sector, with the exception of *Blancomme* and *IPA*, have been taken over by foreign groups when they were not merged within state-controlled entities. *Astral Celluco* was one of the first to sell out to AKZO, *Corona* was taken over by PPG, *Celomer* by International Paints, *Bichon* and *Lefranc Bourgeois* by Sweden's Becker, *Valentine, Julien, Galliacolor* by ICI, *Ripolin Georget Freitag* became part of the *CdF Chimie* group as did *Duco* which has just been sold to *Casco Nobel,* while *La Seigneurie* was taken over by *Elf Aquitaine*.

In 1988 CdF Chimie, later known as ORKEM, took over full ownership of AVI, a profitable company specializing in decorative paints. Another subsidiary of ORKEM was *Lorilleux,* an ink manufacturer merged in early 1988 with *Coates Brothers* to become the third largest group in its field after Dainippon Ink Company (DIC) of Japan and Germany's BASF.

It must be pointed out that all the international groups involved in paints and inks on a worldwide basis produce, in addition, most of the resins and binding agents needed for their formulations. Only the solvents and pigments are likely to be partly brought in from outside sources. The restructuring of the paint industry has, accordingly, been to the advantage of the new groups. On the one hand, it reduced the number of producers and extended the range of products these producers were putting on the market, and on the other, it supplied a captive market for their resins which, until then they had mainly sold to outside customers.

Surface-Active Agents

Used for their good performance, more often than not in formulations, surface-active agents can be classified as specialties even though the quantities consumed in certain cases might connect them with commodities. The structure of the major part of the detergents industry has remained rather stable over the last few years despite some frontier adjustments. The *washing powder* sector, where advertising costs are considerable, is dominated by a small number of substantial soapmakers who came into business as far back as the nineteenth century: the American companies *Procter & Gamble* and *Colgate,* the Anglo-Dutch group *Unilever,* and Germany's *Henkel*. They are all, in various degrees, involved in the major world markets.

Then there are the Japanese companies *Kao Soap* and *Lion Oil,* which remain confined to their own home territory and to a few Southeast Asian countries. Behind these giants, a number of firms catering to their home markets stand out, such as *Purex* in the United States or *Benckiser* in Germany.

As in paints, France is curiously absent from the area. Since Germany's Henkel recently took over the detergents division of the *Lesieur-Cotelle* group and its trademarks Mir, La Croix, and Persavon, after buying up the Savon de Marseille soap flakes of *Union Générale de Savonnerie* (UGS), the French market is now 94% supplied by the big international soapmakers. The few remaining national firms such as *Chimiotechnic* merely sell their products through the supermarkets.

While the sector now seems to be structurally stabilized, washing powder components are fast changing to take into account the new rules and technologies laid down both by governments and consumers. For the companies which supply the soapmakers, these new rules and regulations are having major consequences throughout the world. Just as the requirement of biodegradability had doomed the use of *branched-chain alkylbenzenes* in industrialized countries in the 1960s and caused the shutdown of a large number of dodecylbenzene sulfonate units, so the new rules established by some governments against *tripolyphosphates* in Europe and elsewhere to ward off eutrophication are likely to wipe out the several-hundred-thousand-ton markets of producers like Rhône-Poulenc, Benckiser, Knapsack, or Montedison.

Replacing TPP by new formulations based on *polyacrylic acid* and *maleic anhydride* would, on the other hand, greatly boost companies like Atochem and BASF, which are very much involved in acrylic chemicals. Likewise the use in Europe of washing machines at temperatures that do not exceed 50–60 ºC, like the ones now used in the United States, should have immediate consequences for the formulation of washing powders. *Perborates,* used extensively in Europe as bleaching agents ever since Henkel invented Persil in 1907, are not very efficient at such low temperatures. Activators such as ethylenediamine tetraacetic acid (EDTA), produced by Warwick in England, are needed to hasten decomposition. *Enzymes* which had been very popular in the 1960s in the United States and Europe, then had disappeared in 1971 because they were considered harmful to the skin, have been reintroduced in washing powder formulations because they do help remove certain stains.

The use of *liquid detergents* is more widespread in the United States, where they account for 20% of the market, than in Europe, where their share does not yet exceed 8% on an average. This has consequences on the consumption of nonionic derivatives.

Different habits as well as different regulations have therefore led to frequent changes in the chemicals supplied to soapmakers. Few industries have changed as much as the detergent industry since the end of World War II as it shifted from soap to synthetic detergents, from branched alkyl benzenes to linear alkyl benzenes, from anionic to nonionic. TPP and enzyme regulations were changed; preference was given at times to perborates, at others to chlorine-based products such as bleaching agents.

To develop surface-active agents for industrial use did not require the same financing as was needed for washing powder consumer products. Therefore, producers of all sizes could become involved. Some of these producers were chemical giants who had gone into the business because they had the available raw materials or the right markets. Indeed, surface-active agents use a number of major raw materials to which suppliers attempt to add downstream value.

In Europe, for historical reasons, large chemical groups have become involved in this area. Thus the dye manufacturers had very early added to the range of products sold to the textile and leather industries, wetting agents, softeners, and dye auxiliaries. *BASF*—a pioneer in synthetic auxiliaries with its Nekal, patented in 1917—*Hoechst, Bayer,* and *ICI* were in fact interested at the same time in the markets which surface-active agents opened for their ethylene oxide, higher alcohols, sulfonating agents productions, and in the fact that they help provide better services for their traditional textile clients.

Hüls, the subsidiary of the German holding company VEBA, had no dye tradition. But it nevertheless acquired the Dutch surface-active unit Servo to ensure captive use for at least part of its ethylene oxide and alkyl benzene production. *BP* followed a similar line when it took over the Belgian company Tensia, selling back some of its product lines to ICI. Already involved in surface-active products through its Lissapol for many years, *ICI* has expanded in this sector by buying *Atlas Powder* and its special range of Tweens and Spans. *Shell*'s interest in surface agents went back to the development of its Teepol. It completed its range with ethoxylates, the "neodols" which used both its higher alcohols and its ethylene oxide. *Montedison* was also involved in surface-active agents through its stake in *Mira Lanza*.

In France, however, there was no vertical integration between the great national chemical industry and the surfactant sector. Producers of the latter had to find the necessary feedstock—whether ethylene oxide, alkylphenols, fatty acids, or higher alcohols—from rival companies, while for instance a medium-sized company like Berol Chemie in Sweden, recently acquired by Nobel Industries, had its own source of ethylene oxide, amines, and nonylphenol in Stenungsund to feed its surfactants division.

In the *United States,* vertical integration was not as thorough as in Europe. Although ethylene oxide producers like *Union Carbide, Dow,* or *Texaco* also had their range of ethoxylates, it was mostly specialized firms that produced the surface agents for industrial uses. The same was true in Japan, although a number of producers such as *Nippon Oil & Fats* for fatty acids, *Kao Corporation* and *Lion Corporation* for fatty alcohols and amines had direct access to their main raw materials.

Thus in addition to the large chemical and petrochemicals companies that had chosen downstream integration, there were a number of important surfactant producers that, in varying degrees, were integrated upstream. The most striking example of this, besides the three Japanese companies just mentioned, is Germany's *Henkel.* Its natural fatty alcohol production exceeds 170,000 tons capacity, and besides fatty acids, it produces its own range of carboxymethylcellulose-based thickeners. Recently Henkel even associated in this area with Hercules within a company called *Aqualon*, now fully owned by Hercules, and acquired from Quantum Chemicals in the United States their fatty acids subsidiary *Emery Industries.*

A number of surfactant specialists have also chosen the market approach. Because they are not tied down by their own produced raw materials, they can use those that are the most suitable for the type of surfactant they wish to offer their clients.

An independent producer like the US-based *Stepan* is in a position to provide a complete range of anionic, cationic, and nonionic agents because it has flexible units in four areas of the United States as well as one in southeastern France in Voreppe.

Witco is in the same position, but its own policy has been to develop through acquisitions rather than through internal growth, buying Humko Chemical and Onyx Chemical.

Right from the start *GAF* acquired, from IG Farben, experience in surfactants still of use today. This activity sector, however, was sold to Rhône-Poulenc in 1989.

With a market lacking the uniformity of the United States market, the European producers serve in greater numbers clients with standards and habits varying from country to country. The Tenneco group's *Albright & Wilson* has had to cover France, Italy, and Spain with its *Marchon* subsidiaries. Germany's *Hoechst, Henkel,* and *Schering*, which bought up *Rewo,* also have a number of subsidiaries abroad that produce their surfactants. Hüls's subsidiary *Servo* has only the single production unit in Delden, Holland. But because of the high concentration of its products, Servo manages to carry out three quarters of its sales abroad.

While the range of products offered by these companies is very wide, some of them, nevertheless, focus on specific sectors. Thus the cationic technology acquired in the United States from Armour by *AKZO* and from Ashland by *Schering* has given both these companies a dominant position in the market of textile softeners both in Europe and in the United States.

Companies like *Rhône-Poulenc, Berol,* and Witco are, for their part, interested in the pesticide formulation market. Fatty amines are in the hands of such European firms as *AKZO*, *Kenobel* (Nobel Industries), and *CECA* (Atochem).

Other European companies, such as *ICI* through Atlas and Tensia, *Th. Goldschmidt*, *Rewo*, and *Servo*, have particularly targeted the lucrative area of beauty care. In the United States, *Miranol* has been very successful with the amphoterics (imidazolines, betaines) for baby shampoos, an activity acquired by Rhône-Poulenc in 1989.

America's *Du Pont* and *3M* and Japan's *Sanyo* pay particular attention to the development of fluorine-based surfactants. *Air Products* with its acetylene derivatives Surfynol and *W. R. Grace* with its sarcosinates (Hampshire Chemicals) have also focused on well-defined segments of the business. With world demand exceeding two million tons, the market of surfactants for industry is of a nature to attract a large number of operators, raw material suppliers, processors of these raw materials into anionic, nonionic, and cationic derivatives, or downstream industries that use surfactants in various formulations.

Flavors, Fragrances, and Beauty Products

The sector of flavorings, perfumes, and beauty products has also had its share of restructuring and technological changes over the past 10 years.

Although many of the raw materials needed in this area still come in the form of essential oils from natural sources like jasmine from Grasse, roses from Bulgaria, ylang-ylang from Madagascar, oak moss from Yugoslavia, an increasingly significant role is now being played by semisynthetic or fully synthetic products.

Thus *terpenes* (α-pinene, β-pinene) can be produced from natural turpentine, as is traditionally done by rosin producers such as *Hercules, Glidden,* or *Union Camp* in the United States, or on a smaller scale in France by *Société des Dérivés Résiniques et Terpéniques, DRT*. BASF and Hoffmann-La Roche, however, have demonstrated that starting from acetylene or isobutylene, terpene chemicals can be synthetically reproduced. Both companies are able to produce both their vitamins and perfume bases in this way.

Likewise, *vanillin* is now largely produced synthetically. The world leader in this area is *Rhône-Poulenc*, which has a unit in Saint-Fons to which was added a unit bought from Monsanto in 1986 on the West Coast of the United States. In the latter plant, vanillin is still extracted from paper pulp liquor. *Menthol* from plantations in Brazil and China is also produced by synthesis since *Haarmann & Reimer*, bought by Bayer in 1954, managed to carry out industrially the resolution of racemic menthol, thus isolating the levo-isomer. *Anethole*, synthesized by Hercules from pine oil, is two to three times cheaper than when it is extracted from star anise.

Instead of identically reproducing natural products, chemists have also succeeded in making cheaper substitutes with similar features. Thus nitrated musks and later macrocyclic musks have become substitutes for more rare natural musk. Major chemical companies became interested in the firms that specialized in perfume chemicals. But their involvement in this area was not always successful, for their business views did not necessarily apply to this new activity.

While Bayer's association with Haarmann & Reimer proved successful, it took Hercules several years, from 1973 on, to understand properly how its *Polak Frutal Works* (PFW) had to be managed. Today it is autonomous and prosperous. In contrast, Rhône-Poulenc ended by selling *Lautier* to *Florasynth* in 1981.

When *Tenneco* bought Albright & Wilson in England, it did not see the point of keeping its *Bush Boake Allen* (BBA) aroma chemicals division. BBA, itself the outcome of a merger of several family businesses, was finally sold to the US-based Sylvachem in 1982. Sylvachem[1] already owned *George Lueders*, an essential oils concern Monsanto sold failing proper management. It would seem, therefore, that among the major chemical companies, only Bayer, Hercules and, more recently, BASF, which bought *Fritzsche Dodge & Alcott* in the United States in 1980, have achieved their downstream breakthrough in the flavor and perfume sector.

On the other hand, the pharmaceuticals group Hoffmann-La Roche, which purchased *Givaudan* in 1963, then *Roure Bertrand Dupont* a little later, has managed to rank third in the world in this difficult area. But the leader is undoubtedly *International Fragrances & Flavors* (IFF), an American company that accounts for 10% of the world market. Set up in 1929 by a Dutch immigrant, A. L. van Ameringen, IFF acquired its current form in 1958 and, pushed along by the creative invention of its perfumers and the quality of its compositions, has never ceased growing.

Close on the heels of IFF is the *Unilever* group, which developed in the field through acquisitions. After consolidating in 1983 its three perfume and flavor subsidiaries—PPL, Food Industries, and Bertrand Frères—to form *PPF International*, the group acquired a foothold in the U.S. market in 1984 with *Norda.* Three years later it merged PPF with Holland's *Naarden*, which was on the decline. Called *Quest International* (Unilever) the new company accounts for over 7% of the world market in its area.

Amongst the world leaders, the only privately owned company, the Swiss-based *Firmenich*, ranks fourth. It has retained its independence both because it was held together by the heirs of the founding family and because it produces quality products based on strong research. A number of smaller companies that do not belong to any multinationals are highly competitive. They include Japan's *Takasago*, which began in 1920, America's *Florasynth*, which took over *Lautier*, and Britain's *Pauls Flavours & Fragrances*, which has just established a hold in the United States market by purchasing *Felton International.*

France, which had in its favor the age-old reputation of Grasse and the world image of its perfumes linked to its haute-couture prestige, is nevertheless absent from the fray

---

[1] Sylvachem belongs to the Union Camp group.

of large suppliers in this area, even though it has some Grasse-based companies like *Mane* and *Robertet* and despite the efforts made by the Elf Aquitaine group which has assembled, around Sanofi, firms like *Méro et Boyveau, Tombarel,* and *Chiris.*

Flavors account for a substantial share of the sales of these firms: 30% for IFF, 40% for Givaudan, 50% for Unilever, and 100% for Sanofi-Méro. They are increasingly being used in foodstuffs since the fashion of fruity yogurts and instant desserts began between 1965 and 1970. The internationalization of food habits and the growing industrialization of the food sector have contributed to the development of demand for flavors and to the gradual substitution of natural substances by synthetic products that are less costly to produce and more active in small doses. Just as the perfume industry composes fragrances for its clients, subtle blends of flavors are now devised for the large food companies. Demanding customers, together with stringent regulations and sophisticated technologies, all combine to build up research costs. This explains the restructuring that has taken place in the sector as family businesses have been taken over by powerful international chemical, pharmaceutical, or food-industry groups, leaving only a few independents willing and able to make the necessary research efforts.

Although it still clings to a long tradition, the *world of perfumes* has also changed both in its structures and in its technologies. The highest volume consumption derives from products of the soapmakers. *Procter & Gamble* prepares its own compositions, but its competitors mostly rely on the laboratories of their suppliers for fragrance preparation.

With a few rare exceptions, such as *Guerlain, Chanel*, and *Patou*, the great names as well as the small perfumemakers do likewise. One of the world's largest-selling perfumes, "Anais-Anais" by *Cacharel* (l'Oréal), is prepared by Firmenich, while *Roure-Bertrand-Dupont* has signed two other recent successful perfumes, *Dior*'s "Poison" and *Saint-Laurent*'s "Opium." Launched in 1921, *Chanel No. 5* was the first perfume to carry a synthetic aldehyde note and is still one of the ten world best-sellers. But the market has now moved to floral and oriental fragrances. Perfumes for men with stronger notes have developed spectacularly and now account for 25% of alcoholic perfumery. In addition, the aerosol format has boosted sales of toilet waters and deodorants.

The most varied distribution systems have been developed, ranging from door-to-door sales, which *Avon* started, to sales by mail, a specialty of *Yves Rocher*'s, to sales in large department stores, to sales in selected areas such as perfume shops and pharmacies. Few "nonessentials" have become so indispensable. If they cannot be dispensed with, it is through the efforts of the industry, which relies upstream on the suppliers of both contents and containers, who adapted to all requirements, and downstream on efficient marketing networks. It can also devote to advertising the money that it need not spend on research conducted on the industry's behalf by the chemists.

Although it originated in France, the perfume industry is now mostly in the hands of foreign firms. While *Parfums Dior* and *Givenchy* (belonging to the Moët-Hennessy-Louis Vuitton group), as well as *Guerlain, Lanvin*, *Nina Ricci*, and *Patou* are still under French control, *Cardin Parfums* belongs to American Cyanamid, *Orlane* to Norton Simon, *Chanel* to the Swiss Pamerco group. *Rochas* was owned by Hoechst, which has now sold it, and *Parfums Saint Laurent* is now controlled by Italy's Carlo de Benedetti, who bought it from the US-based Squibb. As for the L'Oréal group, which had taken over the *Lancôme*, *Jacques Fath*, *Guy Laroche*, *Ted Lapidus*, *Cacharel*, and *Courrèges* perfume brands, it has been within the Nestlé orbit since 1973, although it was arranged that until 1993 it would be managed by those representing the interests of the founding family. Not surprisingly, the same great names recur in the area of *beauty products*, including, besides perfumes, *hair care products* and *cosmetics.* Each firm, indeed, wishes to complete its range by acquiring complementary businesses.

The cosmetics industry was born in the United States with the three great "ladies," *Elizabeth Arden, Harriet Hubbard-Ayer*, and *Helena Rubinstein.* In their wake are now *Estée Lauder,* the giant *Avon, Max Factor,* and *Revlon,* founded by Charles Revson. A number of these firms did not survive their founders. Elizabeth Arden was first bought by Eli Lilly and now belongs to the United States *Fabergé* groups; Helena Rubinstein has disappeared after being taken over by *Colgate Palmolive*; following 10 years of poor management and uncontrolled diversification, particularly in pharmaceuticals, Revlon has been grabbed by Pantry Pride, a chain store group belonging to the raider Perelman; Max Factor now belongs to the *Norton Simon* group.

While all this was taking place on the American scene, two groups, one Japan's *Shiseido*, and the other France's *L'Oréal*, were climbing to the rank of leading world producers, raised there by dint of good management and competent research and marketing skills. Although Shiseido was unsuccessful in its bid in the United States to take over the famous *Giorgio* of Beverly Hills, which was acquired by Avon, and the skin care company *Charles of the Ritz*, which Yves Saint-Laurent had sold back to Revlon, it nevertheless ranks second in the world after *L'Oréal* and has very strong positions throughout Asia.

L'Oréal's founder, Eugene Schueller, graduated as a chemist from Institut de Chimie de Paris. He resigned from his job at the Sorbonne to produce a "harmless" hair dye called l'Auréole. The trade name l'Oréal was adopted the following year. A skillful businessman and a true pioneer of ad campaigns, Schueller bought the Monsavon soap factory in 1928 and, before the war, brought on the market the

O'Cap hair lotion, then Ambre Solaire. When he died in 1957, his successors managed to develop the business both through internal growth and an efficient research base and through a series of acquisitions.

In 1961, Monsavon was sold to Procter & Gamble, and L'Oréal purchased the *Cadoricin* firm, which extended its range of hair products, to which were added *Garnier* and *Roja*. Then Lancôme was purchased, introducing high-class products. This was followed by the purchase of other perfumemakers. Tempted by the pharmaceuticals market, the company bought *Synthélabo* in 1973, consisting of four medium-sized laboratories. It is still too early to say whether the money sunk into the sector since then will bring in returns as large as those of the perfumes and skincare business. Mixing the two has not always been successful.

In the United States the marriages between Pfizer and Coty, Colgate and the US Vitamin, Eli Lilly and Elizabeth Arden, Squibb and Charles of the Ritz, Avon and Mallinckrodt, Revlon and Armour Pharmaceuticals all ended in divorce. There was, of course, Bristol-Myers' successful venture with Clairol, and American Cyanamid with its Shulton subsidiary. But these exceptions only confirm the general rule of failure.

In France, while Sanofi can draw satisfaction from its association with Yves Rocher, which enjoys great management freedom, the sector comprising Roger & Gallet, Stendhal, and Charles Jourdan has not yet lived up to the parent company's expectations. Only the British seem to have succeeded in combining such different businesses, possibly because from the start the skincare activities were intimately associated with pharmaceuticals within large groups like Beecham, Glaxo, and BDH.

The Chemistry of Additives

Used in small doses to improve the products in which they are incorporated, additives are to be regarded as specialties with well-stated functions even if, in many cases, they are well-defined chemical entities sold according to specifications. Because of this ambivalence, chemical companies have approached the sector of additives sometimes through the markets they serve, sometimes through the chemicals from which they derive, even at times from both ends.

Additives for Plastics

Additives for plastics have experienced the double approach. The opening up of the markets leading to uniform production of plastics gave worldwide scope to some additive producers. Tasks were shared since polymer producers did not consider it useful to prepare the additives they needed, while additive producers were, as far as possible, careful to avoid competing with their clients in the area of base thermoplastics.

It is true that a major polyolefin producer like *Hoechst* sells its own range of antioxidants and its subsidiary *Riedel de Haen* produces ultraviolet ray absorbers. Likewise, the world ABS leader *Borg Warner*, now acquired by General Electric, has been marketing, since it took over *Weston*, a series of organic phosphites for the stabilization of high polymers. In Japan, *Sumitomo Chemical* is a supplier of large-volume plastics as well as of a rather complete range of stabilizers.

These are exceptions, however. The world's largest additives producer for plastics, *Ciba-Geigy*, remains, for its part, at the sole service of its downstream customers and tries not to appear as a competitor. This is also the position of other additive suppliers like *American Cyanamid, Ferro, Witco, UniRoyal Chemical* in the United States, *AKZO, SFOS (Rhône-Poulenc) in Europe, Adeka Argus,* and *Dai-ichi Kogyo Seiyaku* (DKS) in Japan.

Ciba-Geigy owes its leading position to a number of factors: long perseverance in the specialty, an efficient research base through which the universally used *Irganox* antioxidants were developed, application services adapted to all the polymers requiring stabilization, worldwide production units established within large consumption areas (Europe, America, and Japan). Even where Ciba-Geigy did not invent a product but took a license on it as with *HALS* (hindered-amine light stabilizers), licensed from Japan's *Sankyo*, it developed it to the point of acquiring world supremacy in the area. Ciba-Geigy's success in this activity is all the more remarkable as it has no upstream integration on raw materials used in the synthesis of phenol antioxidants, of phosphites, of thioesters, of substituted benzophenones, of benzotriazoles, or of HALS. But this apparent weakness is fully compensated by the dominant position Ciba-Geigy has acquired in the different types of additives for plastics in its range, either through internal growth by its research, or through license acquisitions (Sankyo), or through purchase of relevant companies (Chimosa in Italy), or again through complementary activities (range of Goodrich's Goodrite antioxidants).

The other producers of plastics additives trail far behind Ciba-Geigy in variety of range or in market coverage. *UniRoyal Chemical* produces antioxidants and blowing agents and has production units in the United States, Latin America, Italy, and Taiwan, but its recent restructuring has cut short its development. *American Cyanamid* which pioneered a number of additives (substituted benzophenones, 2246) sold its European business to Ciba-Geigy in 1982 and now operates only in the American market. *AKZO*'s range is restricted to antistatic agents, PVC stabilizers, and peroxide catalysts, which it acquired through *Armak, Interstab*, and *Noury van der Lande*. *Borg Warner* is mainly focused on phosphites, which it produces solely in the United States; Elf Aquitaine's subsidiary *M & T* is focused on organotins; *Witco*, through its purchase of *Argus Chemical* and *Humko Products*, is involved in heat stabilizers, antistatic agents, and lubricants. *Ferro*, which also produces master batches, has developed specialties such as fireproofing agents and stabilizers for PVC and has recently joined forces with Italy's Enichem to produce and market new lines of polymer additives in the United States.

Other companies came to additives through the chemical tree, such as *Société Française d'Organo-Synthèse* (SFOS), a subsidiary of Rhône-Poulenc, which by isobutylating phenols produces a whole range of phenolic antioxidants as well as special phosphites, or *Ethyl*, which approached the Irganox family of antioxidants and bromine fireproofing agents through its orthoalkylation technology and its access to bromine. Similarly, because the US-based *Olin* was an important hydrazine manufacturer, it became interested in blowing agents like azodicarbonamide and bought up National Polychemicals, which also provided it with a range of phosphites. The blowing agent line has since been sold to UniRoyal Chemicals.

But the interest which major firms like ICI, Bayer, or even Hoechst still have in the sector is restricted by the small number of additives they supply to plastic producers. Under the circumstances, it is more than likely that Ciba-Geigy's lead in the variety of products offered, in research, in customer service, or in geographic coverage will be hard for competitors to catch up with. Indeed, their narrow approach to the market would hardly warrant the heavy investment to fulfill any high ambitions they might have in the area.

But favored by their access to certain raw materials or by their specialization in a very specific range, such competitors can, at least, be assured of a degree of prosperity inasmuch as the standards required for optimum use of plastics are closely related to incorporation in the high polymer of effective additives at a reasonable cost.

Rubber Additives

The specialists in *rubber additives* are distinctly different from the specialists in additives for plastics, even though the same products are sometimes used in both industries: blowing agents (azodicarboamide), phenol antioxidants (BHT, 2246), phosphites (tris-nonylphenyl phosphite). In the first place, additives for elastomers, unlike those which might come into contact with foodstuffs, do not require official approval, which makes it easier to put them on the market. In the second place, most of them are well-known products sold to specifications by a number of producers. The development of new products protected by patents is rarer than in the case of additives for plastics. In fact, while consumption of plastics has been constantly increasing, stagnant demand for both natural and synthetic rubber has not warranted any significant recent research efforts by suppliers of this industry.

If one considers that the automobile sector accounts for 75% of rubber consumption in developed countries, it stands to reason that the longlife radial frame and smaller diameter tires of modern vehicles should require smaller amounts of rubber for the same number of cars produced; in the United States alone, rubber consumption has fallen from 3.2 million tons in 1977 to a little over 2.6 million tons in 1989. This implies a consumption of some 150,000 tons of organic additives.

Faced with such a situation, producers of additives for rubber have either restructured or else rationalized production. In rarer cases, others have offered new products with higher added value than the conventional additives. In the United States, *American Cyanamid* in 1982 halted production in Bound Brook, New Jersey, of its accelerators; *Goodyear* terminated its substituted *p*-phenylenediamine production in Houston in 1984. In 1985, *Allied-Signal* took over *UOP*'s antiozonant unit, while in 1986 *UniRoyal Chemical* became part of *Avery, Inc*., before becoming the object of a leveraged management buyout in 1989. In Europe, Rhône-Poulenc and ICI merged their rubber divisions within a subsidiary called *Vulnax* and then finally sold it to AKZO in 1987. Atochem, meanwhile, was taking a minority stake in *Manufacture Landaise de Produits Chimiques*, henceforward leaving France and Britain with no significant producer with international clout.

In this changed environment, three major additives manufacturers emerged: *Monsanto* with its plants in the United States, Canada, Britain, and Belgium; *Bayer*, which owns two sites in Europe and produces antiozonants in Pittsburgh through Mobay; and *UniRoyal Chemical*, which has production units in Naugatuck, Connecticut, and Geismar, Louisiana, as well as in Canada, Brazil, and Italy.

UniRoyal Chemical was separated in 1966 from the US Rubber, which had provided it with a captive market. But two other tire manufacturers had retained their traditional activities in additives. They were *Goodrich*, which produced only in the United States, and *Goodyear*, which also operates in Europe in its antioxidant units in Le Havre, France. Both these giants sell part of their production through a rubber blend specialist, R.T. Vanderbilt. Goodrich, however, has recently withdrawn from the tire business in order to concentrate on its chemical activities, so that only Goodyear enjoys today the advantage of a captive outlet for the rubber chemicals it produces.

Although Monsanto can rely on only two of its own raw materials, tertbutylamine and *p*-nitrochlorobenzene, for its range of additives, it is regarded as an efficient producer and a pioneer in antiozonants based on *p*-phenylenediamine and prevulcanization inhibitors. It has one of the most complete ranges of additives for rubber and the most modern units to manufacture them.

Because of its long experience in organic synthesis intermediates, *Bayer* is possibly better integrated upstream than Monsanto. Its range of products is just as large, but its production units are essentially restricted to Leverkusen and Antwerp. With the exception of *AKZO, which, through its purchase of Vulnax,* seems to want to improve its range of additives and its geographic coverage, no other major European chemical group has gone beyond a small range of special products.

Like their competitors in Europe and in America, the Japanese producers have focused their attention on accelerators (vulcanization activators and agents) and on antiaging agents (antiozonants, stabilizers). Their automobile exports provide a market for tires that their counterparts in other countries cannot claim to the same extent.

Japan's additives production, however, is too scattered among a large number of producers to be truly profitable. With the exception of two principal companies in the area, *Sumitomo Chemicals* and *Mitsubishi-Monsanto*, firms like *Ouchi, Shinko, Kawaguchi*, and *Seiko Chemical*, which were the first to get into the business in 1930, do not have the required size to be competitive on international markets.

Additives for Lubricants

Additives for lubricants are also greatly dependent on the automobile industry, which alone uses some 60% of the lubricants produced worldwide, whether lube oils for engines (gasoline and diesel oil) or for gear boxes. Since the oil-price rise in 1973, lube oil consumption has been affected by a number of factors: smaller vehicles and therefore smaller engines, falling automobile production, larger intervals between oil changes, implying a higher additives dosage to extend oil efficiency. To these various changes should be added increasing use of diesel fuel in Europe because of favorable taxation. The generalized use of multigrade oils and the introduction of unleaded gasoline, and consequently of catalytic exhaust pipes, should lead to enhanced engine oils. In the circumstances, world consumption of additives for lubricants is likely to remain at around two million tons a year over the next few years, with higher additive doses compensated by extended lube oil efficiency and smaller casing size.

With the exception of *Lubrizol*, the world leader in this area, and *Ethyl,* which came to lube oil additives by buying Edwin Cooper off Burmah Oil in 1968, the main suppliers with extensive ranges of additives are the international oil companies *Exxon*, *Chevron*, *Amoco*, and *Shell*. The business was a natural extension of their lube oil production, which serves as a captive market.

All these oil companies market their additives as a package, the efficiency of which has been extensively tested. Most of the ingredients in the package are produced by the companies themselves: *detergents* (sulfonates, phenates, naphthenates), *dispersants* (succinimides, polybutene, succinates), *antiwear agents* (zinc dithiophosphates, chlorinated paraffins, sulfur, and phosphate hydrocarbons), *anticorrosion agents* (substituted amines, succinic acid derivatives, nitrites). On the other hand, the antioxidants are often supplied separately, as are pour-point depressants (polymethacrylates, PMA) and additives to improve the viscosity index of multigrade oils (PMA, olefin copolymers).

Originally called Graphite Oil Products, *Lubrizol* was founded in 1928 near Cleveland with a capital of $25,000 by six associates. Their close ties with the Case Institute of Technology gave the concern a strong technical orientation. In this manner, *Lubrizol* played a pioneer role in developing lube oil additives and is still today a world leader in its area, with 14 plants installed worldwide, including four in the United States, and testing sites in Wickliffe (Ohio), Hazlewood (Britain), and Atsugi (Japan). In its attempts at diversification, the company recently became interested in biotechnology with the purchase of *Agrigenetics* in 1985 and of a stake in *Genentech*. But it is too early to state whether this choice will bring the same long-term satisfactions as the company's traditional business.

Exxon came to chemical additives for lubricants by producing its Paraflow range of freezing-point depressants as early as 1930 in Bayonne, New Jersey. Through the agreements signed in 1937 with IG Farben, Standard Oil of New Jersey (later to become Esso and then Exxon) acquired the thickeners and additives based on polyisobutylene that improve the oil viscosity index. In the 1960s, Exxon further enlarged its range of lubricant additives and in 1979 set up the Paramins special division, which marketed a series of olefin copolymers (OCP) based on the chemistry of the group's ethylene-propylene elastomers. The object was to compete with the PMA in improving multigrade oils (VI improvers). Based in Houston, Texas, and involved in all world markets, Paramins has become Lubrizol's most dangerous rival.

Chevron approached the oil additives market in the 1930s by supplying metal naphthenates to its parent company Standard Oil of California. Some of these additives were marketed under the trademark Oronite from 1948 onwards. Chevron kept its main research center in Richmond, California, even when, in the 1950s, it spread to international markets through subsidiaries set up with local partners: *Orobis* with BP and *Orogil* with Progil in Europe, *Karonite* in Japan, *AMSA* in Mexico. In 1986, BP bought Chevron's 50% share in Orobis, and more recently Rhône-Poulenc sold to Chevron its 50% share in Orogil.

The interest shown by *Ethyl Corporation* in Edwin Cooper stems from its desire to diversify into the oil sector as unleaded gasoline begins to threaten the future of tetraethyl lead. But in a business in which it is a newcomer, Ethyl still has much to learn before attaining the efficiency and international coverage of its three main rivals. The same is true of *Amoco* and of *Texaco Chemicals*, although they are endowed with a significant captive market through their parent companies, Standard Oil of Indiana and the group made up of Texaco, Caltex, and Getty Oil, respectively.

For its part, the *Royal Dutch Shell* group came to additives after the Second World War in the United States with a range of alkaline sulfonates. Subsequently it enlarged its range with new additives (detergents, dispersants, VI improvers) and fuller geographic coverage through production centers located in Berre, France; in Stanlow, England; and in Marietta, Ohio, and in Martinez, California, in the United States. More recently, a common subsidiary with Lubrizol was set up in Brazil.

Besides these large companies, which offer a range of additives as extensive as possible, if only to recoup research expenses and the high cost of tests required to obtain approval of the "packages," there are a number of chemical companies that have also established a foothold in the market of lube oil specialties. Their reason for doing so was that they had acquired know-how in the chemical sector leading to the products marketed.

Rohm & Haas in Philadelphia developed additives to lower the freezing points of oils and to improve their viscosity index through work carried out as early as 1934 by the chemist Herman Bruson on the properties of PMA produced from higher methacrylates. Other companies, such as *Röhm* in Darmstadt and Melle-Bezons (whose Persan unit in France was bought from Rhône-Poulenc by *Société Française d'Organo-Synthèse* (SFOS) in 1978), also supplied PMA for such applications. Through the chemistry of phenol isobutylation, *Ethyl* and *SFOS* took a foothold in the phenol antioxidant market of oil companies, while *Ciba-Geigy* is developing a significant program in this area. But the need to be thoroughly acquainted with the lube oil business and to be well introduced in the world oil circles narrows the scope of chemical firms that have only a small range of additives to offer and precludes their taking a significant place in such a specialized market.

Food Additives

Because their nature, their uses, and their origins are extremely varied, *food additives* are supplied by a large number of different firms. In what is a fragmented industry, some producers stand out more because of the major place they occupy on the market than because of their range of additives.

In the United States, there are only two producers of *citric acid* (Pfizer and Bayer's subsidiary, Miles) and of *vitamin C* (Pfizer and Hoffmann-La Roche) and a single producer of *saccharin* (PMC, which bought Maumee from Sherwin Williams), *sorbates* (Monsanto), and *carrageenates* (FMC since it acquired Marine Colloids).

Because of the very strict rules that in industrialized nations govern additives used in human food, it has become very expensive to introduce new products. In some countries even some of the older derivatives that used to be considered nontoxic have been questioned. This is the case with saccharin, discovered by Ira Remsen in 1879 and used without drawbacks since then. Because of such limitations, few new producers have ventured into the area over the past few years except through purchase of existing companies that already had approved additives. The giants in the business are generally satisfied with being dominant in certain market sectors through their special technologies (fermentation, extraction, synthesis). The problems of excessively high sugar consumption, however, have induced a number of researchers to look for low-calorie substitutes for sucrose other than saccharin. Accordingly, new *synthetic sweeteners* have been discovered: *cyclamate* (sodium cyclohexyl-sulfamate), synthesized in 1937 and put on the market by Abbott in 1950; *aspartame*, isolated in 1965, produced by reaction of aspartic acid with phenylalanine methyl ester, and developed by Searle, which was subsequently purchased by Monsanto; and Hoechst's *acesulfame K*. Despite lack of coordination in this area among the different national legislatures, these synthetic sweeteners, with their low calorie content and a sweetening power that is 50–200 times as great as that of sugar, should sooner or later take root on international markets.

The use of *gelling* and *thickening* agents in foodstuffs goes back to earliest times. In the last few years, progress has been made in the extraction and purification of plant-based hydrocolloids used for the purpose. In addition, the polysaccharide *xanthane*, produced through fermentation, has been developed over the past 20 years to take its place among the water-soluble gums supplied to the food industry. At the same time, a semisynthetic gum, *carboxymethyl cellulose* (CMC), used in a number of industrial applications, was allowed in its purified form, in human foodstuffs.

The US-based *Hercules*, which started by producing precisely this CMC of which it is the world's leading producer, has gradually extended its range of products by purchasing companies. It is now involved in *pectin*, extracted in Denmark and in the United States from lemon peel; *guar*, prepared in Italy from a bush that grows in India and Pakistan; *carob* developed in Spain; and *carrageenates* extracted from algae growing along the Atlantic and Pacific coastlines. With the exception of pectin, these various gums recently became the business of *Aqualon*, a common subsidiary of Hercules and *Henkel*, its German partner, already involved in the guar and CMC market. In 1989 Hercules became the sole owner of Aqualon.

Hercules never did succeed in developing *Xanthane* through its association with the British company Tate & Lyle in 1979. This gum has remained a specialty of *Rhône-Poulenc* which produces it in Melle, France, and of *Kelco*, a San Diego, California, subsidiary of Merck that also owns *Alginates Industries*.

Other groups have likewise specialized in particular sectors. The Stein Hall subsidiary of Celanese, taken over by the British-based *RTZ Chemicals*, now part of Rhône-Poulenc, has focused on guar, while *Marine Colloids*, a subsidiary of FMC, and *Satia*, of the Sanofi Elf Bio Industries group, specialized in carrageenates. There are a great number of industrial applications for gum, and thus gum producers are usually drawn to the food industry because of their know-how in gum. It is seldom that they have chosen to manufacture thickeners because of their experience in foodstuffs.

The same is true of *antioxidants* like *BHT* (butylhydroxytoluene). Although it is used in purified form in human and animal food, its more common use is as a stabilizer for polymers and lubricants. Only *BHA* (butylhydroxyanisole), α-*tocopherol* (vitamin E), *TBHQ* (tertiary-butylhydroquinone), and *propylgallate*, which are marketed by *Eastman Kodak*, can be considered as purely food antioxidants for the two reasons that they are not toxic and that they are high priced. In fact, Eastman Kodak is the only chemical leader to produce an extended range of food additives: mono- and diglycerides and vitamins.

Producers of *acidulants* came to food applications through chemistry or biochemistry. *Malic acid* is produced by Denka, now owned by Mobay, in the United States and by Croda in England. Like *fumaric acid*, it is a derivative of maleic anhydride production. The major acidulant is *citric acid*, which is also used as a stabilizer. It is a fermentation product that is produced by a few traditional specialists—in Europe by La Citrique Belge, which was bought by Hoffmann-La Roche; in Britain by Sturges, taken over by RTZ Chemicals and now owned by Rhône-Poulenc; and in the United States by Pfizer and Miles.

Phosphoric acid, used in fizzy drinks, is produced in its food quality only by a small number of firms such as FMC and Stauffer, an activity taken over by Rhône-Poulenc in the United States, and by Prayon in Belgium.

Food conservation generally requires the use of chemical additives, although the problem can be solved at times through temperature control (pasteurization or sterilization through heating, freezing, or control of water content [dehydration]). Chemical additives act by working on the metabolism of the microorganisms responsible for food deterioration. More often than not they involve organic acids and their salts, *propionic acid, potassium sorbate, sodium and calcium propionates*, and *sodium benzoate*, traditionally used to preserve cheese, jam, cakes, and fatty materials.

Here again a few large companies such as Monsanto for sorbates and Pfizer for propionates have acquired a leading place on the markets. On the whole, the food additives sector is less open to restructuring and rationalization because it is made up of enterprises that are fundamentally different in size, technologies and in objectives pursued.

Photochemicals

Since the early 1980s, the major photographic companies have made efforts to bring changes to their basic technologies, which had long remained unchanged. The US-based *Eastman Kodak* became interested in reprography, setting up its own range of photocopying machines. It also became involved in electronics and video to counter competition from new Japanese equipment (such as Sony's Mavica filmless cine-camera). The other photographic giants like Bayer's subsidiary *Agfa-Gevaert* and *Fuji Photo Film* have also invested heavily in new areas, the former in magnetic tape and reprography and the latter in photo disks. *Polaroid*, whose founder Edwin Land remained to the day of his retirement an advocate of specialization, is also starting to put a range of videocassettes on the market. These changes, however, are essentially intended for the amateur and mainly concern camera manufacturers. Overall, the sensitive surfaces market should receive no shakeup from these new ventures, for there is still a high demand in a number of areas where photography remains irreplaceable (press and publishing, scientific research, industrial applications, radiographic control devices). The industry's structure reflects this stability. It is not likely to be upset in the immediate future because of the power acquired by the few large multinationals, which vie with one another on international markets and give any newcomer little chance of success.

Unable to compete with Kodak on the American market, *GAF* withdrew from the film industry in 1982. Previously, the first European merger had taken place between Belgium's Gevaert and Germany's AGFA, producing Agfa-Gevaert. Its early years were hard ones, and it is now fully owned by Bayer. Italy's *Ferrania* was taken over by America's *3M*, while Ciba-Geigy was bringing together Britain's *Ilford*, France's *Lumière,* and Switzerland's *Telko* within the Ilford group based in Britain.

Following these restructurings, which, in many cases, took place some time ago, the photographic film industry is now dominated by three giants: *Eastman Kodak*, with units in Rochester, New York, in the United States, in Châlons, France, and in Hemel Hempstead, England; *Agfa-Gevaert*, which produces its photochemicals in Antwerp, Belgium, and Vaihingen, West Germany; and *Fuji Photo Film*, which produces in Japan and has recently set up a film unit in Holland. *Ilford* and *Polaroid*, which went through difficult periods of readaptation; *3M*, which is involved in other areas besides photography; and *Konishiroku* (*Konika*) in Japan, which bought Fotomat in the United States, cannot be regarded as dangerous rivals to the Big Three.

The three major companies follow different policies in matters of raw materials. Fuji Photo film, which has no links with the chemical industry, buys 80% of its supplies outside, while Eastman Kodak and Agfa-Gevaert supply half their needs through their own production. They all produce their most "sensitive" organic derivatives, which are kept secret since they form the basis of emulsion quality.

Although polyester film, introduced by Du Pont under the trade name Mylar in the 1960s, has been added to the traditional supports like paper and cellulose acetate, the principle of photographic film preparation has remained unchanged since "daguerreotype" was developed. The sensitive surface always contains a silver halide crystal emulsion with a *gelatin* binder. Despite all the efforts to replace them, silver salts remain the basis of these emulsions, and film manufacturers still require gelatin, which they consume at the rate of 20,000 tons a year. The suppliers are few, and they are carefully selected. The world leader in this area is *Rousselot*, now a subsidiary of Sanofi Elf Aquitaine, with four units in Europe and one in the United States.

Reducing agents such as *hydroquinone, metol* (*p*-methylaminophenol), and *p-phenylenediamine* are generally purchased from outside producers. Eastman Kodak produces its own hydroquinone, however. The other producers get their supplies from Rhône-Poulenc or from Japanese firms like Sumitomo Chemicals or Mitsui Petrochemicals.

Color photography, now fully perfected, requires a *developer* like *N,N′*-diethylphenylenediamine which reacts with silver salts. The oxidized derivative obtained reacts with a coupling agent made up of groups (–CH=) or ($–CH_2–$) to produce the desired color.

Formulations for sensitive surface emulsions also include *accelerators* (alkaline carbonates, borax), *stabilizers* (sodium bromide, benzotriazoles), *conservation agents* (sodium sulfite), *hardeners*, which improve gelatin behavior (chloromucic acid, substituted 2,4-dichlorotriazines).

The great variety of products used by film manufacturers, their stringent quality standards, and their secretiveness, which prevents them from subcontracting their most advanced formulations, are all factors that keep photochemical producers apart from the rest of the chemical industry and keep newcomers out of their sector.

The Alliance of Chemicals and Electronics

Chemical products used today in electronics seem, at first glance, to be a very ordinary kind. They are different from those generally offered, however, by reason of the extraordinary degree of purity which their producers must achieve in order to satisfy the stringent requirements of the electronics industry. The maximum dose of impurities tolerated in monocrystalline silicon amounts to one part in 10^{13}.

Polycrystalline silicon, produced from silane (SiH_4) or trichlorosilane ($SiHCl_3$) forms the upstream part of the *semiconductor sector*. Monocrystalline silicon is extracted from polycrystalline silicon and sliced into wafers 25 μm thick and 8–10 cm in diameter.

Hoechst's subsidiary Wacker is the world's leading polycrystalline silicon producer, with a capacity exceeding 2,000 tons. The overcapacities that began affecting the electronics industry in the early 1980s forced Monsanto, one of the largest wafer producers, to slow down its silicon production units in 1984. It has since sold this business to Germany's Hüls. Rhône-Poulenc, which had ambitions in the area but lacked the right technology, has withdrawn from the business.

There are enough suppliers of this type of silicon, including, for instance, Dow-Corning, Dynamit Nobel, Shin-Etsu, Tokuyama Soda, Motorola, and Texas Instruments. A possible substitute for the silicon used to produce wafers is *gallium arsenide*, in which Rhône-Poulenc, ICI, and Shinetsu are already involved.

This situation shows how closely suppliers of electronic chemicals need to monitor the very rapid developments taking place in the area; otherwise, their productions run the risk of becoming obsolete before the full payoff.

Photosensitive products are also used for the production of wafers. These *photoresists* polymerize through X-ray treatment. They are called positive or negative according to whether or not they are soluble in solvents when exposed to light. The miniaturization of printed circuits tends to give a boost to positive resins. Germany's Hoechst has pioneered in such photosensitive resins. They are also supplied by Eastman Kodak, Olin Hunt, Ciba-Geigy, E. Merck, and Tokyo Ohka Kogyo.

A great number of chemical firms have set up special divisions to manufacture products for the electronics industry, essentially through acquisitions. For example, *Du Pont* bought *Berg Electronics* in 1972, and a little later *Olin* purchased *Philip A. Hunt*. Some companies, such as Du Pont, Olin, and Ciba-Geigy, have chosen an "integrated systems" approach in this area by providing as wide a range as possible of products and services for the electronics industry. Others have elected to remain strictly within the special areas in which they excel through long experience or proper chemical integration. Thus it was the work carried out before the war with AEG that led *BASF* to make its range of magnetic tapes and gave it the supremacy in chromates which it shares with Du Pont. *Hoechst* came to silicon through Wacker and to gases through Messer Griesheim, and now provides, besides high-purity special gases, a range of photosensitive polymers. Rhône-Poulenc became involved in printed circuits through its polyimide resins and Ciba-Geigy through its epoxy resins.

Most of the companies already producing diethylene glycol terephthalate polymers have launched into the applications of polyester film to video and data processing, Hoechst through its Kalle subsidiary, ICI, Rhône-Poulenc, Du Pont, Japan's Toray, Teijin, and Toyobo, the latter in association with Rhône-Poulenc in Nippon Magphane.

Although Rhône-Poulenc has given up direct upstream development after fruitless association with Dysan in magnetic supports and Siltec in silicon, it still believes it can use its know-how in rare earths to develop their electronics applications. Today, Rhône-Poulenc is the indisputable leader in rare earths, accounting for 40% of the world market. At its units in La Rochelle, France, and Freeport, Texas, it is capable of extracting from lanthanide sands the 14 elements they contain. Over the last few years, *samarium*, for instance, has become essential for microelectronics to the same degree that *europium* and *yttrium oxides* already are for color television.

Whether they approach electronics directly, or through chemicals, or both, chemical companies involved in this business can hope to reap the fruits of their efforts in this area, providing, however, that the sector is spared the technological and economic jolts it has suffered over the past 10 years.

Catalysts

Ever since England's Humphry Davy observed in the early 1800s that water was formed when hydrogen and oxygen react in the presence of a red-hot platinum wire, the phenomenon which Berzelius was to call catalysis has intrigued chemists. The uses of catalysts in industry were first consciously demonstrated by Peregrine Phillips in 1832 when he used platinum to oxidize sulfur dioxide (SO_2) to form sulfur trioxide (SO_3) and by Frédéric Kuhlmann in 1837, when he produced nitric acid from ammonia.

Early in the twentieth century, Germany's Wilhelm Ostwald, France's Paul Sabatier, and America's Irving Langmuir had advanced a step in interpreting the phenomenon of catalysis by showing that it was characterized by an acceleration of the rate of reactions and that it was conditioned by the state of the catalyst's surface. From then on, chemical technology made striking progress through use of catalysts. Between 1905 and 1920, and more particularly in Germany, there was a spurt of new industrial-scale processes, for example, Fischer–Tropsch synthesis and BASF's use of vanadium oxide to produce sulfuric acid.

It is no exaggeration to say that without catalysts Germany would have been in no condition to pursue its war effort until November 1918. Likewise, if Houdry had not developed in the early days of World War II its "catalytic cracking" process, the United States would have found it very hard to provide its bombers with light fuel. It was also through catalytic reforming that the United States managed to obtain from petroleum the toluene needed to produce TNT between 1941 and 1945.

Since then, catalysts have played an essential role, particularly in the production of ethylene oxide from ethylene (Shell, Scientific Design), in the synthesis of hydrogen cyanide and acrylonitrile through ammoxidation (oxidation in the presence of ammonia), of formaldehyde (from oxidation of methanol), and, of course, in the polymerization reactions to produce plastics, elastomers, and synthetic fibers. It is not surprising, under the circumstances, that a catalyst industry should have developed after World War II through internal growth or through acquisitions. The very diversity of catalysts and of their uses has necessarily led to a fragmented sector.

Some oil companies became involved in the production of catalysts because they needed them in their own refineries. *Mobil* has developed the ZSM 5 catalyst based on zeolite following studies which began as early as 1936 on catalytic cracking; *Shell* has used its own technology to develop the sales of its catalysts for hydrogenation cracking. Other companies became involved in catalysts because of their precious metals business. *Johnson Matthey, Engelhard*, and *Degussa* applied their know-how in platinum metals to industrial catalysts. Chemical firms, for their part, approached the area in different ways. *ICI* made the most of its acquired know-how, particularly in methanol and ammonia, by associating with *Nalco* to form *Katalco*, a catalyst supplier; *American Cyanamid* has set up a subsidiary in Holland with Ketjen; *Rhône-Poulenc* has formed *Procatalyse* in joint venture with Institut Français du Pétrole.

In other cases, the involvement in catalysts has been through acquisitions. *W. R. Grace* bought *Davison Chemical* in 1953, and in 1984 Union Carbide purchased *Katalistics International*, BV. One of the three leading United States companies in cracking catalysts, together with Engelhard and Davison, is *Harshaw-Filtrol*, which is the result of the merger of subsidiaries of Gulf Oil and Kaiser Aluminum & Chemical that have specialized in the area.

The developers of new processes have found it at times more expedient to set up their own separate entities to supply the catalysts they were advocating. Allied-Signal's subsidiary *UOP* did so for its platforming; *Houdry* for its catalytic cracking; *Ralph Landau* for the silver catalyst used for direct ethylene oxidation, which was marketed by *Halcon SD* and subsequently taken over by Denka, then by Bayer; and

Phillips Chemical for its polyolefin catalysts, sold through its subsidiary, Catalyst Resources.

Through inert supports, a number of firms have succeeded in creating a niche in catalysts—for instance, *Crosfield*, a subsidiary of Unilever in England and a silica producer; or the German *Süدchemie* group, which specializes in hydrogenation and polymerization catalysts; or again *Condea*, which produces in West Germany alumina of high purity. The sector also includes a few firms which are only involved in a very special sector. Denmark's *Haldor Topsoe* makes catalysts for the synthesis of ammonia and methanol; *Lithium Company of America*, an FMC subsidiary, produces lithium, while *Du Pont* makes boron derivatives.

Linked to the oil industry, to petrochemicals, and to the large commodity chemicals, the catalyst industry can hardly escape the economic ups and downs affecting these three large sectors. Its clients are understandably both demanding and prudent, for the catalytic system is basic to the good running of production units. This explains why it is an area of business that is so difficult to penetrate and run profitably. Its structure should therefore remain rather stable even with the development of catalytic exhaust systems. Introduction in the United States and in Europe of unleaded gasoline and the use of bimetallic systems for catalytic reforming should open up new markets for platinum and rhodium.

Retrospect and Prospect

The economic slump that started in 1973 when OPEC pushed up crude oil prices challenged what were until then regarded as indisputable truths.

First came the realization that just as no tree can climb as high as the sky, so *no growth can be guaranteed to be continuous*. Suddenly investments made at a time of high inflation and low interest rates turned out to be disastrous as demand slowed down simultaneously with monetary erosion.

The scale effect, which until then was assumed to be cost-saving, showed its weaknesses as the giant steam crackers proved more expensive to run at low capacity than smaller units already written off and working at full capacity.

The notion that production costs could be improved by grabbing a greater share of the market turned out to be fatal as the gain in sales was wiped out by severe price erosion.

Likewise the assumption that the fruits of research would be proportional to the funds devoted to the sector was totally invalidated, for never had the world's chemical industry spent so much money in research and development to so little avail. At the same time, the venture capital poured into biotechnology companies has yet to bring in the returns expected. The managers of chemical plants, wary of world petrochemical and heavy chemical overcapacities, believed they would find in a switch to specialties at least partial compensation for the losses incurred through traditional productions. Although they were not all disappointed in their hopes, some of them found that results obtained fell short of expectations, for until then specialties had been the special field of firms that had acquired experience in what were specific and as yet uncrowded sectors.

Manufactured by too many producers, some specialties were becoming commonplace. For a manufacturer, there are only two kinds of products, those that make money and those that do not. The profits that can be made on a sale are closely related to the number of producers on the market and to the day-to-day relationship between supply and demand. Whether a product is deemed a "commodity" or a "specialty," it is all the more profitable for its being offered by a smaller number of producers for a demand that remains unchanged. In this context, there are some pharmaceutically active materials protected by patents and some secret formulations that are genuine profit centers for their producers. Likewise, should a base product become scarce on the market because of an accident on a petrochemical site or because of sudden high demand, prices soar and the fortunate producer can turn out the product to maximum capacity and profit.

Over the last few years, the high cost of installations and of the money needed to finance them was not conducive to the building of new plants on any large scale in industrialized countries. But as demand trends have been moving upwards lately, *petrochemicals* have at long last *returned to profitability*.

The specialties rush of chemical leaders is, on the contrary, more likely than not to produce a surfeit of products, at least insofar as some specialties are concerned. These will shed their "added value" and consequently lose their attraction for the too numerous industry leaders that had decided to follow that path.

Other disappointments are likely to come from the organizational and managerial differences between a purchaser and the specialties firm acquired. The many divestments that have often followed upon hasty acquisitions show how difficult it is to force on an entrepreneurial company the management methods of a large multinational.

One of the paradoxes of the last few years has precisely been that specialties suitable for medium-sized firms capable of being flexible in their approach to daily matters should have fallen into the hands of chemical giants with necessarily heavier structures, while in the United States, for instance, through the "LBO" procedure, a few strong-minded individuals have succeeded in taking over large petrochemical and thermoplastic production units considered until then as the rightful field of the industry's greats.

It is not certain whether the errors of the past will be repeated in the future. The thirst for power could indeed lead some company heads to overinvest, especially if they have public funds at their disposal. They would then recreate the overcapacities that have been so harmful to fertilizer, petrochemicals, synthetic fibers, and plastic producers over the last few years. It is also likely that specialties will continue to attract industry leaders anxious to develop fresh prospects.

Let us hope that all the decision makers will bear in mind that capital funds, whether provided involuntarily by the taxpayer or willingly by the shareholder, are a rare resource that must be judiciously allocated and that success in all things comes from mastery acquired through long patience. In this respect, Germany's chemical industry, which has shown continuity from the time it was established in the last century to the present under the guidance of professionals, is a tried and tested model, showing profits even in the most adverse circumstances.

Drawing inspiration from this example for long years, the United States chemical industry, under the pressure of financial analysts and raiders, has in recent years undergone many upheavals. While they provided new opportunities for the fortunate few, they changed the environment and made people forget that to operate efficiently any industry must set its sights on the long term.

For reasons that were more political than financial, France's and Italy's chemical industries have also undergone too frequent changes over the past 20 years—in their structures, their strategies, and their management teams—to have had a chance of getting through the economic slump unscathed. It is only very recently that they have returned to profits by recovering a measure of stability.

Worldwide, in 1987–1989 the industry, whether in specialties or in basic chemicals, has certainly had its most prosperous years ever. The chemical industry, on the whole, does not, however, enjoy a very favorable image in the eyes of the public. The harmful spillovers caused by untoward accidents are given wider publicity by the media than the benefits the industry provides. In consequence, administrations that were anxious to soothe the more or less justified fears of their citizens have brought out a spate of regulations often more restraining and therefore more costly than is really necessary.

Since one cannot work simultaneously toward a thing and its opposite, no great spate of discoveries useful to humanity should be expected at a time when everything is being done to make it difficult to bring new products onto the market. For a long time the chemical industry was left free to apply its own safety standards and could devote most of its time to the development of new products. In the last few years, it has had to submit to increasingly costly and prolific rules and regulations that require its attention and delay the development of innovations that could save human lives or at least improve our living conditions. Some balance will have to be found between safety requirements and the wider interest of the public.

As in all history, the story of chemicals recalls past events and makes an attempt to explain them. But it can neither create them nor prevent them from recurring. While such history, therefore, teaches us the essential facts that have taken place within two richly endowed centuries, it does not tell us which major facts will form the threads of the next years. It is this unknown factor which makes up the spice of our professional life. We can at least hope that if we conform to reason, to ethics, and to scientific and economic laws for all that is within our scope, each of us will have served this wonderful science that is chemistry to the best of our capacities and in the interests of the greatest number of people.

The Period of the 1990s

The Chemical Industry under Pressure from Public Opinion and Regulatory Authorities

Prior to the mid-1980s the chemical industry experienced some upheavals of which the thalidomide tragedy of the early 1960s and the Bhopal catastrophe in 1984 are two major examples. However, the lessons learned from such sad events led to a spectacular improvement during the 1990s in the safety record of the industry in the Western world.

At the same time much progress has been made in the abatement of pollution in the air and the treatment of effluents from chemical operations in North America, Western Europe, and Japan. Between 1978 and 1988 the content of SO_2 in the atmosphere was reduced by 30% in the United States. Similarly, constant improvements in the way chemicals are manufactured have reduced the amount of by-products resulting from chemical operations and, therefore, of the quantities of effluents to be treated.

Paradoxically, as these improvements were brought about, the chemical industry in the Western world has become the preferred target of environmentalists, and through the influence of the media, its image has been deteriorating in the eyes of the public at large. The time is indeed long gone when a firm like Du Pont could print as a motto on its letter-head "better things for better living through chemistry."

Public opinion was just one area in which the chemical industry of the developed nations suffered setbacks. The industry had to face a more tangible threat in the form of increased pressure from regulatory authorities. For the United States alone, Edgar Woolard, then chairman of Du

Pont, cited a figure of $585 billion for the financial burden incurred by industry in 1993 as a result of federal regulations, and he predicted the figure to reach $660 billion in 2000 ("In Praise of Regulation Reform," *Chemistry & Industry*, 5 June 1995). At the same time the Environmental Protection Agency itself projected that by the end of the 1990s the United States would spend $160 billion per year on pollution control. In 1996 Ben Lieberman, an environmental research associate with the Competitive Enterprise Institute, estimated that in the United States the cost of the phase-out of CFCs in accordance with the 1987 Montreal Protocol on Substances That Deplete the Ozone Layer could reach $100 billion over the next 10 years. Indeed chemical manufacturers had to develop eco-friendly substitutes such as hydrochlorofluorocarbon (HCFC) and hydrofluorocarbon (HFC), which are more costly to make, and hundreds of millions of pieces of air-conditioning and refrigeration equipment using CFCs had to be discarded.

In fact the chemical industry has been affected in many different ways by the flurry of regulatory edicts in the 1990s. For example, the banning of CFCs, together with new restrictions on the use of chlorinated solvents, has forced chemical producers to steer their product mix of chlorinated hydrocarbons away from precursors of these two categories of chemicals. Similar disruptions in the "chemical tree" of derivatives have been caused by the phasing out of tetraethyl lead as an antiknocking additive for gasoline and by its replacement with methyl tertiary butyl ether, whose fate is now held in balance by the authorities in California.

The fear of dioxins, which was born from an accident occurring on July 10, 1976 at the unit in Seveso, Italy, of Givaudan, a subsidiary of the Swiss Hoffmann-Laroche, has also changed the way many chemical operations are conducted. That accident, although it caused no human fatalities, damaged the environment by releasing in the atmosphere some 500 g of a very noxious chlorinated impurity of the family of the dioxins. In order to avoid a release of even the most minute quantities of this particular dioxin again, hexachlorophene, a very useful germicide is no longer manufactured. In addition, pulp and paper mills in North America have been asked to drastically curtail the use of chlorine and chlorates for bleaching the pulp, and industrial and municipal waste incinerators are being submitted to stricter air-pollution controls, which make it necessary to install high-efficiency scrubbers.

The more stringent regulations enacted by Western governments have led, in turn, to a delocalization of chemical activities to places where such rules are less strictly enforced. Nations like India and China have thus become world leaders in the production of some fine chemical intermediates and dyestuffs. Similarly in Mexico many maquiladoras owe their success, in part, to this process of delocalization.

International trade in chemicals has been affected in many other ways by the vagaries of national legislations. For instance, meat treated with hormones is considered safe as food for Americans but not for the citizens of the European Union. Bovine somatropin (BST), a hormone that increases milk production in cows, has long been cleared by the US Food and Drug Administration, yet it is banned by health authorities in Brussels. Products derived from biotechnology have been particularly prone to these inconsistencies. Hoechst spent 7 years and 80 million DM to obtain permission to operate its artificial insulin unit in Frankfurt. Meanwhile one of Eli Lilly's plants located in Strasbourg, France, was regularly exporting an identical product to Germany.

While the use of genetically modified organisms in pharmaceuticals has finally been accepted because of their beneficial effects on human health, environmentalists, particularly in Western Europe, strongly object to the application of such organisms to agriculture. Transgenic crops, which offer increased resistance to herbicides and reduce the need for synthetic insecticides, have received the blessing of public authorities in the United States, Canada, and Argentina and have been planted with enthusiasm by farmers in these countries over the last 10 years. Even though no harmful effect resulting from their use has ever been detected during that period, Greenpeace and other environmental organizations, with public backing, were able to convince European officials that the planting of transgenic crops should be severely restricted until more is known about how they react with the environment. The European Union's sudden conversion to the "precautionary principle" will have important consequences. International trade of transgenic food and feed based on genetically modified corn or soy will be submitted to stricter regulations. Further development of disease- and pest-resistant seeds may take more time to benefit the farmers and, in turn, the final consumers. As these restrictions are being implemented, chemical companies specializing in life sciences will have to consider a drastic restructuring of their portfolio of products.

The Status of Science and Technology

At the end of the 1990s the chemical industry in its main activities had reached a stage of maturation with respect to innovation. Several factors can help explain this situation. First, the pace at which new discoveries were made between the 1930s and the 1960s was not sustainable for the same reason that new elements of the periodic table were not easily found once most of them had already been described. Second, the part of the turnover that chemical firms were able to devote to their research and development budgets became smaller as more of the available funds were used to cover increasing environmental expenditures. Third, the management of companies whose shares were quoted on the stock market had to pay closer attention to the wishes of stockholders and financial analysts, who were often more interested in short-term accomplishments than in ambitious R&D programs, which are necessarily lengthy, costly, and risky.

However, while no major breakthrough was made during the last 10 years of the century in their traditional fields of endeavor, chemical corporations continued to improve the performance of their products by devising new methods for their manufacture. A case in point is supplied by the development of the metallocene catalysts. First described by Walter Kaminsky of the University of Hamburg in the 1980s and pioneered in the field of polyolefins by Exxon and Dow, these organometallic initiators are still more expensive than the conventional Ziegler-type catalysts. They have nevertheless already gained wide acceptance in the field of polymer production because they make possible, owing to their special configuration, the production of a second generation of polyethylene and polypropylene plastics with improved characteristics. The substitution of butane for benzene in the production of maleic anhydride, which began in the 1980s, has at the same time lowered the cost of manufacture of this intermediate and done away with benzene as an objectionable raw material. In the period covered in this chapter, only biotechnology has offered the chemical industry new opportunities for spectacular developments in applications related to both pharmaceuticals and agrochemicals.

In the 1980s, as has been discussed (see p. 342), only a few bioengineered products were developed: human insulin and HGH, both by Genentech, came on the scene, followed by the antithrombotic tissue plasminogen activator (Genentech, 1987) and the red corpuscle producer erythropoietin (Amgen, 1989). About the same time (1985) Abbott introduced a diagnostic test that could detect the AIDS virus in human blood collected for transfusions. By the early 1990s, however, through alliances with pharmaceutical laboratories, all these bioengineered products were already in commercial use with annual sales passing $100 million, and one third of the research projects of the major pharmaceutical companies were based on biotechnology. It is significant that Genentech, the biotechnology company that Roche now controls with ownership of 59% of its shares, has diversified its product range to include oncologic drugs for the treatment of lymphoma and breast cancer as well as cardiovascular products.

The technologies that have thus transformed drug discovery have also been applied to agriculture in such a way that new transgenic plant varieties were produced with characteristics that could not be easily obtained through cross-pollination. Monsanto, which had relied until the end of the 1980s on agrochemicals obtained synthetically, started selling seeds, which yielded transgenic crops engineered in such a way that they either offered outstanding resistance to herbicides or generated insecticides in the form of *Bacillus thuringiensis* toxins. By 1999, 33% of America's corn and 55% of its cotton crop as well as 99% of Argentina's soybeans came from such genetically modified varieties. Contrary to what happened to the pharmaceutical applications, the use of biotechnology in agriculture has been opposed by such pressure groups as Greenpeace and by public-sector agricultural institutions, so it may take some time before genetically modified crops are universally accepted.

In the 1990s the pharmaceutical industry proved to be innovative not only through the use of biotechnology but also through the discovery of new drugs by other methods. It applied the technique of chiral chemistry to isolate from a racemic blend the optically active molecule that is the one desired as a drug. Combinatorial chemistry was another tool the industry began to use: in one stroke thousands of small molecules could be made for screening as drugs rather than having to synthesize the molecules one at a time. It made use of improved drug-delivery systems that could bring new life to older products or maximize the number of potential drug formulations likely to accelerate the path from preclinical trials to final approval.

It even met with luck through serendipity in the field of "lifestyle drugs." Sildenafil, for example, which is being offered by Pfizer under the brand name Viagra as a treatment for male impotence, was initially developed as an antianginal drug before its property of improving male sexual performance was detected. Upjohn first marketed minoxidil (Rogaine) as an antihypertensive before it came to be recommended as a hair growth stimulant for the treatment of male baldness.

The New Landscape

The 1990s were characterized by two main trends in the world economy. First was a move toward globalization made possible by the further development of free trade between nations; this led to a more competitive environment and made it necessary for corporations to streamline their operations and increase their productivity. Second was an expansion of stock markets, with more attention being paid to the financial performance of companies by the pension funds holding their shares and more generally by the various stakeholders.

In order to cope with the changing conditions, the management of many chemical companies of the Western world concluded that they had to operate along new lines. They gave priority to the concept of being the leaders in a few selected fields, and in order to obtain quick results, they came to favor external growth by acquisition at the expense of internal growth by innovation. This in turn led to a flurry of mergers, joint ventures, and divestitures that radically modified the landscape of the industry. Some companies already involved in pharmaceuticals and agrochemicals decided to concentrate on their life science activities exclusively. Oil companies that had previously diversified into fine chemicals and specialties left these fields and limited their ambitions to being strong in petrochemicals. Conversely, various chemical groups well established in basic chemicals tried to divest product lines considered to be too cyclical in favor of specialties. Meanwhile some individual entrepreneurs and financial buyers became interested in the very commodities that chemical giants were divesting.

At the same time these various acquisitions and divestitures were taking place, a minority of more traditional companies decided to retain the various areas in which they had been operating; they decided to try to grow internally by promoting products related to their "chemical tree" or, less often, by developing new molecules through their own research. As a result of these different attitudes, while the products made by the chemical industry remained generally the same over the decade, the ownership of the plants in which these products were made changed hands rapidly, and some well-known names in the industry disappeared, with new ones springing up in their place.

The Trials and Tribulations of the Pharmaceutical Industry

For historical reasons the activities of the pharmaceutical industry in the Western world had been operated either by "stand-alone" pharmaceutical laboratories or by divisions or subsidiaries of diversified chemical groups. To the first category belonged such well-known firms as American Home Products, Bristol-Myers Squibb, Eli Lilly, Merck, Abbott, Upjohn, and Pfizer in the United States; Burroughs-Wellcome, Glaxo, and Beecham in the United Kingdom; Roussel-Uclaf and Servier in France; E. Merck, Schering, and Boehringer, Ingelheim in Germany; and Hoffmann-La Roche (now Roche) in Switzerland. The second category included the life science operations of American Cyanamid, Hoechst, Bayer, Rhône-Poulenc, Elf Aquitaine, ICI, Ciba-Geigy, and Sandoz. However, independently of their origin, all companies involved in pharmaceuticals had to face similar challenges in the period considered in this chapter:

- A growing population of older people with more ailments to be treated;
- Higher R&D expenditures requiring returns between $300 million and $600 million for each approved active pharmaceutical ingredient;
- Patents due to expire for blockbuster drugs, which would invite competition from producers of generic drugs; and
- Higher marketing costs at a time when Social Security institutions and health maintenance organizations insisted on lower selling prices for the drugs being offered.

As if this were not enough, the pharmaceutical industry was also confronted, as were other industries, with the new concepts of globalization and "shareholder value."

In order to increase their geographical coverage, most firms on both sides of the Atlantic resorted to mergers and acquisitions. Thus, beginning in the late 1980s, Squibb merged with Bristol-Myers, SmithKline & French merged with Beecham, Rhône-Poulenc acquired Rorer, and Bayer took over Miles and Cutter Labs. Other transfers of ownership followed in 1994 and 1995: Roche bought Syntex, the Swedish Pharmacia merged with Upjohn, Hoechst acquired Marion Merrell from Dow and completed its control of Roussel-Uclaf to become HMR, and Glaxo and Burroughs-Wellcome formed a single entity. This was also the time when the pharmaceutical, crop protection, and nutrition operations of Ciba-Geigy and Sandoz were combined into a new company called Novartis (1995) and when similar operations belonging until then to American Cyanamid went to American Home Products (1996).

As the process of globalization proceeded, many diversified chemical groups started to pay more attention to the lucrative market of pharmaceuticals at the expense of their traditional chemical lines and to get interested in the promising field of biotechnology. Already in 1985 Monsanto had bought Searle before entering the field of genetically modified seeds. Du Pont also got involved on a limited scale in pharmaceuticals through a joint venture with Merck and moved into seeds. ICI, after the threat of a takeover by the British conglomerate Hanson Trust, split its life science and specialty chemicals operations from the rest of its portfolio by giving birth to Zeneca (1993).

Eastman Kodak decided in 1994 to part with Sterling Drug: the ethical drug division went to Sanofi (Elf Aquitaine), and over-the-counter drugs went to SmithKline Beecham. Dow also divested its pharmaceutical business by selling Marion Merrell to Hoechst, which later was to announce its intention to become a life science company (1997). After some hesitation Rhône-Poulenc followed the same path as Hoechst, and in 1999, having proceeded with the separation of their chemical activities from their core life science business, both firms put together, under the aegis of a new company to be called Aventis, their pharmaceutical divisions and subsidiaries as well as the former crop protection operations of Rhône-Poulenc and of AgrEvo, jointly owned until then by Hoechst and Schering.

These various moves, made under the pressure of financial analysts and with the purpose of enlarging the pipeline of active pharmaceutical ingredients close to approval, caused drastic changes in the structure of the chemical industry. Some of these changes were also a result of the outsourcing by pharmaceutical laboratories of most of their upstream chemical production, which gave a strong impetus to companies specializing in organic synthesis.

Within the pharmaceutical industry itself, the wish of each company's management to see its firm reach what is considered a "critical size" remained a constant feature as the century came to an end and the new millennium began. It led to new national and transnational mergers. In France, Sanofi (Elf Aquitaine) merged with Synthélabo (L'Oréal) and Laboratoires Pièrre Fabre with BioMérieux. Elsewhere the Swedish Astra and the British Zeneca combined their operations; Pharmacia-Upjohn took control of Monsanto; and two giant firms, Glaxo Wellcome and SmithKline Beecham, announced their intention to combine their activities, which would lead to a group with a turnover of $25 billion, exceeded only by Pfizer after its acquisition of Warner Lambert.

However, serious problems remain after such mergers take place. Short-term savings do not necessarily produce long-term growth. Furthermore, the executives and their subordinates of the merged companies may not get along with each other, the best researchers sometimes leave, and most of the time size is not the corollary of creativity. For all these reasons some "contrarians" in the industry have decided to pursue a different course at least for the time being.

Roche, with its capital still controlled by the founder's family, has opted for internal growth. Although still a market leader in vitamins, the company is concentrating on its own research for pharmaceuticals. It is developing simultaneously a line of diagnostics after acquiring Boehringer, Mannheim and has gained access to biotechnology through its controlling interest in Genentech. In order to finance these activities, Roche has at the same time announced the spin-off of its flavors and fragrances business known as Givaudan. Among other contrarians, mention should also be made of Bayer, which with a pharmaceutical operation amounting to more than $5 billion in sales, has until now run the business under its existing structure. Other chemical giants like Solvay with Solvay Pharma, BASF with Knoll, and Akzo Nobel with Organon are following a similar strategy. It is possible, however, that because of the relatively small size of their pharmaceutical businesses they may decide to do otherwise in the future.

Stand-alone pharmaceutical laboratories have also in many cases preferred internal growth to mergers and acquisitions and have prospered by doing so. Such is the case particularly in the United States with Johnson & Johnson, Eli Lilly, Schering Plough, and Merck and Company. In addition biotechnology companies, which have developed products until now commercialized by well-established laboratories, are reaching a size (as with firms like Amgen, Chiron, Genentech, and Genzyme in the United States) that will allow them to consider acquiring their own pharmaceutical companies in the not-too-distant future.

This survey of the Western world should not make us forget that Japan also has a thriving pharmaceutical industry. Three firms in particular—Takeda, Sankyo, and Yamanouchi Pharmaceutical Company—are of international repute. Takeda, the largest, is also the oldest, having been founded by the Takeda family in 1781. It has a joint venture with Abbott in the United States. Out of 25 blockbuster drugs currently available in the United States, six were discovered in Japan. But Japan remains weak in biotechnological developments, which are the source of much Western drug innovation.

The Reshuffling of Chemical Assets

Many large groups, once deprived of their life sciences activities, were left with sizable chemical activities that had to be dealt with. One widely used solution was the creation of new chemical entities through various spin-offs. Thus, between 1993 and 1999, Kodak gave birth to Eastman Chemical Company, American Cyanamid to Cytec Industries, Ciba-Geigy to Ciba Specialty Chemicals, Sandoz to Clariant, Monsanto to Solutia, and Rhône-Poulenc to Rhodia. Some more complex cases had to be solved differently. Hoechst, for example, proceeded in successive steps: in 1994 it set up a 50-50 joint venture with Bayer, called Dystar, for the purpose of managing their respective dyestuffs businesses. In 1997 it transferred its specialty chemicals operations to Clariant in return for a 45% stake in the enlarged company; finally, in 1999, it passed its industrial chemicals assets on to Celanese—the US corporation it had acquired 12 years earlier—and passed Herberts, its coatings company, on to Du Pont.

ICI also was faced with difficult problems after the creation of Zeneca, as it had an impressive range of commodity chemicals to dispose of and wanted to acquire activities that would be more lucrative and less cyclical than the ones it was left with. In 1997, having changed its top management, ICI was able to acquire from Unilever, which wanted to focus on consumer products, three profitable businesses: National Starch (adhesives), Quest International (fragrances), and Unichema (oleochemicals), now called Uniqema. While taking advantage of this opportunity to re-enter the specialty chemicals markets, ICI began to divest its bulk chemicals and polymer operations: polyester fibers and films went to Du Pont; fertilizers and ammonia to Terra Industries; explosives to AECI and Orica; autopaints to PPG Industries; polyurethane, aromatics, and titanium dioxide (TiO_2) to

Fig. 1.1 Sir James Whyte Black, one of the recipients of the 1988 Nobel Prize in physiology or medicine. While at ICI, his pioneering work in analytical pharmacology led to the discovery of β-adrenoceptor and histamine type 2 antagonists. Courtesy Nobel Foundation

Huntsman; and acrylics, fluorochemicals, Crossfield silicas, and chloralkali units to a privately held financial company called Ineos. At the end of 2000, with this ambitious program of divestitures completed, ICI had become an efficient specialties and paints company, and the three letters *I*, *C*, and *I*, which originally stood for Imperial Chemical Industries, are the only reminder of its past glory. Other groups did not even retain their former names (for example, American Cyanamid, Sandoz, Hoechst, and Rhône-Poulenc) and therefore will soon survive only as entities in the memory of old-timers. Many other changes were to occur as the reshaping of the chemical industry continued (Fig. 1.1).

The Impossible Marriage of Pharmaceuticals with Agrochemicals

Placing all life science activities under one roof appeared to be a logical decision, and it at first had the blessings of financial analysts. However, it proved unwise. Indeed, no sooner had Novartis, AstraZeneca, Pharmacia-Upjohn-Monsanto, and Aventis been created than their managements were made aware of some important facts:

- There was very little synergy between the pharmaceutical and the crop protection parts of the business.
- The margins generated in the human health sector were far higher than the ones obtained in selling products to farmers.
- Pesticide sales are cyclical because they depend on the weather and on commodity prices.
- The backlash begun in Europe against genetically modified foods has decreased—at least temporarily—the expectations for high-tech seeds.

As a result all the above-mentioned companies have decided to divest their agrochemical operations through either mergers, spin-offs, or straight sales, and new entities devoted entirely to crop protection are being born: Syngenta from the merging of the agrochemical divisions of Novartis and AstraZeneca; Aventis CropScience, the merger of the crop protection activities of Rhône-Poulenc and of Hoechst Schering, formerly known as AgrEvo, once the problem raised by the participation of Schering in AgrEvo had been solved; and Monsanto, to be left essentially as a separate unit by Pharmacia-Upjohn shortly after having been acquired by them.

While these divestitures are taking place, large chemical groups less spoiled by the high margins of the pharmaceutical industry are reinforcing their position in agrochemicals. Such seems to be the case with BASF, which in March 2000 acquired the pesticides line of American Cyanamid from American Home Products, and of Bayer, which has added strength to its range of fungicide products by buying the Flint product line of Novartis. Other groups like Du Pont that are very active in seeds and Dow Chemical, which purchased Eli Lilly’s remaining share in DowElanco in 1997, have also kept faith in the future of agrochemicals, a market worth $30 billion worldwide. Even smaller firms can find the agrochemical business particularly rewarding, provided they focus on “niche” products as FMC and Uniroyal Chemical (now part of Crompton Corporation) have done. In any case the trend seems to be for life science companies to focus on pharmaceuticals and for large chemical groups with relatively small operations in human health to negotiate a withdrawal and concentrate on agricultural chemicals, an activity closer to their traditional practice. The announcements at the end of 2000 that Knoll, part of BASF, will be sold to Abbott and that Bayer and Du Pont are considering a separation of their pharmaceuticals from the rest of their businesses provide further proof for this line of thinking.

The Fate of the Dyestuffs Sector

Still in 1992 there were six major producers of dyestuffs in Europe—BASF, Bayer, Hoechst, Ciba-Geigy, Sandoz, and ICI—and an American manufacturer of smaller size, Crompton & Knowles. However, the pressure of small competitors in India and China has since led to major changes in the industry. Indeed, with their low overhead, favorable labor costs, and lack of consideration for environmental issues, these Far East firms began to offer intermediates for dyestuffs, then finished dyes, at very attractive prices. The established producers of the Western world reacted initially by using these cheaper intermediates in their own production, but they soon realized that more drastic moves were necessary. In the case of ICI the textile dyes were transferred to Zeneca Specialties, which in turn made toll-manufacturing arrangements with BASF before being taken over by venture capitalists (Cinven and Invest Corp) in 1999 to form Avecia. As has already been mentioned, Hoechst and Bayer combined their textile dyes in a 50–50 joint venture called Dystar (1994), which BASF later joined. The dyestuffs lines of Ciba-Geigy and Sandoz went, respectively, to Ciba Specialty Chemicals and Clariant after the creation of Novartis in 1995. Crompton & Knowles sold its textile dye business in 2000 to a British manufacturer, Yorkshire Chemicals, now called Yorkshire Group.

These various changes of ownership were accompanied by plant closures and the transfer of production to such places as Brazil and the Asia-Pacific area, which offered cheaper labor costs and less severe regulatory constraints. Smaller producers of textile dyes in Europe also lowered their ambitions, with Holliday Chemical Holdings, now part of Yule Catto, closing their historical site of Huddersfield in the United Kingdom and Yorkshire Chemicals discontinuing production at its unit in Tertre, Belgium. Conversely, large firms like Ciba Specialty Chemicals and Clariant, while delocalizing their commodity dyes to more propitious regions, were able to link their dye expertise with their pigment technologies and to develop more sophisticated products for use in such applications as plastics, paints, and inks.

Consolidation in the Field of Specialty Chemicals

In their pursuit of “shareholder value” chemical companies have generally favored developing specialty chemicals over manufacturing large-volume commodities, which are considered to be less profitable and too cyclical. However, for historical reasons, the balance between these two categories of products varied from one group to another. This was particularly evident in the case of the newly born companies that resulted from the split of life science operations. Thus Ciba Specialty Chemicals was from the beginning a company focused on specialties and performance chemicals. It became even more focused with the acquisition in 1998 of Allied Colloids, a British company specializing in flocculants, and with the sale of its epoxy resins, a line of polymers in which it had been a pioneer, to Morgan Grenfell Private Equity. Clariant, with the product ranges inherited from Sandoz and more recently from Hoechst, belonged in the same league. It further enhanced its position by acquiring the former British Tar Products (now BTP) with its fine chemicals arm Archimica, a leading manufacturer of active molecules for the pharmaceutical industry, and PCR, a high-tech product specialist. The new ICI under the guidance of Brendan O’Neill had been re-created into a leading specialty chemicals corporation.

Other newly born companies were not so lucky, and their image remained blurred in the eyes of financial analysts because their product mix still included a significant proportion of commodities. Eastman Chemical Company, for example, still relied heavily on polyethylene terephthalate plastics and cellulose acetate in the field of polymers and on the acetyl chain and oxoalcohols as far as high-volume chemicals were concerned. Its management, by acquiring Peboc in the United Kingdom from Solvay Duphar as well as two other fine chemicals units in the United States, had made plans to enter the sector of organic synthesis for pharmaceuticals, only to give it up a few years later. Instead, Eastman Chemical Company decided to increase its presence in the resin market after it acquired Lawter International, an ink resin manufacturer; in 2000, Eastman also acquired the rosin esters and hydrocarbon resin lines of Hercules.

Rhodia, because of earlier mergers, is also regarded as a conglomerate rather than a purely specialty chemical firm. Having placed its toluene diisocyanate unit outside its perimeter of activity and sold its chloralkali business to a US investor, La Roche, it still retains such commodity chemicals as phenol, phosphoric acid, and sodium tripolyphosphate (STPP) and maintains a presence in such unrelated fields as rare earths, styrene butadiene rubber latex, and cellulose acetate tow. Its core businesses are also very diversified; they include nylon fibers, polyamide engineering plastics, biopolymers (guar and xanthan gum), diphenols and derivatives, silica, silicones, and surfactants. In the future Rhodia intends to develop its activities in organic synthesis for pharmaceuticals, and it has recently acquired ChiRex, a US company that complements the units it already operates in the United Kingdom and in France (ICMD, Industrie Chimique Mulhouse Dornach). Rhodia recently made a successful bid for the old British firm Albright & Wilson, which gives it a leading position in the field of phosphoric derivatives, while the surfactant line will be resold to Huntsman.

Solutia, a firm created through a spin-off by Monsanto in 1997, also looks more like a conglomerate than like a specialty chemicals company because it retains large operations in nylon and acrylic fibers as well as in upstream commodities. It has entered its phosphorus chemicals operation into a joint venture with FMC called Astaris. Another joint venture, this time with Akzo Nobel, runs its former rubber chemical business under the name of Flexsys. Solutia has acquired, from Akzo Nobel, CPFilms and has purchased from Deutsche Morgan Grenfell the line of Vianova resins that belonged initially to Hoechst. The prospects for the lucrative polyvinylbutyral film for safety glass (Saflex) look particularly promising.

Another newly born corporation, Cytec Industries, having sold its acrylic fibers to Sterling Chemicals, looks more entrenched in the field of specialties than either Rhodia or Solutia. As a spin-off of American Cyanamid, Cytec has gained a leading position in the production of acrylamide derived from acrylonitrile and is a major producer of flocculants destined for water treatment. Recently, however, it sold its line of paper chemicals to Bayer. Among the new entities that appeared after the restructuring of the life sciences operations of several large groups, the former specialty chemical business of ICI, renamed Zeneca Specialties and then Avecia after it was acquired by Cinven and Investcorp, should be mentioned.

Aside from the above-mentioned changes from outside causes, consolidating this sector of the chemical industry has also been to a great extent the work of long-established firms already well positioned in the field of specialties and performance chemicals. In the United States, for instance, Rohm & Haas through its acquisition of Morton International in mid-1999 has greatly enhanced its status as a leading producer of chemical specialties—a status it had earned through its past presence in such lines as plastics additives, biocides, agrochemicals, and electronic chemicals. With the purchase in 1999 of Lea Ronal and of the photo-resist business of Mitsubishi, Rohm & Haas has further increased its presence in the electronic materials market, and its decision to divest the salt production activities of Morton International can only strengthen its position as a supplier of specialty chemicals.

Another American firm, Great Lakes Chemical Company—whose most profitable product, tetraethyl lead, made by its subsidiary Octel was being threatened by regulatory authorities—has also decided to concentrate its efforts on specialties and performance chemicals. Great Lakes, after deciding to focus on plastics additives, acquired successively Société Française d'Organo-Synthèse (SFOS), a Rhône-Poulenc subsidiary; LOWI, an independent German firm; and the antioxidant and ultraviolet absorber lines of Enichem in Italy. Finally, in 1998, Great Lakes demerged Octel and became a company in which over 80% of its sales was devoted to specialties. Ethyl, Great Lakes' main competitor in bromine chemistry, took a different path. In 1994 it set up two autonomous companies: Albermarle Chemical Company, which took over the polymer and fine chemicals businesses, and Ethyl, which was to specialize in petroleum additives.

Olin, one of the few conglomerates left on the American scene, clarified its structure in 1999 by regrouping under a new company called Arch Chemicals all its fine chemicals operations, while retaining the metal, ammunition, and chloralkali operations under the existing organization. A year later Arch Chemicals took over the British company Hickson International, which was involved in wood-treating chemicals, coatings, and fine chemicals. The trend toward further consolidation has led FMC to plan the separation of its chemical divisions, which represent almost 50% of its turnover from its machinery and engineering operations.

Other large groups, however, have chosen to maintain their specialties in their traditional structure. General Electric, for example, generates over $6 billion worth of plastics and chemical sales, including sales of acrylonitrile butadiene styrene resins and plastics additives acquired from Weston Chemical (a former Borg Warner subsidiary) and now managed by GE Plastics. Similarly Riedel de Haen, the fine chemicals German company taken over from Hoechst by Allied-Signal is run as part of that conglomerate whose turnover of specialty chemicals exceeds $1 billion. Several other American firms belong to the category of companies with at least $1 billion worth of sales in specialties and performance chemicals, including W.R. Grace, Lubrizol, Crompton, Hercules, and B.F. Goodrich. The fate of these companies is worth considering.

W.R. Grace has shrunk considerably from its earlier days as a conglomerate under Peter Grace, and it has to be considered as a specialty chemicals company from now on. Lubrizol is still a lube additives specialist with some growing activity outside its core business. But Crompton is the outcome of two mergers that have greatly expanded the portfolio of products it now manages.

In 1996 Crompton & Knowles, then a manufacturer of dyestuffs and polymer-processing equipment, took over Uniroyal Chemical, a company resulting from a leveraged management buyout after the tire business, which had gone to Michelin, was split off. Three years later Crompton & Knowles and Witco were merged through an exchange of shares, Witco having become vulnerable as a result of unsuccessful restructuring. The new entity, now called Crompton, no longer includes the oleo-chemicals of Witco, which went to Th. Goldschmidt (SKW Trostberg), or the textile dyes of Crompton & Knowles, sold to Yorkshire Chemicals. With a turnover of around $3.3 billion it is nevertheless a well-diversified specialty chemicals company, with leading positions in ethylene propylene rubber, plastics additives, and rubber chemicals as well as crop protection products and silanes.

At the same time that Crompton appeared as a strong contender with its wide range of specialties, two well-known American firms were approaching the end of their existence as specialty chemicals manufacturers. B.F. Goodrich, originally a tire maker, had diversified successfully into chemicals. However, at the end of 2000, it had become a major aerospace industry supplier with only $1.2 billion worth of specialty chemicals sales left. This business unit has now been sold to an investor group led by AEA Investors Inc., a closely held business, and B.F. Goodrich has divested the last of its former operations. The other firm,

Fig. 1.2 Crompton's silane unit, Termoli, Italy. Courtesy Crompton Corporation

Hercules, was originally a prominent and innovative producer of chemicals and polymers with leading positions in several niche markets. When Thomas Gossage, from Monsanto, took over as chairman in 1991, Hercules had already sold its dimethylterephthalate (DMT) business and the 50% share it owned in Himont, a polypropylene joint venture with Montedison. From 1991 to 1996 the company went through a restructuring that combined the sale of various parts of its portfolio of activities (aerospace, flavors and fragrances, polypropylene packaging films, electronics, printing materials, and so forth) with an ambitious share repurchase program. This program raised the market value of Hercules stock to a peak of $65 per share in 1995, a year before the chairman retired (Fig. 1.2).

By that time, however, the company was left with only a few businesses, which although quite profitable, offered little growth potential. To give the company a boost, the new chairman considered entering the field of water treatment. In 1997 Hercules made an offer for Allied Colloids, but it was outbid by Ciba Specialty Chemicals. A year later, however, Hercules, by paying a high premium, was able to acquire BetzDearborn, a water-treatment specialist formed when Betz Laboratories bought the Dearborn units of W. R. Grace. The deal boosted Hercules' revenues from $1.9 billion to $3.5 billion, but it also increased its debt substantially at a time when competition in the water treatment market had become much more acute owing to the consolidation of the two entities Calgon and Nalco Chemical Company under the aegis of the French group Suez Lyonnaise des Eaux. The management of Hercules, hoping to improve the company's debt ratio, decided to part with other assets: their food gum operation is now a joint venture with Lehman Brothers Merchant Banking partners as majority shareholders; the resins operation, for which they held a leading position, has been sold to Eastman Chemical; and FiberVisions, the world's largest producer of thermally bonded polypropylene fibers, is for sale. As was to be expected, the valuation of Hercules stock has been severely downgraded, reaching a low of $14 per share. This decline has in turn attracted the attention of the well-known raider Samuel Heyman, the chairman of International Specialty Products, which is the new name for the former GAF Corporation. With Thomas Gossage now back in the driver's seat, Hercules may not have any other choice but to be sold and disappear as a going concern after a long and often brilliant existence that began in 1912.

This story is typical of what happens to an otherwise healthy company when it is managed for too long under the pressure of short-term financial considerations. The fate of Laporte Chemicals, a British company founded in 1888 by the chemist Bernard Laporte to produce hydrogen peroxide (H_2O_2) for use as a bleach, was not any better than that of Hercules. Yet by 1995 Laporte had grown into a uniquely large specialty chemicals corporation with a turnover of £1 billion. It had quit the phthalic anhydride and TiO_2 sectors in the 1980s and in 1992 dissolved its joint venture with

Solvay, Interox making H_2O_2. Having made several acquisitions in the fields of organic peroxides, adhesives, and process chemicals, Laporte was operating through 16 strategic business units. Between 1995 and 1998, however, the management of the company started to divest some businesses that were underperforming, the remaining portfolio being focused on fine organics, peroxide initiators, construction chemicals, electronic materials, pigments, and additives. This reshuffling of activities culminated in 1999 with the acquisition of Inspec, a British firm that resulted from the leveraged management buyout 7 years before of several chemical operations and in particular of parts of British Petroleum's specialty chemicals and of Shell's fine chemicals divisions. No sooner had this last acquisition been completed than the new management in charge of Laporte, with a view to reducing the outstanding debt, decided to part with those units that were not directly involved with fine organics and performance chemicals. As a result the units that represented half of the total turnover were sold at the beginning of 2000 to Kohlberg Kravis Roberts & Co., a private equity investor. Finally, having reached a perfect size for being taken over, Laporte, or rather what was left of it, was acquired a few months later by Degussa AG and disappeared from the British scene as an independent chemical corporation after 112 years of operation.

Degussa AG is the new name of a German group that is slated to become one of the largest fine chemicals and specialty chemicals corporations in the world. But to achieve this status, further restructuring is needed. Indeed Degussa itself, which celebrated its 125th anniversary in 1998, was by then a member of VEBA AG, a large German energy concern that owned 100% of Chemische Werke Hüls. The next step was for Degussa and Hüls to be merged through an exchange of shares that gave VEBA a stake of 62.4% in the new company. Meanwhile, another energy company, VIAG AG, established in Bavaria, had taken full control of SKW Trostberg, to whom it was supplying power. In 1997 VIAG gained a controlling interest in a family business, Th. Goldschmidt, which it merged with SKW Trostberg 2 years later. VIAG thus owned 63.7% of the merged entity. Then VEBA and VIAG announced their intention of combining their chemical assets and gave birth at the end of 2000 to what is now called Degussa AG with initial sales of 14 billion Euros per year. In order to streamline its operations into a true specialty chemicals group, Degussa AG will need to dispose of several commodity activities (e.g., phenol, fertilizers, salt and metallic chemicals, dimethyl terephthalate, C-based alcohols, and plasticizers) and to sell its pharmaceutical subsidiary Asta Medica as well as its automotive catalyst sector. It will then be left with leading positions worldwide in such diversified fields as amino acids, carbon black, precipitated silicas, organosilanes, specialty polymers, flavors (Mero & Boyveau) oleochemicals (Th. Goldschmidt), and hydrogen peroxide.

As the consolidation of the specialty chemicals industry continued, it affected many firms in which the founder's family still held substantial portions of the capital, even though some of them kept on thriving. One example was Union Chimique Belge, whose management had been wise enough in the late 1980s to focus on three profitable lines: the pharmaceutical sector, with a blockbuster antiallergic drug called Zyrtec; the chemical sector, with ultraviolet curing and powder-coating technologies; and the film sector, with a leading position worldwide in oriented polypropylene and cellophane films. The German company Wacker, in which Hoechst had been a partner from the start, also managed to retain its independence, although it was still active in some commodity-type products (acetic acid, vinyl monomers and polymers, and silicon carbide), it chose to emphasize lines that showed greater profitability, such as silicones, hyper pure silicon for semiconductors, and specialty and fine chemicals. In France the well-run starch and derivatives producer Roquette Frères, after a short flirtation with Rhône-Poulenc, was able to recover its autonomy.

Other firms were less fortunate; they lost their independence because the founder's successors either could not agree on a management plan or had to sell their shares to pay inheritance taxes. Thus, Compagnie Française des Produits Industriels, which belonged to the Hess family, was sold in 1996 to Fernz Nufarm, a firm from New Zealand specializing in agrochemicals. As has been discussed, Th. Goldschmidt of Germany was taken over by VIAG. Another German firm, Raschig, which was founded in Ludwigshafen in 1891 and which had remained under the control of its family owners since then, was integrated in the mid-1990s into the PMC Group of Philip Kamins, a California-based entrepreneur with interests in plastics and specialty chemicals. Even a powerful company like the venerable Henkel had to reconsider its position. In 1999 it set apart under the name of Cognis its specialty chemicals activities, including production of an important range of oleochemicals. Henkel, still controlled by the heirs of Fritz Henkel, has made clear its intention to leave this field in order to better compete in the consumer products area with such large groups as Unilever and Procter & Gamble.

The specialty chemicals industry is estimated to have an annual turnover of $200 billion worldwide, with production scattered among a great variety of suppliers. Consolidation of that industry was to be expected owing to the pressure of shareholders and to management eager to rationalize their portfolio of products. Some corporations like Ciba Specialty Chemicals, Clariant, ICI, Rohm & Haas, Crompton, and Great Lakes have succeeded in being considered as specialty chemicals companies in their own right. Others, such as Rhodia and Degussa AG, still have a product mix that is much too diversified to reach that status. They will have to restructure their portfolio further if they want to improve their share value. Still others, because their management was led by short-term objectives or made unwise divestitures and acquisitions, have been penalized by the stock markets and will no longer survive as independent entities: Laporte and Hercules are among these sad cases. Aside from bowing to financial considerations, the consolidation of the industry also had at its root the necessity of reducing the number of participants once their customers became less numerous and acquired a worldwide presence. For instance, with only three major tire manufacturers left—Michelin, Goodyear, and Bridgestone—the number of the main rubber chemicals manufacturers also had to shrink: Solutia and Akzo Nobel merged into Flexsys, facing Uniroyal Chemical (Crompton) and Bayer as their major competitors.

The Case of Fine Chemicals

Fine chemicals, sold in relatively small volumes and at rather high prices, are obtained through various organic chemical reactions. Their production does not require heavy investments as they are made in multipurpose units equipped with glass-lined and stainless-steel reactors. For this reason this sector of the chemical industry has attracted skilled entrepreneurs who know how to surround themselves with teams of good chemists capable of conducting multistep syntheses leading to the formation of complex molecules used by pharmaceutical laboratories and agrochemical firms. Over the years, however, these customers of the fine chemicals industry realized that they should make their active ingredients themselves and so created their own chemical departments to carry out at least the final steps of the synthesis of the products they had developed.

During the 1990s a drastic change took place in that many pharmaceutical companies saw their R&D and marketing expenditures rise considerably, so they decided to outsource the synthesis of their proprietary drugs. At the same time many patents protecting well-known ethical drugs expired, and the same drugs came to be offered on the market as generics by competitors. As a result the chemical subsidiaries of major laboratories were offered for sale, and the toll-manufacturing activities of well-established fine chemicals firms increased spectacularly.

Thus, in 1997, Glaxo Wellcome handed over its Greenville site in the United States to Catalytica and its Annan plant in Scotland to ChiRex in exchange for supply contracts running for a 5-year period. Similar deals were made when in 1997 Warner Lambert sold Sipsy Chimie Fine, the chemical operation of its French subsidiary Laboratoire Jouveinal, to PPG Industries and a year later when the British firm BTP bought the French firm Hexachimie from Bristol-Myers Squibb. Prior to its merger with Sandoz, Ciba-Geigy had, for its part, divested its French subsidiary La Quinoléine, which was acquired by Organo-Synthése. Sanofi sold its operation in Spain, Moehs SA, to PMC.

While these divestitures were taking place, traditional fine chemicals manufacturers were able to enjoy new contracts for the synthesis of active ingredients from companies involved in crop protection products and pharmaceuticals. Many European companies were the beneficiaries of the new trend. In Switzerland the leader was Lonza, which became independent in 1999 when its mother company, Alu Suisse, merged with the Canadian aluminum manufacturer Alcan. Other Swiss firms—such as Siegfried AG, EMS-Dottikon, Cilag, and Orgamol, Röhner, now part of Dynamit Nobel—took advantage of these developments. Clariant entered the field on a big scale in 2000 with its acquisition of BTP.

In France as well, some companies that had been traditionally involved in fine chemicals through their specific expertise engaged in similar activities. The state-owned Société Nationale des Poudres et Explosifs decided to make use of its skills in phosgene and nitration chemistry by buying Isochem from the Wirth family and making inroads in the field of polypeptides. Rhodia, which had inherited ICMD from its former links with Rhône-Poulenc, acquired ChiRex at the end of 2000. Another firm, Produits Chimiques Auxilliaires de Synthèse (PCAS), which belonged to Dynaction, a medium-sized French conglomerate, also expanded by taking over Pharmacie Centrale de France's chemical subsidiary and a plant in Limay belonging to the German Schwartz Pharma.

In Holland, under the leadership of Simon de Brée, DSM chose to diversify into fine chemicals first by acquiring the Austrian Chemie Linz from ÖMW, then in 1999 by taking over Gist-Brocades, a Dutch supplier to the pharmaceutical industry. Finally, with the purchase of Catalytica's pharmaceutical subsidiary in 2000, DSM has become a leader in this field and now generates a turnover of $1.3 billion.

Conversely, the United Kingdom has seen many of its fine chemicals companies disappear from the map: Hickson International is now a part of Arch Chemicals, Courtaulds Chemicals has been taken over by Akzo Nobel, and the fine organics and performance chemical lines of Laporte were recently acquired by Degussa AG. Peboc got sold by Duphar to Eastman Chemical Company, which now wants to dispose of it. A similar trend has occurred in Germany where Riedel de Haen was acquired in 1995 by Allied Signal, Raschig by PMC, and Boehringer Mannheim by Roche.

Italy, long the source of expertise in multistep synthesis built up over the years by individual entrepreneurs, did not fare any better. Francis in Milan, after having been owned by Laporte, is now controlled by Degussa AG; Profarmaco was acquired by Nobel Chemical, which in turn is owned by

Cambrex; OPOS went successively to Hoechst Marion Roussel and then to Holliday Chemical Holdings, itself a part of Yule Catto.

The United States was not as strongly endowed with fine chemicals expertise as were many European countries since the skills of the organic chemist were less cultivated in America than those of the chemical engineer and only a few American companies have been successful in this field in the past. But mention should be made here of Cambrex, a company that began with the acquisition of the former Baker Castor Oil Company and of Nepera Chemicals, a manufacturer of pyridine and derivatives. In 1993 Cambrex decided to become a global supplier of products to the life sciences industry. It made two major moves to accomplish this: in 1994 it acquired from Akzo Nobel the Nobel pharmaceutical chemistry business, with its lucrative subsidiary in Italy, Profarmaco, a producer of generics; and in 1997 it purchased BioWhittaker, a supplier of human cells and cell cultures, which gave Cambrex a promising future in the field of bioscience. Thus the sector of fine chemicals has been the scene of many divestitures and acquisitions, and since its consolidation is far from complete, it still offers opportunities to those firms that either enjoy some lucrative niche markets or that reach a size that allows them to compete successfully with the established leaders.

The Further Concentration of the Industrial Gas Business

At the end of the 1990s there were four major producers of industrial gases in the world: L'Air Liquide, British Oxygen Company, Praxair, and Air Products. In addition to these leaders the German firms Messer Griesheim, the Swedish AGA, and the Japanese Nippon Sanso played an important role in their respective markets. If we consider that Praxair was the former Linde Division of Union Carbide Corporation, these companies were well-established firms, the youngest among them, Air Products, having celebrated its fiftieth anniversary in 1990. Since it is not economical to transport liquefied gases long distances, all of these producers operated on a worldwide basis, with plants scattered throughout the five continents.

British Oxygen had invested more heavily than its competitors in the former Commonwealth and in the Asia-Pacific area, while AGA had invested more heavily in the Nordic countries and in Latin America. Air Products was the only company with a sizable part of its total turnover devoted to commodity and specialty chemicals, which represented over 30% of its yearly sales of $5 billion. L'Air Liquide, although initially involved in chemicals, had sold its fertilizer subsidiary, Société de la Grande Paroisse, to Elf Atochem and its pharmaceutical arm, Lipha, to E. Merck, in Darmstadt, Germany. Although the competition for industrial gases was fierce between these companies, new applications were developed in the fields of electronics, semiconductors, health, oil refining, and food processing, and all the firms with standardized plants and similar technologies were enjoying a thriving business.

At the end of 1999, however, British Oxygen, which was overinvested in Southeast Asia, approached Praxair with an offer to merge the two groups. Dating back to 1992, Praxair was a spin-off of Union Carbide, made necessary when in 1985 in order to fight a takeover bid from Sam Heyman, the majority owner of GAF, Union Carbide decided to buy back its shares and as a result considerably increased its outstanding debt. The proposed merger of Praxair and British Oxygen would have had a considerable impact on the industrial gas market. Feeling threatened, Air Liquide and Air Products made the unusual move of a joint bid for British Oxygen, but antitrust considerations led the parties to give up their plan after a long period of negotiation.

Meanwhile less ambitious schemes have materialized. Indeed, in 2000, Linde took over AGA and Messer Griesheim, which was owned initially by Hoechst, now part of Aventis (66.6%), and the Messer family (33.3%) cut a deal whereby two financial associates, Allianz and Goldman Sachs, will buy from Aventis the shares formerly held by Hoechst. The new landscape offered by these various moves may still change somewhat as the companies just merged will for financial or antitrust reasons have to dispose of assets that should be of great interest to the majors.

The Changing Ties Between the Oil and Chemical Industries

Beginning with the World War II effort, oil companies in the United States became involved in petrochemicals and polymers and contributed to major innovations in those fields. The process of downstream integration was then seen by them as a way to upgrade their feedstocks and to develop new activities. Over the years oil majors like Exxon and Shell, either through a frenzy of acquisitions or through internal growth, became important producers not only of petrochemicals, synthetic rubbers, resins, and plastics but also of specialties and even fine chemicals. In the 1990s low crude-oil prices and increased competition on a worldwide scale led the management of the oil companies to reconsider their position, with the idea of restricting their chemical operations to base commodities produced at the site of their main refineries. In order to achieve this goal, a vast program of divestitures was begun either through straight sales or spin-offs or through joint ventures run at arm's length from the companies' owners.

Thus in 1994 Texaco sold its chemical operations to Huntsman. ARCO Chemical Company, after floating its Texas-based petrochemical activity under the name of Lyondell in 1989, sold to that new entity 9 years later the remaining operations of ARCO Chemical. The joint-venture route was used by Shell when in 1989 it bought the stake Montedison held in Montell, their polypropylene subsidiary, before merging it with the polyolefin operation of BASF, now called Basell. Similarly, Occidental Petroleum, which with Oxychem still holds strong positions in chloralkali products and specialties inherited from the former Hooker Chemical, is trying to unload some of its petrochemical operations. To this end it has become a shareholder of Equistar Chemicals, now the largest US producer of ethylene, which is as a result a joint venture between Lyondell (41%), Millennium (29.5%), and Oxychem (29.5%). Millennium Chemicals was itself born in 1996 from a demerger of Quantum Chemicals, formerly owned by Hanson Trust, and like Oxychem it wishes to find a buyer for its shares in Equistar.

The consolidation of the oil industry, which started at the end of 1998 with the merger of BP-Amoco and was followed in 1999 by BP-Amoco's acquisition of ARCO and then by the still larger Exxon-Mobil merger, can only further dilute the part of petrochemicals in the product mix of these new entities. The announced takeover of Texaco by Chevron will have the same effect. The chemicals of Chevron and Phillips Petroleum have already been combined under a new entity, Chevron Phillips Chemical Company. Oil companies everywhere have followed a similar trend.

Thus in 1999 the Spanish Repsol, by acquiring YPF in Argentina, and the Norwegian Norsk Hydro, by taking over Saga Petroleum in Norway, have both contributed to a decrease in the share of chemical sales in their consolidated turnover. Norsk Hydro, still strong in fertilizers, no longer considers its ethylene, vinylchloride, PVC, and chlorine operations as core businesses. ENI in Italy is also prepared to part with some of Enichem's activities: it disposed of its acrylic fibers and sold at the end of 2000 the polyurethane business of Enichem to Dow Chemical while trying to take over Union Carbide's share of Polimeri Europa, following Dow's acquisition of Union Carbide and the resulting antitrust rulings. Similarly, the Nesté Chemicals subsidiary of the Finnish Nesté Oil has been divested and acquired by a Nordic investment firm, Industri Kapital, which merged it with Dyno, a recent acquisition specializing in explosives and fine chemicals.

The only important exception to this policy of unloading chemical activities not strictly related to the immediate downstream production of refinery operations appears to be the French oil company TOTAL. Historically, TOTAL, the former Compagnie Française des Pétroles, had only a limited presence in petrochemicals and polymers through aromatics produced at the Gonfreville refinery and minority participations in alkylbenzene, polyolefins, and butyl rubber joint ventures in France. The situation changed in 1990 when ORKEM, a company born from the restructuring of the state-owned French chemical industry, was split and its assets transferred to TOTAL for the specialties and to Elf Aquitaine's chemical arm Atochem for the petrochemicals and polymers. Through a deal completed in June 1999, TOTAL made a public exchange offer for Petrofina of Belgium. This was followed a few months later by the merger of the newly created TotalFina with Elf Aquitaine. That way a new group, TotalFinaElf, was born, with an impressive chemical arm, AtoFina, having global sales of $16 billion a year.

Ranking as the fifth largest chemical company in the world, after BASF, Bayer, Du Pont, and Dow, once its merger with Union Carbide is completed, AtoFina is a very diversified producer. Its product mix includes fertilizers and a variety of agrochemicals, four of the major thermoplastics (polyethylene, polypropylene, polystyrene, and PVC), acrylics and polymethyl-methacrylates, a chloralkali chain, oxygenates (H_2O_2 and organic peroxides), hydrazine and derivatives, and fluorinated products. While commodity petrochemicals and plastics represent less than 20% of its business, AtoFina has inherited from TOTAL a lucrative portfolio of specialties in such varied fields as resins (Cray Valley), adhesives (Bostik), radiation curing coatings (Sartomer), paints (Kalon), and rubber articles (Hutchinson). Originating from various restructuring steps and acquisitions over the last 10 years, AtoFina, because of its size and diversity, does not resemble any of the chemical operations of the major oil companies.

Its future within TotalFinaElf will depend on whether the management of the group feels that its chemical activities should be closer to the cracker or can be run as a full-fledged chemical operation. In any case some streamlining cannot be avoided. In fact, prior to the merger, TOTAL had already sold its Ink Division (Coates Lorilleux) to Sun Chemical, a subsidiary of the Japanese Dainippon Ink & Chemicals. More recently a department specializing in metal treatment has gone to Henkel, whereas unsuccessful attempts have been made to dispose of the agrochemicals of ElfAtochem. At this stage anyway and for historical reasons the approach of TotalFinaElf can be considered unconventional when we compare it with that of the oil companies. With improving crude oil prices and better capacity utilization in their refinery units, the oil companies are concentrating their major reinvestments on what they consider to be their core business and limiting their production of chemicals to the base commodities derived from the olefins of the steam cracker or other feedstocks.

The Role of Entrepreneurs and Private Equity Funds

The various mergers and divestitures described in this chapter have provided opportunities for entrepreneurs and venture capitalists to make acquisitions in fields that were no longer of interest to major chemical firms. Already in the 1980s a similar situation had given birth to new chemical entities launched by such daring entrepreneurs as Gordon Cain and Jon Huntsman. Cain Chemical, the result of a highly leveraged management buyout, was for a while a large petrochemical concern before being taken over in 1988 by Oxychem.

Huntsman Chemical, however, has proved longer lasting. Its first steps in the chemical business started with the acquisition of the polystyrene activities of Shell, followed 3 years later by those of Hoechst. By 1999, through external growth, Huntsman Chemical ranked as the largest privately held chemical company. That year its revenues reached $8 billion as a result of the acquisition of ICI's polyurethane, titanium dioxide (TiO_2), aromatics, and petrochemical operations by Huntsman ICI Holdings (70% Huntsman Chemicals and 30% ICI). ICI's Wilton, Teesside cracker has in particular given Huntsman a strong position in olefins.

With the cash obtained from the sale of its styrene business to the Canadian NOVA Chemicals, Huntsman could afford to proceed with more acquisitions, and in 2000 it took over the surfactants operations of Albright & Wilson, a British firm that is now part of Rhodia.

Another American entrepreneur, D. George Harris, who had managed the SCM chemical conglomerate before it was acquired and dismantled by Hanson Trust, had also shown an interest in basic commodities being divested by the majors. During the 1990s Harris was able to set up operations in such fields as soda ash, phosphate, and boron, both in the United States and in Europe. In 1998, facing a difficult period, Harris managed to sell these operations to IMC Global, an American firm active in agribusiness.

Recently, new opportunities were offered to private equity institutions and to ambitious managers who can use the funds that venture capitalists provide to revamp companies in need of restructuring and make substantial money when selling them a few years later. In Europe the PVC field has offered such an opportunity because it suffers from intense competition and is submitted to the pressure of such environmentalists as Greenpeace. As a result several PVC producers have been anxious to sell their assets or at least to participate in the restructuring of the sector. They have found buyers among the private equity institutions. With a consortium of financiers, including Advent International and Candover, George Harris acquired in December 2000 Vestolit, the vinyl business of Degussa-Hüls. For its part Advent has also bought Vinnolit and Ventron from Wacker and Celanese, respectively. European Vinyls Corporation (EVC), the troubled joint venture formed by Enichem and ICI, may well be the next to go to private capital for the consolidation to be further advanced.

The fibers sector is another area in which European producers are suffering because of the presence in Asia of strong competitors, a situation still aggravated for these producers by the gradual elimination of the protective textile quotas set up under the Multi Fiber Agreement. Here again the situation is attracting the attention of financial investors as well as that of private textile firms in countries enjoying the advantage of low labor costs. As an example CVC Capital Partners, in January 2000, purchased a 64% stake in Acordis, the new entity resulting from the takeover of Courtauld's fiber business by Akzo Nobel and its subsequent float on the stock market. Du Pont, in its effort to unload part of its activities in polyester fibers, set up between 1999 and 2000 joint ventures with two foreign firms: Alpek, a subsidiary owned entirely by the Mexican firm ALFA, the largest private petrochemical concern in Latin America, and Haci Ömer Sabanci Holding, an important financial and industrial conglomerate in Turkey.

The increasing role played by private equity capital can be explained by the important sums of money they can collect from pension funds and other institutions. It may seem strange that they should direct investments toward chemical operations that are being spun off by the big players. However, some of the assets being divested are offered at attractive prices, and the managers who are put in charge as well as their backers can hope to restore rapidly some profitability to the operations through a thorough program of cost cutting and restructuring. In any event these investors are generally inclined to take a short-term view of the business, and with the exception of entrepreneurs like Huntsman a quick profit is their main objective.

The Everlasting Presence of Contrarians

In the past 10 years many chemical companies have changed their product mix and even their names and have taken entirely new directions in order to cope with a new environment and in the hope of pleasing financial analysts. During that same period, however, some large chemical corporations, while prepared to streamline their operations and take advantage of new opportunities, have maintained their faith in the traditional activities that had been the key to their success. These contrarians have taken a long-term view of the business they are in, in the hope that their approach will in the end meet with the approval of the stock market because internal growth, which takes time to deliver, is an essential part of any industrial undertaking.

The German chemical giant BASF ranks among the top contrarians. Its management still believes in the *verbund* concept of linking products from the raw material down to the most elaborate derivatives. It is a concept particularly suited to a large site like Ludwigshafen where vertical integration can be easily practiced, with products flowing from one unit to another. BASF has also maintained its connection to oil as a raw material, which started in 1952 through Rheinische Olefin Werke in Wesseling, BASF's joint venture with Shell, and was amplified in 1969 when the Wintershall refinery was acquired. The fact that BASF remains diversified in various lines of organic and inorganic chemicals does not prevent it from restructuring its traditional operations when the need arises. Joint ventures were thus entered into at the end of the 1990s, in polyolefins with Shell (Basell), in PVC with Solvay (Solvin), and in dyestuffs with Bayer and Hoechst (Dystar). BASF finally decided to part with its pharmaceutical subsidiary Knoll, sold at the end of 2000 to Abbott, and it is negotiating with Kali und Salz its withdrawal from some fertilizer operations. With a turnover around $30 billion per year in group sales, BASF retains a leading position in the world chemical industry.

Dow Chemical has followed a similar path. Over the last 5 years it has reinforced its position in Europe in basic chemicals and commodity polymers by the acquisition and reconstruction of the Buna Sow Leuna Olefinverbund complex in eastern Germany, which has been fully owned by Dow since June 2000. At the same time the olefin crackers in Terneuzen, Holland, have been greatly expanded, while methylene diisocyanate (MDI) capacities for use in polyurethanes are being increased in Stade, Germany, and in Estarreja, Portugal. Above all, Dow has seen at the beginning of 2000 its acquisition of Union Carbide approved by both the Federal Trade Commission in Washington and the European Commission in Brussels. This merger, leading to yearly sales of $28.4 billion, puts the new Dow Chemical in the same league with BASF and Du Pont. The approval of the merger implies that Dow sell its ethylene amine business, which will go to Huntsman Chemical, and its ethanol amine operations, which will be acquired by Ineos, a financial institution. It also commits the new entity to part with the 50% stake Union Carbide held in Polimeri Europa, a joint venture with Enichem to manufacture polyethylene. At the same time Dow will further strengthen its position in polyurethanes by acquiring the polyurethane line of Enichem, which includes TDI, MDI, and polyols.

Like BASF, Dow is a firm believer in vertical integration. It has always tried to connect its thermoplastic and elastomer productions to upstream steam crackers while making its chloralkali operations less dependent on PVC than other producers by developing various lines of chlorine derivatives on site. Even before BASF, Dow left the field of pharmaceuticals with the sale of Marion Merrell to Hoechst and has increased its stake in crop protection chemicals through the acquisition of Eli Lilly's interests in DowElanco. The announced takeover of Rohm & Haas fungicide and insecticide lines will further enhance Dow's presence in this sector.

Among the contrarians, Bayer has had for historical reasons a different approach to the problem of remaining a well-rounded producer of chemicals. It has always been a leader in such specialty chemicals as plastics and rubber additives and supplies master batches through its subsidiary Rhein Chemie. Its polymers are mainly centered on synthetic elastomers and specialty plastics like polyurethanes and polycarbonates that Bayer itself develops. These are not subjected to the cutthroat competition that commodity thermoplastics are. Bayer has remained involved in aroma chemicals with a well-known subsidiary, Haarmann & Reimer. In pharmaceuticals the company occupies an honorable rank, although in the future its management may want to separate that branch from the rest of its activities in preparation for further alliances. Like BASF, Bayer considers its agrochemical line as a core business. However, unlike BASF, it is less interested in integrating upstream production to oil raw materials and has therefore decided to sell to its partner BP-Amoco its 50% stake in their EC Erdölchemie joint venture. When necessary, Bayer is also prepared to divest noncore activities as it did in mid-1999, with 70% of Agfa-Gevaert or with its Dralon fiber operation sold to the Fraver Group in Italy. Taking these divestitures into account, Bayer's yearly turnover should approach that of the new Dow Chemical—Union Carbide merger.

Another chemical giant, Du Pont, while still broadly based, has been less inclined than its above-mentioned rivals to remain focused on its traditional lines of activity and has changed its course of action several times over the last 10 years. In 1981 Du Pont purchased Conoco in order to protect itself from oil shocks; 20 years later its management was no longer interested in maintaining that business, which was then spun off. Du Pont also seemed eager at one time to enter the pharmaceutical sector and had made an alliance with Merck in that field. In 2000 Du Pont tried to acquire the former Romainville research center of Roussel-Uclaf in France, which belonged by then to Aventis. After failing in that attempt, Du Pont became somewhat disenchanted with its pharmaceutical business, which had yearly sales of only $1.6 billion out of a total turnover of $27 billion; it is predicted that they will eventually give up that business. Du Pont remains determined, however, to become a leader in seeds after its acquisition in 1999 of Pioneer Hi-Bred. In spite of some temporary setbacks in Europe for genetically modified organisms, this branch of life science should provide the company with a bright future. Like other Western synthetic fiber manufacturers, Du Pont has been trying to unload some of its assets in that area, which has suffered from the impact of

Asian producers. And as discussed earlier, it has to that effect made joint ventures in polyesters with such private groups as ALFA in Mexico and Haci Ömer Sebanci in Turkey.

Akzo Nobel, with yearly sales about half as large as those of the four majors, belongs nevertheless to the league of contrarians in that it occupies leading positions in many traditional lines of the chemical industry. Under a 5-year plan initiated in the second half of the 1990s, the company, which resulted then from the merger of the Swedish Nobel Industries with the Dutch AKZO, undertook an ambitious restructuring program. The merger had given Akzo Nobel the number-one position worldwide in coatings and a valuable stake in surfactants and in pulp and paper chemicals. From the former AKZO the company had inherited a lucrative pharmaceutical line with the Organon prescription drugs and with Diosynth, a producer of raw materials for pharmaceuticals, which was reinforced by the Buckhaven unit of Courtaulds in the United Kingdom. The animal health care business Intervet International is also profitable. The restructuring of the group involved the creation of joint ventures for rubber chemicals with Monsanto (Flexsys) and for PVC with Shell (Rovin), acquired in late 1999 by the Japanese Shin-Etsu. The soda-ash operation was sold early in 1998 to Brunner, Mond. The problems of the fibers division were treated first by a merger with the British Courtaulds, followed by a spin-off of the fibers operations of the joint entities, sold in 2000 as Acordis to its new owners, which included—aside from a minority stake held by the management—the financial institution CVC Capital Partners and Akzo Nobel itself.

Another contrarian, DSM in the Netherlands, had been a state-owned company before it became privatized, a process that began in 1989 and was completed in 1996. From its past it had inherited positions in fertilizers, industrial chemicals, and such intermediates as melamine and caprolactam as well as polyolefins, with access to basic olefins through its own crackers in Geleen, Netherlands. In 1997 DSM acquired the polyethylene and polypropylene operations of Hüls (VEBA) with the Gelsenkirchen site. The company had also diversified into elastomers, having purchased in the United States the Copolymer Rubber and Chemical Corporation, which contributed to DSM's expansion into the fields of ethylene propylene, styrene butadiene, and nitrile rubbers. DSM is also a supplier of industrial resins and engineering plastics products. It had made inroads in the fields of fine chemicals in 1985 when it bought Andeno, a producer of synthetic intermediates for the pharmaceutical industry. By the end of the 1990s this initial acquisition had been followed by that of Chemie Linz; of Gist Brocades, a biotechnology firm; and in 2000 of Catalytica, thus giving DSM a yearly turnover of $1.3 billion in life science operations out of a total yearly turnover of around $7 billion. The surge in such activities will probably force DSM to slow down its heavy investments in olefins and petrochemicals, giving the company a better balance between commodities, performance, materials, and fine chemicals.

The last contrarian worth mentioning, with a turnover at the level of that of DSM, is Solvay, a company that dates back to 1863 and is still involved in most of its traditional lines. It is organized around four main sectors—chemicals, plastics, processing, and pharmaceuticals—with no intention of spinning off any of them. Solvay remains a leader in such commodities as soda ash, chloralkali chemicals, and oxygenates (H_2O_2 and peroxides). For its lines of polymers, it set up joint ventures with BASF (PVC) and Petrofina (high-density polyethylene) in order to improve its access to raw materials and acquire new technologies. Solvay also has considerable experience in plastics processing, an activity that represents more than one third of its total sales. In the health sector the group—after selling its crop protection line to Uniroyal Chemical (now Crompton) and the U.K. fine chemicals producer, Peboc, to Eastman Chemical—has decided to concentrate all its efforts on Solvay Pharma, its pharmaceutical unit that already accounts for 22% of its earnings and specializes in such areas as gastroenterology, psychiatry, and gynecology.

It takes courage these days on the part of the management of the contrarians to go on thinking of the chemical activities in which they operate as made up of several parts to be kept under one roof. Indeed many investors have come to feel, as do financial analysts, that a split-up of the activities of large and diversified chemical groups would unlock more value in the shares they hold. This makes the temptation to dismantle existing organizations sometimes hard to resist.

The Case of Japan

In the 10 years discussed in this chapter, the economy of Japan languished, and the Japanese chemical industry, the third largest in the world, could not avoid the consequences of that situation. It also suffered from several factors specific to Japan. First, the largest chemical companies belonged to the "Keiretsu" system, which had replaced the "Zaibatsus" of pre-World War II, and featured a complex network of cross-shareholdings dominated by banks. Second, these companies remained under the influence of the Ministry of International Trade and Industry (MITI), which proved to be a hindrance when the time came to restructure the chemical sector in the early 1990s. Third, internationally, with the exception of such firms as Dainippon Ink & Chemicals and Shin-Etsu, Japanese chemical groups had not developed a strong presence outside Asia and depended heavily on indirect exports through their domestic customers or on licensing for their activities overseas. Fourth, Japanese society, used to lifetime employment, was not prepared for the social upheavals inherent in any serious restructuring. For all these reasons the changes required from the chemical industry by the new conditions of world trade have been slow in coming and were only partially accomplished by the end of the 1990s.

As an example, the necessary merger of Mitsui Petrochemicals and Mitsui Toatsu Chemicals became effective only in 1997 and was long delayed because of the reluctance of each senior management group of these two members of the Mitsui group to merge. Similarly difficult to achieve has been the consolidation of Mitsubishi Chemicals and Mitsubishi Petrochemical, which finally took place in October 1999 and led to the birth of the eighth largest chemical producer in the world, with a yearly turnover equivalent to $15.7 billion in fiscal year 1999–2000. The proposed merger of Sumitomo Chemical with Mitsui Chemicals should create an even larger group when it is implemented.

These three majors of the Japanese chemical industry can be considered as contrarians in their own right since they maintain a strong presence in all facets of their industry. Contrary to their American or European counterparts, they are not submitted to the pressure of investors eager to obtain the best value for their shares. While this could be an advantage in that it allows the management to undertake long-term projects, it may lead to some complacency at a time when cost cutting and restructuring are the necessary requirements for companies to survive in a fiercely competitive environment.

The Chemical Industry at the Beginning of the Third Millennium

The general trends described at the end of the previous chapter have been amply confirmed over the last 10 years. Regulatory burdens have multiplied, especially environmental ones, causing higher costs to the industry and yielding either no benefits or negative consequences. Globalization has proceeded at a faster pace, and as products have matured and technologies have become more readily available, competition between chemical firms has been fiercer on the international scene. The expansion of stock markets, with the increased interest paid to profitability by financial analysts, pension funds, managers, and individual investors in the Western world, has given prominence to the concept of “share value.” The management of chemical companies on both sides of the Atlantic could not remain indifferent to these trends and reacted in several ways to the heavier pressure they implied. In an effort to meet regulatory expenditures and at the same time remain competitive while pleasing their shareholders, Western chemical groups had to resort to severe cost-cutting measures.

Except in such fields as biotechnology R&D budgets were made smaller, and as a result organic growth was somewhat sacrificed, with fewer innovative products coming to the market. Preference was given instead to growth by acquisition, which offered companies an immediate way of achieving a better rank and a higher turnover in some selected activities, if not always better profits. More generally, short-term considerations took precedence over long-term ones, and chemical groups went into a frenzy of asset transfers through mergers, acquisitions, divestitures, and joint ventures. Since 1994 the total value of merger and acquisition deals across the chemical industry in the United States and in Europe has reached a figure of $200 billion, and there is no end in sight to this restructuring fever, greatly facilitated by the ample availability of funds. These moves provided an opportunity for new players to enter the game.

Individual entrepreneurs not accountable to outside shareholders tried their luck by acquiring at rebate prices businesses that larger companies did not want to keep anymore. Private equity institutions bought pieces of business that were for sale with the hope of combining and rationalizing them before returning them to the market 3–5 years later through an initial public offering. Financial services groups played their role either by financing some of the deals or by acting as brokers between the parties. While profitable to some individuals or banking institutions, these transfers of assets did not always meet the expectations of those who had engineered them or for that matter of the new shareholders.

Indeed the creation of pure play companies focusing on selected fields in which they come to command leading positions cannot achieve the required results if the remaining competitors keep fighting each other in order to maintain or even to improve their market share in such a way that the benefits of the consolidation are lost through the erosion of selling prices. When the number of suppliers is reduced, there is often a corresponding reduction in the number of customers, and the pressure on prices remains unabated. Furthermore, when acquisitions are purchased at too dear a price, they affect the purchaser’s balance sheet by denting the bottom line and lowering the operating profit. Aside from these facts that can be assessed quantitatively, there is the less visible aspect of mergers and acquisitions, which has to do with the “morale” of the employees involved in the deals.

Sometimes at all levels valuable people do not like the new “environment” in which they would have to operate, so they decide to leave, offering their talents to competition and thereby weakening the newly born organization. Some chemical groups acting as contrarians have opted for a more stable course in the way they shape their future. Although they are prepared to take advantage of opportunities in order to enlarge or streamline their activities, they keep their diversified operations under one roof, making use of existing synergies and managing their assets with a medium- to long-term objective. It is to be hoped that these contrarians can remain successful because the chemical industry is too vital to our well-being to be submitted entirely to the whims of short-term financial considerations.

Safety Considerations in the Chemical Process Industries

2

John F. Murphy

Introduction

Why Process Safety Is Important

There is an ongoing emphasis on chemical process safety as a result of highly publicized accidents such as the recent BP Deep Water Horizon well blow out in the Gulf of Mexico that resulted in a fire and explosion that killed 11 people and a massive leak of oil that caused catastrophic damage to the environment and economy of the Gulf Coast. Public awareness of these accidents has provided a driving force for industry to improve its safety record. There has also been an increasing amount of government regulation.

The chemical industry is one of the safest industries, but its safety record in the eyes of the public has suffered. Perhaps this is because sometimes when there is an accident in a chemical plant it is spectacular and receives a great deal of attention. The public often associates the chemical industry with environmental and safety problems, which results in a negative image of the industry.

So why is process safety important? It is important because good process safety performance, the lack of major process safety incidents, allows a company the freedom to manage its business without the interference of government regulators, litigation, and adverse public opinion. By avoiding injuries to people, major property loss, and business interruption loss, process safety results in the creation of positive business value for a company. The actions that are required to manage process safety well are the same actions required to manage business well.

Occupational Safety Versus Process Safety

It is important to differentiate between occupational safety which involves accident prevention through work systems which are aimed at minimizing the risk of injury to workers and process safety which involves the prevention and mitigation of fires, explosions, and accidental chemical releases that can have far reaching impacts. Occupational safety focuses on the prevention of worker injuries and occupational illness, primarily relating to trips, slips, falls, cuts, burns, etc. These injuries result from the failures in the control of traditional work procedures. Process safety focuses on the prevention of leaks, spills, process upsets, toxic releases, and equipment failures which may or may not injure or result in fatalities to workers or others at or near the site. This chapter deals primarily with process safety.

Process Safety Technology Issues

The Internet provides considerable information on incidents, good industry practice, and design guidelines. The best practices in industry are briefly discussed in this chapter. Details are readily available from resources listed in the references section at the end of the chapter. Hazards from combustion and runaway reactions play a leading role in many chemical process accidents. Knowledge of these reactions is essential for the control of process hazards. Much of the damage and loss of life in chemical accidents are caused by a loss of containment that results in a sudden release of hazardous material at high pressures, which may or may not result in fire; so it is important to understand how loss of containment and sudden pressure releases can occur. Loss of containment can be due, for example, to ruptured high pressure tanks, runaway reactions, flammable vapor clouds, or pressure developed from external fire. Fires can cause severe damage to people and property from thermal radiation. Chemical releases from fires and pressure releases can form toxic clouds that can be dangerous to

J.F. Murphy (✉)
Process Safety Services, Punta Gorda, FL, USA
e-mail: hamjfm@embarqmail.com

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_2,

people over large areas. Static electricity often is a hidden cause of accidents. It is very important to understand the reactive nature of the chemicals involved in a chemical facility.

Process Safety Management Issues

Chemical process safety involves both the technical and the management aspects of the chemical industry, and this chapter addresses both. It is not enough to be aware of how to predict the effect of process hazards and how to design systems to reduce the risks of these hazards. It also is important to consider how chemical process safety can be managed. Technical and management people at all levels in an organization have process safety management responsibility, and can contribute to the overall management of safer chemical processing plants.

Loss of containment due to mechanical failure or misoperation is a major cause of chemical process accidents. The publication *One Hundred Largest Losses: A Thirty Year Review of Property Damage Losses in the Hydrocarbon-Chemical Industry* [1] cites loss of containment as the leading cause of property loss in the chemical process industries.

It has become clear that process safety can be and must be managed as any other part of the business. A process safety management system is focused on preparedness for the prevention and mitigation of catastrophic releases of chemicals or energy from a process associated with a facility. It also includes the response to and restoration from these events. The term process safety management was first recognized on a broad scale in the late 1980 after Bhopal (see case histories). It formed the basis for many of the American Institute of Chemical Engineers' Center for Chemical Process Safety's guideline books and eventually led to US regulations (OSHA PSM) in 1992.

Barrier Analysis and Layers of Protection [2]

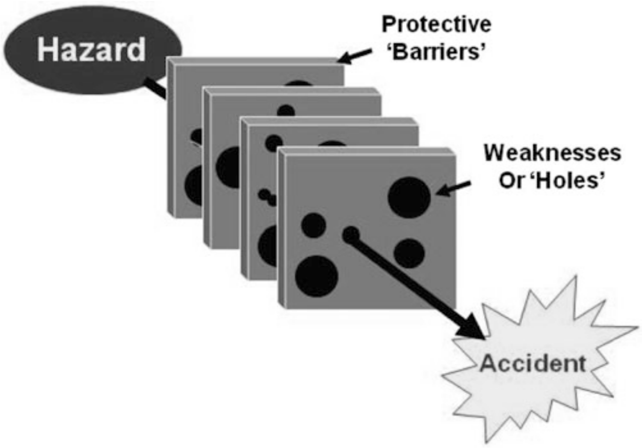


Fig. 2.1 Barrier analysis

There are certain concepts that people and companies have found useful in preventing and minimizing process safety incidents. The US Department of Energy has published a comprehensive and useful report on barrier analysis methodology (Document EH-33, Office of Operating Analysis and Feedback, 1996). The first concept is that of the use of barriers (see Fig. 2.1). Each of these cards represents pieces of Swiss cheese such that when the holes line up an incident will result. The objective of process safety management is to remove and/or minimize the sizes of the holes so that the hazard cannot propagate and become an incident. These barriers include systems for prevention, mitigation, and recovery. Examples of prevention barriers include control systems, procedures, alarms, and maintenance. Examples of mitigation barriers include dikes and containment, facility siting, gas detectors, and fire protection systems. Examples of recovery barriers include medical capability, mutual aid, spare part systems, and insurance.

Another way of looking at barriers is the use of an onion skin model shown in Fig. 2.2. The layers of protection include the tanks, vessels, and piping systems; the basic process controls; and various safety systems both preventive and mitigative. This is sometimes also called defense in depth. All of these barriers or layers of projection must be effectively managed.

Process safety is part of every facet of design and operation of a chemical processing facility during its lifetime as illustrated in Fig. 2.3. The risk of chemical processing must be managed at an acceptable level by the application of inherently safer design strategies, risk reduction measures, and risk-based process safety management, all of which will be discussed further in this chapter.

Anatomy of an Incident [3]

The Anatomy of an Incident Model

One definition of process safety is the sustained absence of process incidents at a facility. The anatomy of an incident is a useful model that explains how process incidents occur. The model will be used to logically introduce the technical elements of process safety. Figure 2.4 illustrates the model. To understand the model, several definitions are necessary:

- An incident is an unplanned event or sequence of events that either resulted in or had the potential to result in adverse impacts.

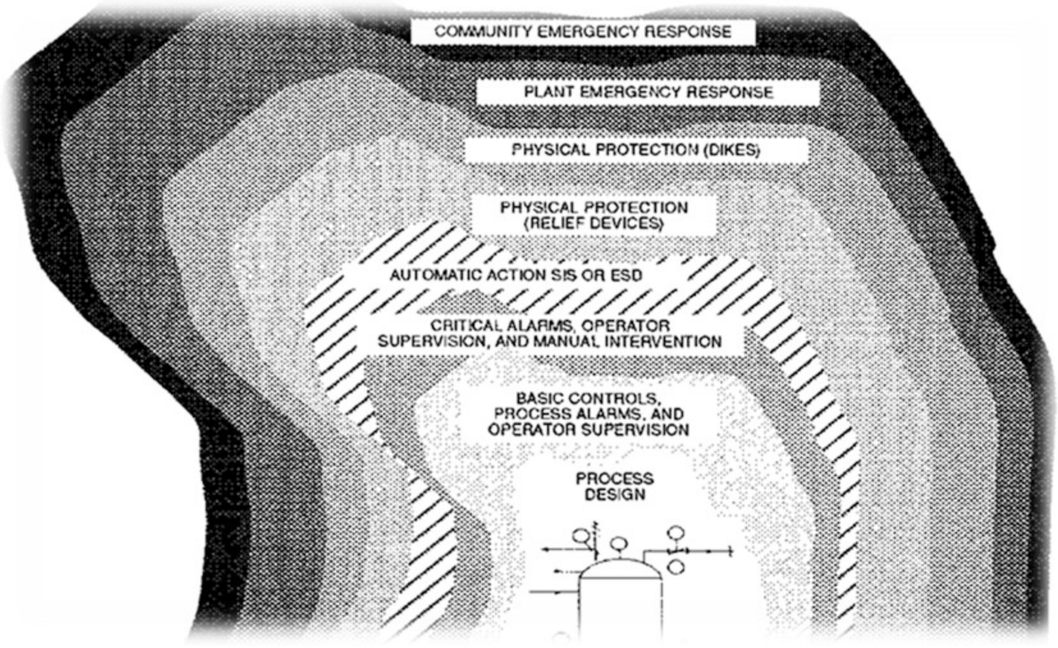


Fig. 2.2 Layers of protection (Copyright 1993 Center for Chemical Process Safety (CCPS) of the American Institute of Chemical Engineers (AIChE), www.aiche.org/ccps,and used with permission)

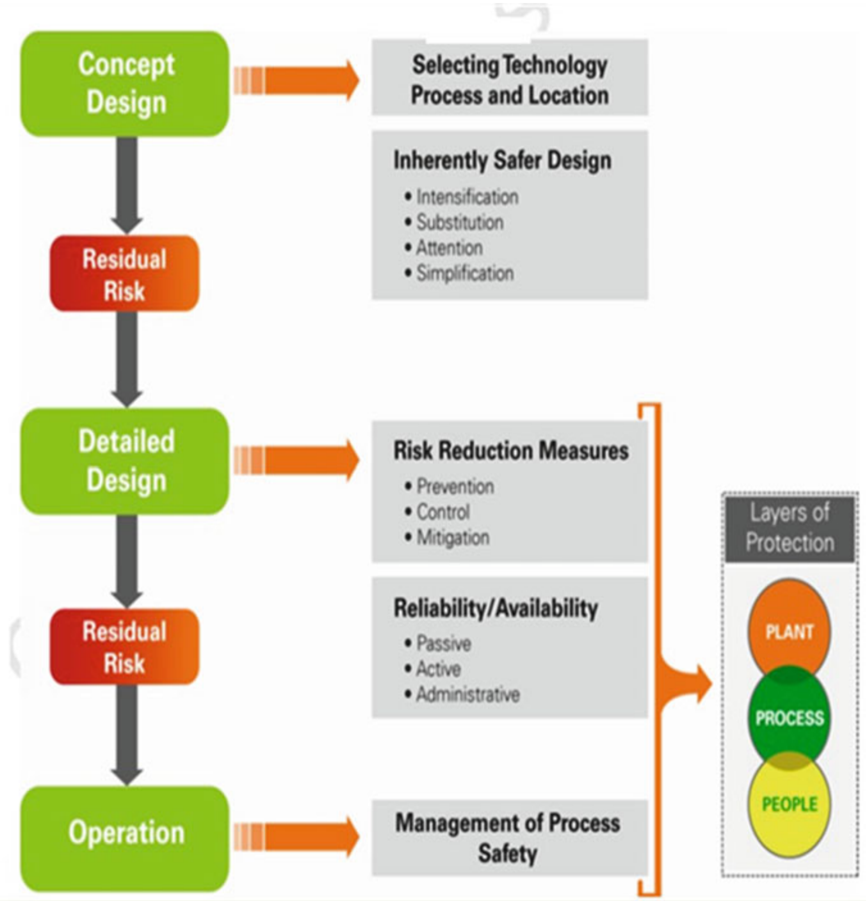


Fig. 2.3 Process safety key principles (Copyright 2010 Center for Chemical Process Safety (CCPS) of the American Institute of Chemical Engineers (AIChE), www.aiche.org/ccps, and used with permission)

Fig. 2.4 Anatomy of an incident (Copyright 2008 Center for Chemical Process Safety (CCPS) of the American Institute of Chemical Engineers (AIChE), www.aiche.org/ccps, and used with permission)

- *Loss events* will result if a deviation continues uncorrected and the process is not shut down. Loss events are generally *irreversible process material/energy releases* but may also be related to production or equipment failures.
- A process hazard is the presence of a stored or connected material or energy with inherent characteristics having the potential for causing loss or harm.

Without safeguards, a cause (such as a process upset) will result in a deviation (such as an increase in pressure rupturing a tank) that will result in the process hazard (such as a flammable liquid in storage) becoming a loss event (e.g., a leak of a flammable liquid that ignites becoming a pool fire) with subsequent impacts (such as injuries and economic and/or environmental damage). However, in a well-designed and operated chemical processing facility there are numerous prevention safeguards to prevent the loss event from happening and mitigation safe guards to minimize the impacts if there is a loss event.

Loss of Containment [2]

The major thrust of process safety is to prevent and mitigate accidental or unintentional loss of containment of hazardous materials. Chemical plants are designed to avoid this consequence but loss of containment has resulted in all the process safety incidents that have occurred and those that are likely to occur in plants in the future. Good process safety design prevents and/or mitigates all forms of loss of containment. Unexpected releases of hazardous materials can occur as a result of:

- Mechanical failure of the pressure envelope (the pressure/temperature limit of a vessel or pipe).
- Process upset causing overpressure, high temperature or volume increase.
- Human error resulting in the direct release of material to environment.

In some cases the leaks are small to begin with and slowly increase such as a small flange leak which continues to open over time. Other sources, such as a major vessel failure, have maximum leak rates initially and as the pressure of the source diminish the leak rate decreases. The type of release can change as well. A line rupture can result in a jet release which could eventually degrade to a slower continuous liquid release.

Common loss of containment events include:

- Opening a maintenance connection during operation
- Piping failure from corrosion
- Overfilling vessel or knockout drum
- Mechanical failure of a process vessel due to thermal or mechanical shock
- Overpressure due to process upset
- Failure to shut valve after transfer operation
- Hot work within berm ignites vapors from tank
- Leaking roof seams
- Leaking floating roof seal
- Non-uniform mixing of contents—temperature anomaly
- Internal explosion due to violent chemical reaction
- Rupture of furnace tube
- Rupture of tube in heat exchanger
- Failure of an internal baffle due to corrosion
- Vacuum due to various causes
- Excess flow into vent system
- Mechanical impact

Containment and Control [2]

The first layer of protection is to control the process so that it remains within its normal operating conditions. During the operational mode or normal operation, the objective is to maintain normal operation and to keep hazards contained and controlled. Layers of protection that help to attain this objective include the:

- Basic process control system
- Maintenance procedures

- Inspections, tests, maintenance
- Operating procedures
  - Training people in procedures
  - Conducting a procedure or operating a process correctly and consistently
  - Keeping a process within established limits
- Guards, barriers against external forces
- Management of change

The design of the instrumentation is extremely important. The control system addresses process deviations as they occur and has either built in control systems or operator actions to bring the process back into control. Process control systems maintain normal operation with manual operator controls relying on operator procedures and process alarms or automated process control systems utilizing basic process control systems, distributed control systems, and programmable logic controllers.

Initiating Cause or Event [2]

If an initiating cause develops then the control system is the first layer of defense to help bring the process back into control. Some possible initiating events include:

- Equipment malfunctions
  - Pumps, compressors, agitators, valves, instruments, sensors, control systems
  - Spurious trips, vents, reliefs
- Loss of utilities
  - Electricity, nitrogen, water, refrigeration, air, heat transfer fluids, steam, ventilation
- Human errors
  - Operations
  - Maintenance
- External agencies and events
  - Vehicle impact, extreme weather conditions, earthquake, knock-on effects, vandalism/sabotage

The safeguards and alarms built into the basic process control system and/or operator actions will attempt to bring it back under control. Many protective features are built into basic process control systems. However, these are not truly independent safety systems. If control is lost because of some instrument failure, that same instrument will normally not be able to function as a safety device. In a basic process control system there is usually no independence between control and protection because they are all controlled by the same control function.

Prevention of Loss of Containment Events [2]

Preventive safeguards come into play once an initiating cause results in an abnormal situation that cannot be controlled by the basic control system when a process deviation has occurred. Bringing the process into a safe condition requires the application of preventive safeguards. Preventive safeguards normally attempt to shut the process down when other safeguards have been unsuccessful in bringing the process back under control.

Safety Instrumented Systems [2, 4]

Typically Safety Instrumented Systems or SIS are the normal means employed to shut the process down. Although an operator may be in charge of initiating shutdown manually, the operator sometimes faces conflicts of interest in deciding whether to let the process operate or shut it down and he may not always be able to reliably assess the data available in a short time. The SIS will try to act such that any loss of containment is averted (Of course the reason the SIS may be shutting the process down is because it was already caused by a loss of containment event!). This includes any loss of containment through relief systems if possible. If the SIS works, then the system will have to go through another startup sequence before it operates again.

There are specific standards that regulate the use of SIS. Testing capability is designed from the beginning. If the SIS can only be tested during shutdowns, there may be a long wait. A safety instrumented system:

- Achieves (or maintains) a safe state of the process
- Is designed and managed per
  - ANSI/ISA 84.00.01-2004
  - *Guidelines for Safe and Reliable Instrumented Protective Systems*, CCPS
  - Other applicable practices
- Requires rigorous management system with respect to inspection, testing, and maintenance to justify risk reduction claim
- Requires that all components of the “system” must be included in the certification of the system (not just the final element) and maintained and tested to justify the risk reduction claim

The key feature of an SIS is that it is totally independent of control. It should have its own independent sensor(s) and final element(s). Because of the additional complexity the plant may endure more spurious trips. Voting systems are sometimes employed to improve reliability while still maintaining the safety levels required.

Redundant Instrumentation and Control Systems [5]

Computer-controlled chemical plants have become the rule rather than the exception. As a result, it is possible to measure more variables and get more process information than ever, and chemical plants can be made safer than ever before. However, it must be kept in mind that instruments and control components *will* fail. It is not a question of *if* they will fail, but *when* they will fail, and what the consequences will be. Therefore, the question of redundancy must be thoroughly considered. *The system must be designed so that when failure occurs, the plant is still safe*.

Redundant measurement means obtaining the same process information with two like measurements or two measurements using different principles. Redundant measurements can be *calculated* or *inferred* measurements. Two like measurements would be two pressure transmitters, two temperature measurements, two level measurements, and so on. An example of inferred measurement would be using a pressure measurement and vapor pressure tables to check an actual temperature measurement.

A continuous analog signal that is continuously monitored by a digital computer is generally preferable to a single point or single switch, such as a high level switch or high pressure switch. A continuous analog measurement can give valuable information about what the value is *now* and can be used to compute values or compare with other measurements. Analog measurements may make it possible to predict future values from known trends. Analog inputs may be visual, and one can see what the set point is and what the actual value is. The software security system should determine who changes set points, and should not be easy to defeat.

A single point (digital) signal only determines whether switch contacts are open or not. It can indicate that something has happened, but not that it is going to happen. It cannot provide information to anticipate a problem that may be building up or a history about why the problem happened. Single point signals are easy to defeat. Some single point measurements are necessary, such as fire eyes, backup high level switches, and so on.

As a rule, it is best to avoid:

- Both pressure transmitters on the same tap
- Both temperature measuring devices in the same well
- Both level transmitters on the same tap or equalizing line
- Any two measurements installed so that the same problem can cause a loss of both measurements

It is a good idea to use devices that use different principles to measure the same variable, if possible.

An alarm should sound if any time redundant inputs disagree. In many cases the operating personnel will have to decide what to do. In some cases the computer control system will have to decide by itself what to do if redundant inputs disagree.

The more hazardous the process, the more it is necessary to use multiple sensors for flow, temperature, pressure, and other variables.

Since it must be assumed that all measuring devices will fail, they should fail to an alarm state. If a device fails to a nonalarm condition, there can be serious problems. If a device fails to an alarm condition, but there is really not an alarm condition, it is also serious, but generally not as serious as if it fails to a nonalarm condition, which can provide a false sense of security.

Usually it is assumed that two devices measuring the same thing will not fail independently at the same time. If this is not acceptable, more than two devices may be used. If this is assumed, one can consider the effects of different levels of redundancy:

Number of Inputs Consequence

One failure provides no information on whether there is an alarm condition or not.

Two failures of one device show that there is a disagreement, but without more information, it cannot be determined whether there is an alarm condition or not. More information is needed; the operator could "vote" if there is time.

Three failures of one device leave two that work; there should be no ambiguity on whether there is an alarm condition or not.

Pressure Relief Systems

If the process cannot be shut down in time, then a relief system may be called into action. Although this is also a "loss of containment," the fluid is discharged to a specific safe location. Sometimes defining a "safe" location is difficult.

The design of relief systems involves, in general, the following steps:

1. Generate scenario. What could reasonably happen that could cause high pressures? This could be fire, runaway reactions, phase changes, generation of gases or vapors, leaks from high pressure sources, and so on.
2. Calculate the duty requirements—the pounds per hour of material that has to be vented, and its physical condition (temperature, pressure, ratio of vapor to liquid, physical properties). This is a rather involved calculational procedure.
3. Calculate the relief area required based on the duty, inlet and outlet piping, and downstream equipment. This is also a rather involved calculational procedure.
4. Choose the relief device to be specified from vendor information.

A group of chemical companies joined together in 1976 to investigate emergency relief systems. This later resulted in the formation of The Design Institute for Emergency Relief Systems (DIERS), a consortium of 29 companies under the auspices of the AIChE. DIERS was funded with $1.6 million to test existing methods for emergency relief system design and to "fill in the gaps" in technology in this area, especially in the design of emergency relief systems to handle runaway reactions [6]. DIERS completed contract work and disbanded in 1984.

Huff was the first to publish details of a comprehensive two-phase flow computational method for sizing emergency relief devices, which, with refinements, has been in use for over a decade [7–10]. The most significant theoretical and experimental finding of the DIERS program was the ease with which two-phase vapor–liquid flow can occur during an emergency relief situation. *The occurrence of two-phase flow during runaway reaction relief almost always requires a larger relief system than does single-phase flow vapor venting*. The required area for two-phase flow venting can be from two to much more than two times larger to provide adequate relief than if vapor-only venting occurs [7]. Failure to recognize this can result in drastically undersized relief systems that will not provide the intended protection.

Two-phase vapor–liquid flow of the type that can affect relief system design occurs as a result of vaporization and gas generation during a runaway reaction or in many liquid systems subjected to fire (especially tanks that are nearly full). Boiling can take place throughout the entire volume of liquid, not just at the surface. Trapped bubbles, retarded by viscosity and the nature of the fluid, reduce the effective density of the fluid and cause the liquid surface to be raised. When it reaches the height of the relief device, two-phase flow results. Fauske and Leung [11] described test equipment that can be used to help determine the design of pressure relief systems for runaway reactions that often result in two- or three-phase flow.

Blow Down Systems/Flare Systems/Incinerators [2]

Relief devices most often discharge into collection/treatment systems called blow down systems. These collection/knockout systems are usually pressure vessels. Treatment systems, usually scrubbers and/or flares/incinerators, are located downstream of the collection/knockout vessels. More blow down systems are being designed to collect relief valve discharges especially two-phase discharges. Collection systems can at times be pretty sophisticated. Retrofit of these systems can be very difficult in placing them within existing process equipment. Some design considerations related to relief systems include:

- Relief capacity based on simultaneous release of several PSVs from a single contingency.
- System approach used to determine limiting scenario.
- Blocked in circuit, fire, or loss of cooling. Which one is quicker acting?
- May not take credit for instrumentation or human intervention.
- System capacity must consider backpressure at all points of entry.
- PSV’s have full capacity at 10% overpressure.
- Step configuration for variable loads.
- Flare systems and incinerators [2].

A flare system is used to collect and burn excess flammable vapors and safely disperse the byproducts to atmosphere. A blow down (piping manifold) system is used to collect surplus vapors from a process. Most releases occur under emergency or upset (unplanned) conditions and are directed from the outlet of one or more relief valves. Entrained liquid is removed from the vapors (knock out drum) and the vapors are burned in a central stack. The flare stack is usually elevated to reduce thermal radiation effects on the ground and equipment and also to disperse smoke.

An incinerator is a waste treatment device that involves the combustion of organic materials in the solid or gaseous state. It is usually a direct fired piece of equipment where feedstock undergoes combustion. Processes may or may not recover the heat released from combustion.

Mitigation of Loss of Containment Events [2]

Mitigative safeguards limit the extent of the loss event. Mitigative safeguards include both physical as well as administrative components. As the incident progresses, mitigation is successful or unsuccessful. Mitigative safeguards include:

- Isolation of piping systems and equipment
- Detectors and alarms
- Flame arrestors/suppression systems
- Explosion containment rooms
- Fire protection
- Water or steam curtains
- Emergency response

Design for Emergency Isolation of Piping Systems [2]

Design of piping systems with its associated equipment must include provision for safe and rapid isolation of the contents of the system should the need arise. This can be accomplished by strategic placement of emergency block valves (EBVs). EBVs are typically located at

- Loading/unloading lines in hazardous service
- Furnace crack gas outlet lines
- Inlet and outlet of compressors
- Inlet and outlet of reactors
- Inlet of pumps from vessels with 10,000 lb of flammable material
- Major lines entering a system of vessels containing more than 10,000 lb of flammable materials which operate together
- Battery limits for pipelines containing hazardous materials

Through proper design, it would be possible to provide for isolation of an individual item or section, or to effectively shut down an entire operation. The location of block valves must include consideration of potential hazards to operating personnel. Therefore, manually operated devices must be easily and safely accessible and easy to operate. Remotely or automatically operated block valves, usually recommended in situations where operating personnel would be exposed to hazardous conditions, should be placed in locations where, insofar as possible, they will not be subject to conditions which might impair their operation. If there is a possibility that they might be exposed to fire in the case of an emergency, block valves should be fireproof or fire resistant to the extent necessary to ensure dependable operation under such conditions. It should be apparent that proper use of block valves has the potential to greatly reduce loss of containment with its resultant explosions and fires because they.

- Provide isolation between different hazards within a system
- Quickly interrupt flow through a system or prevent gross movement of hazardous material into an exposed location
- Block in specific pieces of hardware that may be involved in an incident
- Cause an orderly shutdown of equipment

Categories of EBVs include:

  - Manually actuated valve at the equipment to be isolated
  - Manually actuated valve located at a safe distance from the equipment to be isolated
  - Mechanically assisted valve at or near the equipment to be isolated
  - Mechanically assisted valve with remote activation capability

Gas Detection [2]

Gas detection is used to determine the presence of undesired vapors and gases at some specified concentration. It is also used to support some action or decision. Sensors need to be located where gas is most likely to accumulate. With gas detectors, the concentration of test gas at point of measurement at a specific time is known, but how much gas is present, how far the gas cloud extends, the concentration profile within the cloud, what other gases are present, or how fast the gas is moving are not known.

Flame/Detonation Arrestors [2]

There are two types of arrestors: flame arrestors and detonation arrestors which are more massive and robust. The arrestor forces the gas velocity to decrease through increasing the diameter and then provides sufficient surface to cool the gas. If the upstream fire continues eventually it could overwhelm the arrestor as the internal parts heat up. Flame arresters absorb kinetic energy from a fluid and prevent deflagration/detonation transition. Common types include crimped ribbon, parallel plate, expanded metal, packed bed, hydraulic, perforated plate, sintered metal, and liquid seal. Criteria for design include vapor composition, operating pressure and temperature, piping configuration, and flow rate. All flame arresters must be tested in their final environment.

Explosion Venting, Mechanical Isolation, and Explosion Suppression [12]

Explosion vent panels are pretty common especially for dust applications. In dust applications there is a possibility of a buildup of solids on the inside panels which can add weight and resistance. Some materials such as hydrogen can't be vented using explosion panels because of its high flame speeds. In general these systems are designed to discharge in an open environment. The panels are designed per NFPA 68. They are lightweight, tethered panels which are non-fragmenting.

Fast acting valves are another means of stopping flame propagation. They are dependent on very quick acting sensors that are able to detect change of pressure in the millisecond range.

Suppression systems such as halons cause the fire to starve because air is displaced by an inert. Having a space that is effectively sealed to minimize leakage is very important.

Certain types of foam or plastic media can be applied to a liquid spill in a dike to restrict vaporization.

Explosion containment rooms are used in special applications, particularly high pressure operations. They are also used in laboratories and pilot plants, sometimes with blow out walls and roofs.

Fire Protection [2]

Fire protection must be a functional, rational, and consistent system across a facility. A "holistic" fire protection strategy must be adopted to deal with all possible exposures that could threaten people, equipment, the operation, and the environment. Compliance to codes and standards is a minimal requirement. There must be an integrated fire protection strategy that includes:

- Prevention: Process selection, equipment layout, good engineering, quality construction
- Detection and response: Isolate fuel source
- Suppression: Prevent fire from spreading—extinguish original fire

All of these activities are essential to success. Obviously the first consideration is to prevent as many opportunities for loss of containment of flammables and possible fires as possible. Foam systems are sometimes necessary to extinguish fires. Fire protection includes:

- System for ensuring the early detection, extinguishment, and protection from fires
- Deluge protection recommended in critical areas such as pump bays
- Deluge systems should be easy to operate and should be augmented with detectors
- Heat actuated sprinkler systems
- Extinguishment in motor control centers and control rooms
- Fireproofing on high value equipment (2–4 h)

A typical fire water distribution system is usually installed underground. Valves are placed so that water supply can get to any point from two directions within the loop. A 4 h supply of firewater is usually specified as recommended by insurance companies. The pumping system should be redundant with both electric and diesel backup pumps. With fire protection systems it is important to ensure the ongoing integrity of all fire protection hardware by preventive maintenance and testing. Ownership must be clearly established to ensure that proper testing and maintenance is done on fire protection systems.

All operating personnel must be familiar with the functionality of fire protection equipment that they may be required to use.

Emergency Response [2]

Emergency response involves scenario-based planning. It requires understanding the hazards and risks and on-site response capabilities. It also requires securing support from outside parties, for example, fire departments, and mutual aid organizations. Emergency response also requires the establishment of safe havens, evacuation routes, command centers, and the development of emergency response procedures. Emergency response training includes emergency drills that must be conducted. These drills include tabletop exercises and full-scale drills. Feedback from drills should result in improvement and modification of emergency procedures as necessary. See Chap. 3 for a more detailed discussion of emergency response.

Management Systems (Risk-Based Process Safety) [2, 13]

Management systems are necessary to ensure that process safety concepts and practices are implemented in an organization. There are several models for process safety management systems. The OSHA process safety management model is a legal requirement in the United States (see section on “Regulations” in this chapter). Risk-Based Process Safety is the model developed by CCPS. It is not a regulatory requirement but is consistent with the OSHA PSM standard and the chemical processing industries’ good practice. The recent book entitled *Risk Based Process Safety,* CCPS, 2007 discusses the model proposed by CCPS that will be used in this discussion of process safety management systems. The concept of Risk Based Process Safety is that each company needs to tailor its process safety management systems to address its risk and risk tolerance. In other words the higher the risk of a major process safety incident, the more robust your management system should be. Also companies with the same risk of a major process safety incident may have different management systems because of different risk tolerance. For example, a company that has a chlorine unloading operation adjacent to another business or residential area should have a more robust management system than a company with a similar unloading operation located in a remote area. Of course hazardous operations need to be located in as remote areas as possible.

CCPS defines process safety management as a management system that is focused on the prevention of, preparation for, mitigation of response to and restoration from catastrophic releases of chemicals or energy from a process associated with a facility. The key concept is “catastrophic” releases of chemicals or energy from a process associated with a facility. A process is a sequence of activities that leads to a desired outcome. That outcome often involves the creation of a marketable product. In practice, processes of concerned include the shipping, handling, storing, mixing, separating, and disposal of chemicals. Risk-based process safety uses risk-based strategies and implementation tactics that are commensurate with the risk-based need for process safety activities, availability of resources, and existing process safety culture. Based on the company’s perception of risk, each element of RBPS needs to be designed and implemented to fit the risk. Each company will have different looking management systems based on their risk perception and culture, but all 20 elements need to be addressed. The goal of RBPS is to design, correct, and improve process safety management activities.

The 20 elements of RBPS are organized under the pillars “commit to process safety,” “understand hazards and risk,” “manage risk,” and “learn from experience.” Each of the 20 elements under the four pillars will be discussed (see Fig. 2.5).

Commit to Process Safety

The first pillar of RBPS is to *Commit to Process Safety*. To commit to process safety, facilities should focus on:

- Developing and sustaining a culture that embraces process safety.
- Identifying, understanding, and complying with codes, standards, regulation, and laws.

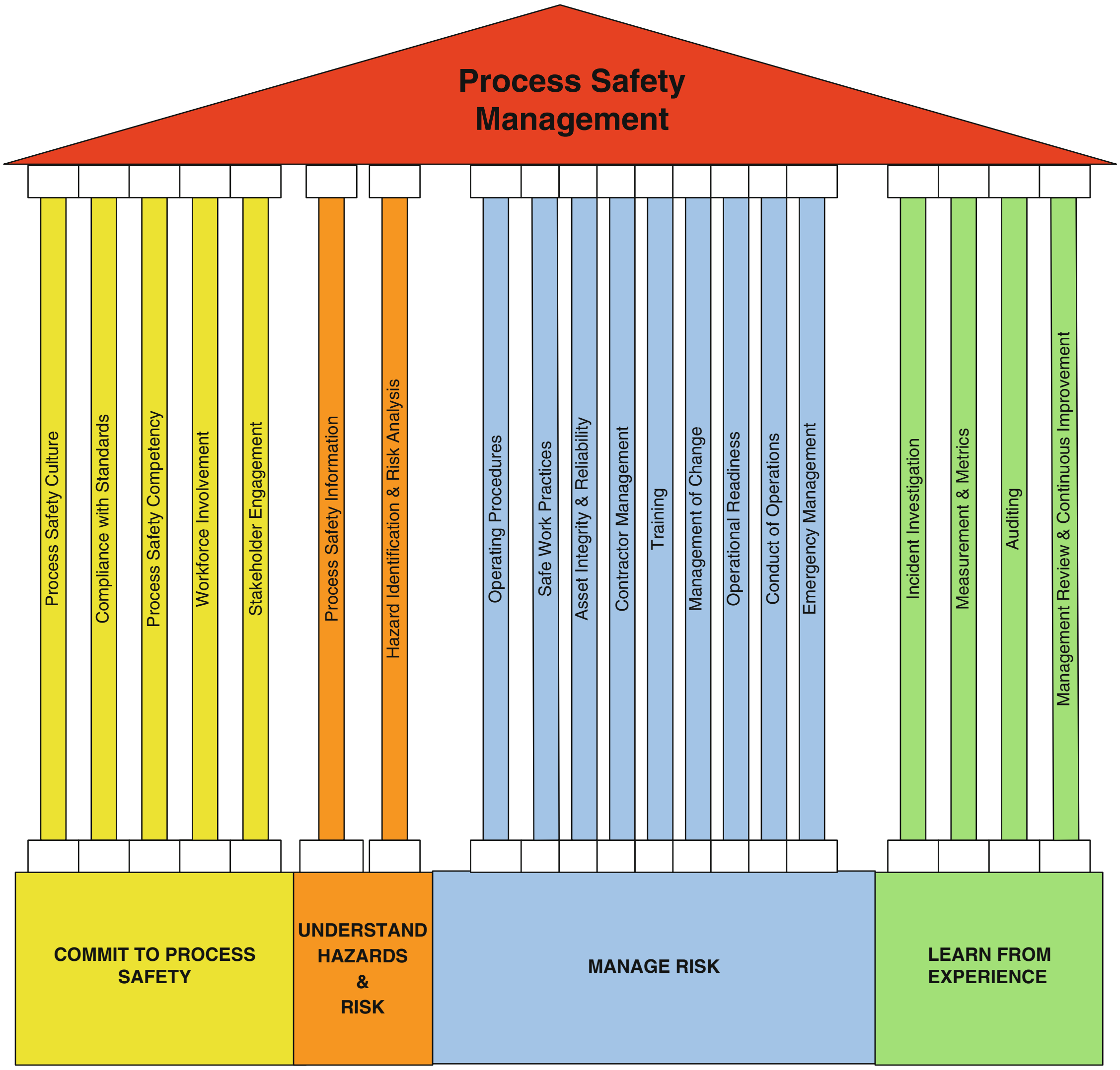


Fig. 2.5 Risk based process safety (David Guss, Nexen Inc. (2008), from CCPS, Process Safety Boot Camp Course (2010) and used with permission)

- Establishing and continually enhancing organizational competence.
- Soliciting input from and consulting with all stakeholders, including employees, contractors, and neighbors.

Process Safety Culture

What is process safety culture? Process safety culture is the combination of group values and behaviors that determine the manner in which process safety is managed. It is often described as how we do things around here, what we expect here, or how we behave when no one is watching. Why is process safety culture important? Investigations of catastrophic events have identified common process safety culture weaknesses that are often factors in other serious incidents. Examples of process safety culture weakness include warning signs such as lack of enforcement of process safety standards, no sense of vulnerability, poor communications, and delayed or no response to process safety issues and concerns.

Human Factors [2]

When examining process safety culture, it is important to understand human factors. This is a very important subject that people sometimes have difficulty in understanding. It encompasses a broad range of topics. Human errors will happen and must be eliminated or the probability or consequences of those errors must at least be reduced when handling hazardous materials. The perception of the application of human factors as ill-defined, difficult to apply, and expensive is a misconception. Those companies that have actively applied human factors have seen a true business value. Human factors look at the ways to identify and control potential human errors and conditions which affect the outcome of human activity. The human is the ultimate variable in every human–machine interface. The potential for human error exists wherever there are several choices or degrees of freedom. There are three groupings of human interactions that all overlap and are intertwined with one another. These are people, facilities, and management systems. The potential for human errors in all these groupings must be examined. The CCPS book *Guidelines for Preventing Human Error in Process Safety* by Center for Chemical Process Safety (CCPS), August 2004, is an excellent source of information on human factors.

Compliance with Standards

The next element under the pillar "Commit to Process Safety" is "Compliance with Standards." A standards system is a system to identify, develop, acquire, evaluate, disseminate, and provide access to applicable standard, codes, regulations, and laws that affect process safety. It addresses internal and external standards; national and international codes and standards; and local, state, and federal regulations and laws. It interacts with every RBPS management system element.

"Recognized and generally accepted engineering practices" or RAGAGEP is a regulatory term that means the engineering practices that are prevalent in the industry must be followed. It is included in the Process Safety Information element as applied to equipment used in the process in the OSHA PSM standard. A standard system as discussed above will help to ensure that RAGAGEP is implemented in the facility.

Why is a standards system important? Knowledge of and conformance to standards helps a company operate and maintain a safe facility and to consistently implement process safety practices. It also minimizes legal liability. Changes in standards must be current so the company can adjust its compliance activities. The standards system also forms the basis of the standards of care used in an audit program to determine management system conformance.

Inherently Safer Concepts [2, 14]

One important accepted practice is inherently safer design. The concept of inherent safety was first coined by Trevor Kletz. Trevor Kletz worked for ICI as a safety consultant and is still active in process safety today. Recently CCPS has updated a guidelines book on inherently safer chemical processes (see references). Inherently safer is a principle that continues to be important in the reduction of overall risk for any company.

In the narrow definition, inherently safer designs permanently and inseparable *reduce or eliminate process hazards* that must be contained and controlled to avoid loss events. To quote Trevor Kletz "The essence of the inherently safer approach to plant design is the *avoidance of hazards* rather than their control by added-on protective equipment" [15].

Process safety strategies include:

- Inherent (hazard elimination or reduction)
- Passive (process or equipment design features that reduce risk without active functioning of any device)
- Active (engineering controls)
- Procedural (administrative controls)

The above is the hierarchy of strategies to control hazards or risk. Least reliable is an administrative approach. From a reliability stand point, engineering controls follow then passive controls such as dikes. Finally, inherent safety is the most reliable. In practice all four strategies are used to design and operate a safe chemical processing facility.

Sometimes what appears to be an inherently safer approach creates its own hazards and risks which may result in lower risk overall but it will not reduce risk to zero. Sometimes the frequency of a loss of containment will be increased by the change. For example, supplying chlorine, a hazardous material, from many small cylinders instead of one large tank increases disconnect frequency, thus the chance of a leak, but the consequences of a release if a leak should occur will be significantly reduced because the quantity release would be less. Another example is a reactor that underwent runaway reaction would have produced much less consequences if it had failed at only a few pounds pressure. Instead the vessel was designed to 50 psig, to decrease the likelihood of a release and when it failed the pressure rise was exponential producing a much greater amount of stored energy and consequently resulted in much greater damage.

Inherently safer principles put forth by Trevor Kletz include:

- Use less hazardous materials
- Minimize inventories
- Reduce operating severity
- Simplify equipment design

The implementation of these principles will result in serious process incidents occurring less frequently. Examples of the application of the principles include:

Intensification: Reduce quantity of hazardous materials.

Substitution: Use of safer materials.

Attenuation: Running at safer operating conditions.

Limitation of effects: Changing equipment layout to reduce consequences.

Simplification: Avoidance of multiproduct operations.

Error tolerance: More robust equipment to tolerate upsets and errors.

Avoid knock-on effects: Open construction, layout.

Prevent incorrect assembly: Piping systems to reduce human error.

Ease of control: Less hands on control. It is extremely important to apply inherently safer principles at the laboratory stage or during the process conceptual design. The earlier inherently safer design concepts can be applied, the easier implementation will be, although, the principles can be applied throughout the life of a plant.

Process Design [2]

Standards and practices are an important aspect of process design. There are many different process safety topics that are part of design. Some are addressed very early in the design process such as layout and spacing, while others are done during detailed design, e.g., bonding and grounding. Some of these safeguards are passive, e.g., grading and drainage and some require engineering control, e.g., system isolation.

The design topics covered briefly in this chapter include:

- Layout and spacing
- Infrastructure
- Grading and drainage
- Equipment selection/sizing
- Design for pressure protection
- Design for mechanical integrity
- Fireproofing/firewalls
- Inerting
- Electrical area classification (EAC)
- Bonding and grounding system isolation
- Review of design alternatives

Process design is a key element in regard to how to contain and control potential loss of containment events.

- *Layout and spacing*

Layout and spacing are critical to safe design. Congestion and confinement are major contributors to high pressures from vapor cloud explosions. Layout and spacing concepts are much easier to apply with a new grass roots facility, but with an existing facility debottlenecking seems to be more prevalent and the job much more difficult.

An I, L, or H pattern provides ease of access and minimal congestion. Layout involves determining the most logical way of configuring the plant given various constraints from the site plot. Every additional foot of piping (and valves) increases the risk of failure. Important strategies include:

- Unit layout in I, L, or H pattern preferred
- High value equipment separated from high hazards if possible
- Major inventory sources well spaced and separated
- Minimize amount of congestion and confinement
- Egress, maintenance and fire fighting access
- Arrange equipment by function
- Arrange piping by category of service
- *Grading and drainage*

Equipment containing large volumes of flammable materials should be located upstream of sloped areas. The rate and quantity of firewater should also be considered in the design of the sewer system or catch basin. The grading plan must be completed prior to plant layout. Key points regarding grading and drainage include:

- Grading should slope away from fuel sources or critical equipment
- Sewer system should be sized to handle storm runoff or maximum release from major equipment failure
- Multiple catch basins reduce the travel time/distance and reduce the surface area for a potential fire
  - Remember sewers are costly; catch basins are not
- Need to segregate incompatible materials
- Sewer design must recognize plugging potential; provide means of clearing
- Sewers must not permit the passage of flammable vapor
  - Use traps
- *Integrated approach to layout, grading, and sewers is required*
- Equipment sizing

Larger vessels are more adiabatic and self-heating and unwanted reactions more difficult to control. Sometimes mixing becomes more difficult in large reactors. Some factors to consider in equipment sizing include:

- Determined by inventory requirements and holdup considerations
- Driven by economics but limited by transportation systems
- Large size equipment is more prone to failure from localized stresses
- Difficult to establish control within all parts of large systems
- Difficult to establish equilibrium in large vessels and reactors

The advantages of small equipment include:

- Reduce hazardous inventory
- Reduce size to save capital and reduce maintenance
- Reduced size gives quicker response—more predictable behavior
- Lower internal stresses

The advantages of larger equipment include:

- Increase size to cushion equipment against upset
- Increased size may reduce number of procedural steps

These are some of the considerations that must be examined. Sometimes goals of the plant and business area are in conflict. The business area wants to have a large inventory so the customer is always able to be supplied on time. The plant wants to minimize storage to have a more inherently safer plant. A facility is probably less prone to human error in a single large system compared to a group of smaller systems.

- *Pressure protection design*

This topic has already been discussed as a mitigative safeguard.

- *Design for mechanical integrity*

Equipment must be designed for mechanical integrity. The design must tolerate continuous exposure to process fluids at normal and upset operating conditions. Normal modes of failure must be anticipated, i.e., impeller wear, corrosion. What tests are necessary to maintain mechanical integrity and how frequently they must be done needs to be established during design. In many cases instruments, like vibration sensors, are designed and alarmed if the machine goes out of its normal operating range. Pressure vessel cycling is another consideration especially for batch reaction systems. Replacement of systems over time and how that can physically be accomplished must be addressed during design. How critical systems must be tested—on-line and off-line—and how frequently testing will occur must also be addressed.

How relief valve testing will be done and access to relief valves is another design consideration. Key concepts to keep in mind regarding design for mechanical integrity include:

- Commit to quality engineering standards
- Material selection to match conditions
- Design to full range of service conditions
- Design for pressure/temperature cycles
- Minimize temperature gradients/local stresses/vibration
- Quality control during fabrication and construction
- Design for ease of testing and maintenance
- Dimensional tolerances in rotating equipment
- *Fireproofing and fire walls*

  Key aspects of fireproofing include:
- Fire hazard zones must be determined based on credible release scenarios
  - Consider drainage and system inventory
- All load-bearing members which reside in a fire hazard zone must be passively protected for 2 h min per UL 1709
- Must extend fireproofing to highest structural member that supports fuel containing equipment—include cross members
- Include whole structure footprint unless clear break point exists
- Fireproofing materials must have good thermal insulation properties and must resist mechanical impact and erosion
- Concrete is an ideal choice for most applications but it can add considerable weight to a structure
- Mixed cementaceous fireproof materials (such as Pyrocrete 241) may be used if they are properly applied
- Contour or surface application around structural steel is recommended to avoid moisture collecting
- All fireproofing applications are subject to verification and approval by UL 1709

Firewalls are solid barriers that shield equipment (*usually high value*) from direct fire radiation and prevent the spread of fires. Firewalls are generally of concrete construction and provide 2–4 h of fire protection. Full partition firewalls may be used to segregate EAC zones.

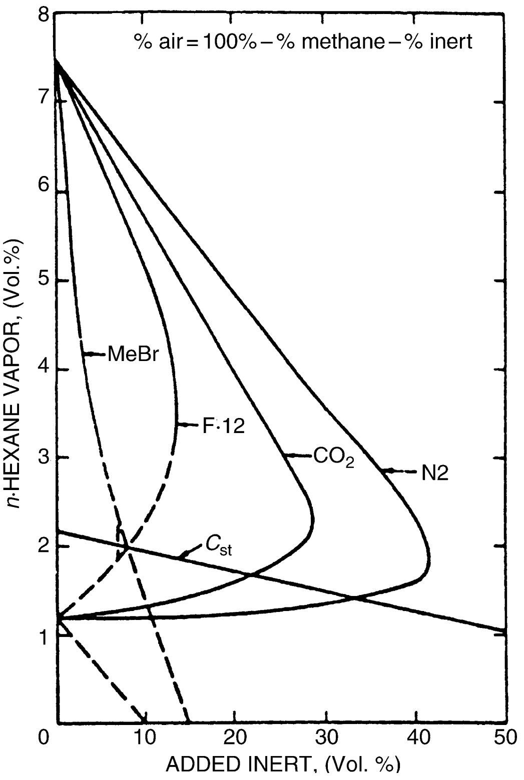


Fig. 2.6 Limits of flammability of various methane–inert gas–air mixtures at 25°C and atmospheric pressure. (Courtesy Bureau of Mines)

Fire stops are barriers installed in pipe racks or conduit runs. They prevent the passage of fuel, air, or fire but permit circuit continuity. Their integrity is a function of design and quality of installation.

Inerting: The addition of inert gases to a mixture of flammable gases and air affects flammability limits. Carbon dioxide causes a greater narrowing of the flammable range than does nitrogen. Water vapor is an acceptable inert gas if the temperature is high enough to exclude much of the oxygen, which requires a temperature of 90–95°C. Because water vapor and carbon dioxide have a higher heat capacity than nitrogen, they are somewhat more effective as inerting agents than nitrogen. Some halogen-containing compounds also can be used for inerting materials at relatively low concentrations. An example of this is the use of Freon-12 (CCl_2F_2). Caution must be used with halohydrocarbons because of the possibility of the halocarbons themselves burning, especially at high pressures. Environmental considerations are making the use of halogenated hydrocarbons for inerting increasingly undesirable. Materials are being developed that are considered environmentally acceptable. Figures 2.6 and 2.7 [16] show flammability envelopes for methane and *n*-hexane for various air–inert mixtures at 25°C and 1 atm. All flammable

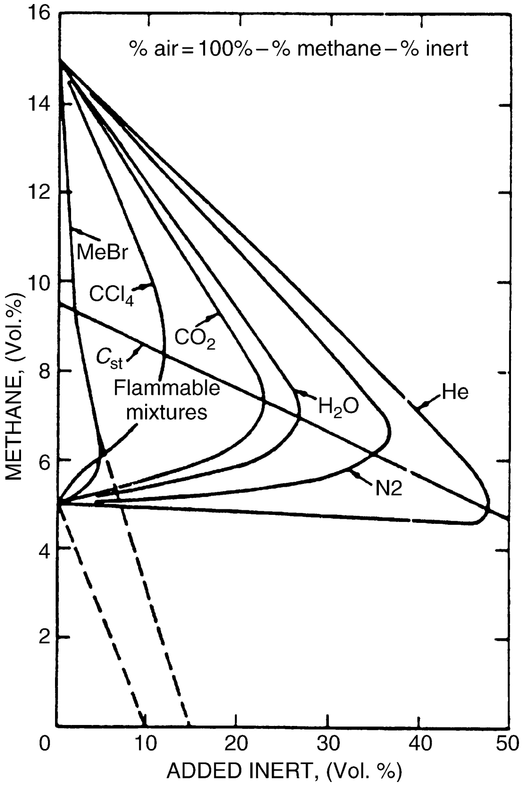


Fig. 2.7 Limits of flammability of various *n*-hexane-inert gas mixtures at 25°C and atmospheric pressure. (Courtesy Bureau of Mines)

envelopes are similar to Figs. 2.6 and 2.7 except in minor detail. The lower limit is virtually insensitive to added inerts. The upper limit, however, decreases linearly with added inert until the critical concentration of inert is reached beyond which no compositions are flammable. In these graphs, C_{st} means the stoichiometric composition.

The limits of flammability are dictated by the ability of a system to propagate a flame front. Propagation does not occur until the flame front reaches about 1,200–1,400 K. Since the typical terminal temperature for hydrocarbons at stoichiometric conditions is about 2,300 K, it can be seen that having only one-half the fuel or oxidizer present will produce about one-half the flame temperature, which is too low to propagate flame.

A useful rule to remember is that the lower flammable limits of most flammable vapors are close to one-half the stoichiometric composition, which can be calculated. Another easy rule to remember is that about 10% oxygen or less in air (assuming the rest is mostly nitrogen) will not support combustion of most flammable hydrocarbon vapors.

The flammability limits of hydrocarbon-type fuels in oxygen and inert gas atmospheres are a function of the inert gas and any fuel or oxygen in excess of that required by the stoichiometry of the combustion process. In systems where fuel content is fixed, inert material having a high heat capacity will be more effective at flame suppression than inert material having a low heat capacity.

Many of the flammable limits reported in the literature are somewhat too narrow, and certain gas compositions regarded as being nonflammable are in fact flammable when given the proper set of circumstances. *In other words, take data on flammability limits from the literature with a grain of salt.* It is best not to design closely on the basis of most available data on flammability limits.

The use of inert gases can cause some serious hazards that must be recognized if inerts are to be used effectively and safely. Considerations in the use of inert gases include:

1. *An inert atmosphere can kill if a person breathes it*: Precautions should be taken to ensure that personnel cannot be exposed to the breathing of inert atmospheres.
2. *Some products need at least a small amount of oxygen to be stored safely*: This includes styrene and some other vinyl monomers, which must have some oxygen in them to make the usual polymerization inhibitor for styrene (t-butyl catechol, or TBC) effective. If pure nitrogen, for example, is used to blanket styrene, the inhibitor will become ineffective. TBC customarily is added to styrene monomer to prevent polymer formation and oxidation degradation during shipment and subsequent storage; it functions as an antioxidant and prevents polymerization by reacting with oxidation products (free radicals in the monomer). If sufficient oxygen is present, polymerization is effectively prevented (at ambient temperatures); but in the absence of oxygen, polymerization will proceed at essentially the same rate as if no inhibitor were present. The styrene may polymerize and can undergo an uncontrolled exothermic reaction, which may generate high temperatures and pressures that can be very hazardous. The inhibitor level of styrene must be maintained above a minimum concentration at all times. The minimum concentration of TBC in styrene for storage is about 4–5 ppm.
3. To be effective, inert atmospheres must be maintained within certain composition limits. This requires the proper instrumentation and regular attention to the system.
4. *Inerting systems can be quite expensive and difficult to operate successfully*: Before the use of inert systems, alternatives should be explored, such as the use of nonflammable materials or operating well outside, preferably below, the flammability range.

Electrical Area Classification [2]: One of the first exercises during design between the electrical engineers and the process engineers results in an electrical classification plot plan diagram. Vehicular traffic in the plant is another source of ignition that should be considered in layout. A plot plan drawing with electrically classified buildings, rooms, etc. is usually developed. More detailed drawings may be needed for special process equipment and discharges. EAC is a system used to control potential electrical ignition sources in close proximity to flammable substances. EAC is usually done by process engineers in conjunction with electrical engineers. It determines electrical hardware which may be used and where it may be located. Non-sparking equipment is defined. Explosion proof enclosures provide small confined enclosure over electrical contacts. Sealing devices must also be specified.

There will be different electrical system requirements depending on the type of material and whether it is a flammable gas or dust. The Zone system is replacing the US Division system.

- Class I—Flammable gas vapor or liquid is present
- Class II—Combustible dust is present
- Group A—Acetylene
- Group B—Hydrogen
- Group C—Cyclopropane, ethyl ether
- Group D—Acetone, butane, hexane, natural gas, fuel oil
- Group E—Combustible metals
- Group F—Carbonaceous materials, including coal dust
- Group G—Flour, starch, plastic
- Division 1 (Zone 0 or 1)—Flammable or combustible concentrations exist under normal operating conditions
- Division 2 (Zone 2)—Flammable or combustible concentrations exist under abnormal operating conditions or have a low likelihood of occurrence

Static Electricity, Grounding and Bonding

Introduction

Many apparently mysterious fires and explosions have eventually been traced to static. In spite of the large amount of information about static electricity, it remains a complex phenomenon not often understood and appreciated. Static electricity is a potential source of ignition whenever there is a flammable mixture of gas or dust.

When two different or similar materials are in contact, electrons can move from one material across the boundary and associate with the other. If the two materials in contact are good conductors of electricity and are separated, the excess electrons in one material will return to the other before final contact is broken. But if one or both of the materials are insulators, this flow will be impeded. If the separation is done rapidly enough, some excess electrons will be trapped in one of the materials. Then both materials are "charged." Electric charges can build up on a nonconducting surface until the dielectric strength is exceeded and a spark occurs. The residual charge could ignite flammable mixtures.

The two materials or phases in initial contact may be a single liquid dispersed into drops, two solids, two immiscible liquids, a solid and a liquid, a solid and a vapor or gas, a liquid and a vapor or gas.

The important thing to keep in mind is that whenever there is contact and separation of phases, a charge may develop that could be disastrous. Three conditions must be met before an explosion caused by static electricity can take place:

1. An explosive mixture must be present.
2. An electric field must have been produced due to the electrostatic charge that had been generated and accumulated in a liquid or solid.
3. An electric field must be large enough to cause a spark of sufficient energy to ignite the mixture.

In designing preventive measures, all three factors should be controlled.

Static electricity is essentially a phenomenon of low current but high voltage and high resistance to current flow. A low-conductivity liquid flowing through a pipeline can generate a charge at a rate of 10^{-9} to 10^{-6} A (A). A powder coming out of a grinding mill can carry a charge at a rate of 10^{-8} to 10^{-4} A. At a charging rate of 10^{-6} A, the potential of a container insulated from earth can rise at a rate of 1,000 V/s and a voltage of 10,000 V or higher can readily be obtained in this way.

Several electrostatic voltages and energies commonly encountered are typified by the following examples:

1. A person walking on dry carpet or sliding across an automobile seat can generate up to 5,000 V in dry weather. An individual having a capacitance of 100 pF, a reasonable figure, could generate a spark energy of 1.25 mJ. This is far more than is needed to ignite some flammable vapor–air mixtures.
2. A person can accumulate dangerous charges up to about 20,000 V when humidity is low.
3. A truck or an automobile traveling over pavement in dry weather can generate up to about 10,000 V.
4. Nonconductive belts running over pulleys generate up to 30,000 V. The voltage generated by a conveyor belt can be as high as 10[5 V; the system can in effect act as a Van der Graaf generator.
5. The energy in the spark from an ordinary spark plug is 20–30 mJ.

The capacitance and the energy for ignition of people and of common objects are important. The capacitance of a human being is sufficient to ignite various flammable gas mixtures at commonly attained static voltages.

There are several hazard determinants relating to static electricity

Capacitance: The capacitance of an object is the ratio of the charge of the object to its potential. The capacitance gets larger as the object gets larger. With a given charge, the voltage gets higher as the capacity of the object gets smaller. For a sphere, capacitance is given by

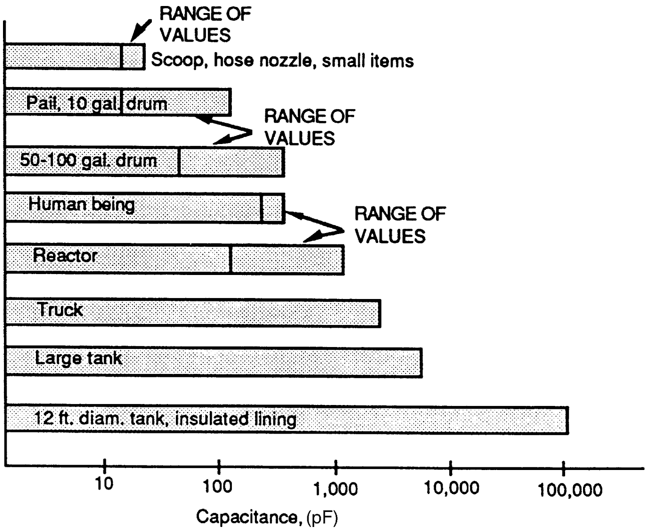


Fig. 2.8 Some typical values of electrical capacitance. (Data from Eichel [21])

$$C = \frac{Q(10^{-3})}{V}$$

The energy stored in a capacitor is [17]

$$W = 0.5\,CV^2(10 - 3) = 500\,Q^2/C$$

where

C = capacitance, pF (1 pF = picofarad = 10^{-12} F)

Q = charge, microcoulombs (1 C = 1 A/s = charge on 6.2×10^{18} electrons)

V = voltage in kilovolts

W = energy, millijoules (mJ)

This energy may be released as a spark when the voltage gets high enough. The minimum sparking potential for charged electrodes is about 350 V and occurs at a spacing of 0.01 mm. Sparks from an equally charged nonconductor are less energetic and may contain only part of the stored energy. These comparatively weak sparks are not likely to ignite dust clouds but can ignite flammable gases (Fig. 2.8).

The energy that can be stored by capacitance of an object can be compared with the minimum ignition energies of flammable gas–air mixtures and of dust–air mixtures to determine the probability that a spark discharge may have sufficient energy to cause ignition. If the charged object is a poor conductor, the calculation of energy available to produce a spark may not be possible because the charge often is not uniformly distributed, and the resistance to flow of current is high. Figure 2.8 shows some typical values of electrical capacitance [18].

Relaxation time: When a liquid is flowing in closed metal pipes, static electricity is not a hazard. When the liquid enters a tank, it may become a hazard. Charges caused by liquid separation during pumping, flow, filtration, and other effects such as splashing and agitation can accumulate on the surface of the liquid in the tank and cause sparking between the liquid surface and the tank or conducting objects in the tank. The charge thus generated can be dissipated by relaxation or via discharge through a spark or corona discharge. The relaxation time is the time required for 63% of the charge to leak away from a charged liquid through a grounded conductive container. The half-time value is the time required for the free charge to decay to one-half of its initial value. The half-time is related to the relaxation time by the relationship

$$T_{\mathrm{h}} = T_{\mathrm{r}} \times 0.693$$

where T_{h} = half-time T_{r} = relaxation time. Relaxation times vary from small fractions of a second up to minutes and even hours for some highly purified hydrocarbons that have very low conductivity.

It is important to recognize that a large charge can accumulate in the liquid even in a grounded container. In fact, it was reported that the majority of accidents attributed to static electricity in the petroleum industry have been with liquid in grounded containers [19].

Relaxation time can be calculated as follows:

$$T_{\mathrm{r}} = E(E_0/k)$$

where

T_{r} = the relaxation time, in seconds; the time for 63% of the charge to leak away

E = relative dielectric constant, dimensionless

E_0 = absolute dielectric constant in a vacuum = 8.85×10^{-14} to less than 1×10^{-18}

K = liquid conductivity, Siemens per centimeter (S/cm)

Siemens (S) are also called mhos

Example: Benzene in a large tank could have a specific conductivity as low as 1×10^{-18} mho/cm and as high as 7.6×10^{-8} S/cm. The corresponding relaxation times for the two conductivities can be calculated as follows. Pure benzene has a dielectric constant of 2.5 to less than 1×10^{-18}. Using the above equation:

1. $T_{\mathrm{r}} = (8.85 \times 10^{-14})(2.5)/(7.6 \times 10^{-8}) = 2.91 \times 10^{-6}$ s
2. $T_r = (8.85 \times 10^{-14})(2.5)/(1 \times 10^{-18})$

   $= 2.21 \times 10^5$ s (this is in excess of 60 h)

Benzene typically contains some water and has a higher conductivity than in the above example and has a much lower relaxation time.

The purity of a liquid has a great effect on its relaxation time, and thus its static hazard potential. In actual practice, relaxation times of a few seconds to an hour are encountered, depending on the purity and dryness of the liquid. This emphasizes the dangers of open sampling of tank contents soon after filling. If it is likely that the liquid being used has a low conductivity, it is important that enough time elapses between activities that can produce a static charge, such as loading a tank, and any activity that could cause a spark, such as sampling from the top of the tank.

In case (2) in the above example, a conductivity of 10^{-18} S/cm is so low that there may be little charge separation and little charge formation, and there may be no hazard even though the calculated relaxation time is extremely long. Materials with a half-time value of less than 0.012 s have been reported not to cause a hazard. A useful rule to remember is that the concept of relaxation is very important because it is possible for liquid in a tank to retain an electric charge for a long time if the liquid is a poor conductor, even if the tank is grounded. The specific conductivity, and therefore the relaxation time, is greatly affected by impurities. For example, the specific conductivity of benzene can vary from as long as 1×10^{-18} to about 7.6×10^{-8} S/cm, depending on its purity. It can vary significantly with the amount of water or other materials dissolved in the benzene [17].

Resistivity: The extent of charge separation is dependent on the resistivity of the liquid. Some materials have a sufficiently high conductivity to render them safe in terms of static buildup. If the resistivity is low, charge separation is easy, but so is charge recombination through the liquid. If the resistivity is high, there may be appreciable charge separation without immediate recombination, leading to a high charge. If the resistivity is extremely high, there may be no charge separation, and there is no buildup of a charge. If the conductivity of a liquid falls in the hazardous range, it is possible to modify it by the use of a very small amount of an additive. Additives usually are a combination of a polyvalent metal salt of an acid such as carboxylic or sulfonic acid and a suitable electrolyte. Additives of this type can impart a conductivity of 10^{-8} S/m (Siemens per meter) in a 0.1% solution in benzene [19].

A useful rule to remember is that when the resistivity of a liquid exceeds 10^{15} Ω centimeters (Ω-cm), or is less than 10^{10} Ω static generation or accumulation is negligible. Between these limits, the net generation of charges increases with the *maximum charge generation at* 10^{3} Ω. Styrene, for example, a commonly used monomer, has a resistivity of 4×10^{13} at 20°C [20], and therefore is capable of building up a potentially hazardous charge.

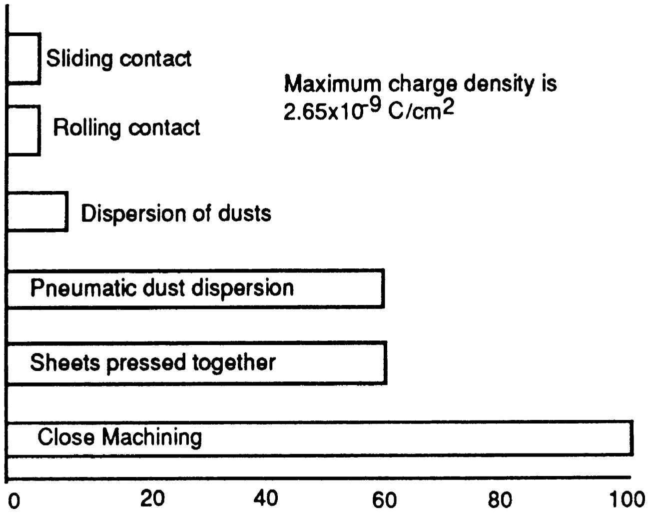


Fig. 2.9 Percentage of maximum theoretical charge produced by various operations. (Data from Eichel [21])

Static charge development: Static electrification of solids can occur in various ways. Different operations will produce the percentages of the theoretical maximum charge density shown in Fig. 2.9 [21].

It should be noted that pure gases do not generate significant static electricity in transmission through pipes and ducts. Gases contaminated with rust particles or liquid droplets produce static, but this is not a problem in a closed, grounded piping system. *If these gases impinge on an ungrounded, conductive object, dangerous charges can accumulate on that object.* Wet steam, which contains water droplets, can develop charges. *If the water droplets contact an ungrounded conductor, that object can develop a static charge.*

Flammable gases may ignite when discharged to air during thunderstorms, even without a direct lightning hit. Dry hydrogen and occasionally other gases may ignite when they are discharged to air in normal weather. This may be so because the electric field developed by the ejected gases can develop a corona discharge which can cause ignition. The minimum ignition energy of hydrogen is only 0.02 mJ. A toroidal ring developed by the National Aeronautics and Space Administration (NASA) is reported to prevent unwanted discharge and subsequent ignition of a vent-stack outlet [22].

Humidification: The conductivity of electrical nonconductors, such as plastics, paper, and concrete, depends on their moisture content. Relatively high moisture in these materials increases conductivity and therefore increases dissipation of static electricity. With relative humidity of 60–70% or higher, a microscopic film of moisture covers surfaces, making them more conductive.

Humidification can and often should be practiced to reduce the hazard of static electricity, but should not be relied on entirely to remove all possibility of static discharge.

In winter, cold air brought into a building and heated to normal room temperature is extremely dry, often less than 5–10% relative humidity. When processing solid materials that can develop a static charge, this air should be humidified to reduce static hazards as well as improve the comfort of personnel.

Filling liquid containers: A fire during top loading of a flammable liquid into a tank constitutes a serious problem if there could be a flammable mixture in the vapor space. Static electricity can be generated by splashing if the liquid is top-loaded, so it is normal practice to fill with a dip pipe positioned so the tip of the dip pipe is near the bottom of the tank. This may not be sufficient to prevent static charge buildup, as a charge may be generated in the bottom of the tank before the pipe tip is fully submerged, and it is possible for the liquid to acquire a charge before it reaches the tank.

Product filters using cotton, paper, felt, or plastic elements are prolific generators of static electricity. It is considered that at least 30 s is necessary to dissipate this charge, although with dry nonconductive liquids, it may require as long as 500 s.

Loading a less volatile liquid into a tank where there was previously a more volatile liquid is particularly hazardous because the more volatile liquid may form a flammable mixture, and the less flammable material is often a poor conductor and will not readily dissipate static charge. This type of loading accounts for 70–80% of severe losses at terminals [19]. This appears to occur most often when the compartments are one-fourth to one-third full, and when the temperature is close to -1°C.

Inerting the tank while it is being filled will reduce the possibility of ignition by static electricity and is highly recommended when it is possible and practical. However, this is not always practical. In any case, if inerts are to be used, they must be added carefully, as the following example illustrates. Two firemen were fatally injured when an explosion occurred as they were attempting to use portable CO_2 fire extinguishers to inert a tank truck. The source of ignition was believed to be a spark from the horn of the extinguisher to the latch on the tank truck. It was found that the voltage on the horn increased as the carbon dioxide "snow" passed down the horn to the outlet side.

Grounding and bonding lines, although very important, will not immediately dissipate the charge on the surface of a nonconducting liquid in a tank. A relaxation time for charge to be dissipated should be allowed after filling or other operations to permit static charge on the liquid surface to dissipate to the dip pipe or tank shell. *The minimum time is 1 min, but longer periods are advisable with some liquids that have extremely low conductivity. Bottom loading may reduce the static electricity hazard but does not eliminate it* [23].

Review of Design Alternatives: Hazards should be considered and eliminated in the process development stage where possible. This would include considerations of alternative processes, reduction or elimination of hazardous chemicals, site selection, etc. By the time the process is developed, the process designers already have major constraints imposed on them. Hazards should also be identified and removed or reduced early in the design. Adding protective equipment at the end of the design or after the plant is operating can be expensive and not entirely satisfactory. Allowing time in the early stages of design for critical reviews and evaluation of alternatives would involve studies such as an early hazard and operability (HAZOP) study, using flowsheets, before final design begins [24]. Fault tree analysis," quantitative risk assessment (QRA), checklists, audits, and other review and checking techniques can also be very helpful. These techniques are extensively discussed in the technical literature and will not be discussed in detail here.

Process Safety Competency

The next element under the pillar "Commit to Process Safety" is "Process Safety Competency." Process safety competency encompasses three interrelated actions; ensuring the appropriate information is available to people who need it, continuously improving knowledge and competency and consistently applying what has been learned. It is important for several reasons. Catastrophic process safety incidents are relatively rare, but losses can be devastating. Because of this learning must be proactive so that lessons must not be forgotten. Only competent people can transform information into knowledge. Knowledge management, not information management, helps organizations understand and manage risks and remain competitive.

The following principles must be followed. Facilities should implement management systems to help identify learning needs that are critical to process safety, support efforts to learn or obtain critical knowledge, maintain knowledge in a manner that helps promote risk-informed decision-making, and share information with other facilities (in some cases even competitors). Activities that help maintain and enhance process safety competency must be executed and evaluated. Results of the evaluation must be shared and plans adjusted as necessary.

Workforce Involvement

The fourth element in the pillar "Commit to Process Safety" is "Workforce Involvement." Workforce involvement includes developing a written plan of action regarding worker participation, consulting with workers on the development of the RBPS management system, and providing workers and their representatives access to all information developed under the RBPS system. In the US PSM standard workforce involvement is called employee participation and is a PSM regulatory requirement. This element provides for a consultative relationship between management and workers at all levels of the organization. It is important because it provides an equitable mechanism for workers to be directly involved in protecting their own welfare. It facilitates access to information only available through the unique experience of operating the process every day. It also provides a mechanism for workers to access necessary information and reinforces the process safety culture.

Stakeholder Outreach

The last element under the pillar "Commit to Process Safety" is "Stakeholder Outreach." Stakeholder outreach involves seeking out individuals or organizations that can be or believe they can be affected by company operations and engaging them in a dialogue about process safety. It also involves establishing a relationship with community organizations, other companies and professional groups, and local, state, and federal authorities. Stakeholders must be provided with accurate information about the company and the facility's products, processes, plans, hazards, and risks. This is important for several reasons. Sharing information with industry peers will promote better process safety for everyone. Sharing information in proactive ways with community and government stakeholders will build trust and commitment. By promoting openness and responsiveness, an effective outreach program will increase all stakeholders' confidence in the company.

Understand Hazards and Risk

The second pillar of RBPS is *Understand Hazards and Risk*. To understand hazards and risk, facilities should focus on:

- Collecting, documenting, and maintaining process safety knowledge
- Conducting hazard identification and risk analysis studies

Process Safety Knowledge

The first element is "process safety knowledge." Process safety knowledge includes written technical documents and specifications, engineering documents and calculations, specifications for design, fabrication, and installation of process equipment, and other written documents such as material safety data sheets (MSDSs). In the OSHA PSM standard process safety knowledge is called "Process Safety Information." To comply with the standard, employers must compile considerable documented process safety information on the hazards of chemicals used in a covered process as well as information on the process technology and equipment before conducting the process hazard analyses required by the standard. Process safety knowledge is important because risk understanding depends on accurate process safety knowledge. Process knowledge also supports other RBPS elements such as procedures, training, asset integrity, management of change, and incident investigation.

Process safety knowledge includes understanding the characteristics that are inherent to each material that is used in a process. Material hazards include flammability and combustibility hazards, reactivity hazards, toxicity hazards, and corrosivity hazards. Each type of hazard will be discussed.

Combustion Hazards

The enchanting flame has held a special mystery and charm the world over for thousands of years. According to Greek myth, Prometheus the Titan stole fire from the heavens and gave it to mortals—an act for which he was swiftly punished. Early people made use of it anyway. Soon the ancients came to regard fire as one of the basic elements of the world. It has since become the familiar sign of the hearth and a mark of youth and blood, as well as the object of intense curiosity and scientific investigation.

Suitably restrained, fire is of great benefit; unchecked or uncontrolled, it can cause immense damage. We respond to it with a powerful fascination coupled with an inbred respect and fear. A good servant but a bad master is Thoreau's "most tolerable third party" [25].

Fire [26]: Fire or combustion is normally the result of fuel and oxygen coming together in suitable proportions and with a source of heat. The consumption of a material by a fire is a chemical reaction in which the heated substance combines with oxygen. Heat, light, smoke, and products of combustion are generated. The net production of heat by a fire involves both heat-producing and heat-absorbing reactions, with more heat being produced than is absorbed. Energy in the form of heat is required:

1. To produce vapors and gases by vaporization or decomposition of solids and liquids. Actual combustion usually involves gases or vapors intimately mixed with oxygen molecules.
2. To energize the molecules of oxygen and flammable vapors into combining with one another and so initiating a chemical reaction.

The amount of energy required to cause combustion varies greatly. Hydrogen and carbon disulfide can be ignited by tiny sparks, or simply may be ignited by static generated as the gases or vapors discharge from pipes into air. Other materials, such as methylene chloride, require such large amounts of energy to be ignited that they sometimes are considered nonflammable. Fire also can result from the combining of such oxidizers as chlorine and various hydrocarbon vapors; oxygen is not required for a fire to take place.

There are exceptions to the general rule that a solid must vaporize or decompose to combine with oxygen; some finely divided materials such as aluminum powder and iron power can burn, and it is generally accepted that they do not vaporize appreciably before burning.

Products of combustion: Heat, light, smoke, and asphyxiating toxic gases are produced by fire. In a hot, well-ventilated fire, combustion usually is nearly complete. Almost all the carbon is converted to carbon dioxide, and all the hydrogen to steam, and oxides of various other elements such as sulfur and nitrogen are produced.

This is not the case in most fires, where some of the intermediate products, formed when large complex molecules are broken up, persist. Examples are hydrogen cyanide from wool and silk; acrolein from vegetable oils; acetic acid from timber or paper; and carbon or carbon monoxide from the incomplete combustion of carbonaceous materials. As the fire develops and becomes hotter, many of these intermediates, which are often toxic, are destroyed (e.g., hydrogen cyanide is decomposed at a significant rate at 538°C).

Small airborne particles of partially burnt carbonaceous materials form smoke, which is often thickened by steam, when there is only partial combustion of fuel.

Solids: Ordinarily, combustible solids do not combine directly with oxygen when they burn. They give off vapor and gaseous decomposition products when they are heated, and it is the vapors or gases that actually burn in the characteristic form of flames. Thus, before a solid can be ignited, it usually must be heated sufficiently for it to give off flammable concentrations of vapors. Glowing, which is combustion in the solid state, is characteristic of materials in the final stages of a fire's decay when flammable gases have been burned away, or when the production of gases and vapors has been suppressed.

Solids with larger surface areas, in relation to their volume, burn more readily than those that are more compact when exposed to heat and oxygen in the air. Common materials such as textiles in the form of fibers or fabrics, foamed rubber, foamed plastics, thin sheets of plastic, paper, corrugated cardboard, combustible dusts, dry grass and twigs, and wood shavings are examples of materials with large surface areas in relation to their volume. In a well-established fire, materials with relatively small surface areas, such as chunks of coal or logs, burn readily.

Combustion is self-propagating; burning materials produce heat which causes more of the solid to produce flammable vapors until either the fuel or oxygen is exhausted, or until the fire is extinguished in some other way.

Dusts: Most combustible solids can produce combustible dusts. Combustible dusts are particularly hazardous; they have a very high surface area to volume ratio. When finely divided as powders or dusts, solids burn quite differently from the original material in the bulk. Dust and fiber deposits can spread fire across a room or along a ledge or roof beam very quickly. Accumulations of dust can smolder slowly for long periods, giving little indication that combustion has started until the fire suddenly flares up, possibly when no one suspects a problem.

Many combustible dusts produced by industrial processes are explosible when they are suspended as a cloud in air. Even a spark may be sufficient to ignite them. After ignition, flame spreads rapidly through the dust cloud as successive layers are heated to ignition temperature. The hot gases expand and produce pressure waves that travel ahead of the flame. Any dust lying on surfaces in the path of the pressure waves will be thrown into the air and could cause a secondary explosion more violent and extensive than the first.

Liquids: A vapor has to be produced at the surface of a liquid before it will burn. Many common liquids give off a flammable concentration of vapor in air without being heated, sometimes at well below room temperature. Gasoline, for example, gives off ignitable vapors above about −40°C, depending on the blend. The vapors are easily ignited by a small spark or flame. Other liquids, such as fuel oil and kerosene, need to be heated until sufficient vapor is produced.

Many liquids can be formed into mists that will burn at temperatures where the vapor pressure is insufficient to produce a flammable mixture of the vapor and air.

For any flammable vapor there are maximum and minimum concentrations of the vapor in air beyond which it cannot burn. When the mixture of vapor in air is too weak, there is insufficient fuel for burning; when the mixture is too strong, there is insufficient oxygen for burning.

If the density of a flammable vapor is greater than that of air, as is normally the case, flammable concentrations may collect at low levels, such as at floor level or in basements, and can travel a considerable distance to a source of ignition, from which flames will then flash back.

Gases: Flammable gases usually are very easily ignited if mixed with air. Flammable gases often are stored under pressure, in some cases as a liquid. Even small leaks of a liquefied flammable gas form relatively large quantities of gas, which is ready for combustion.

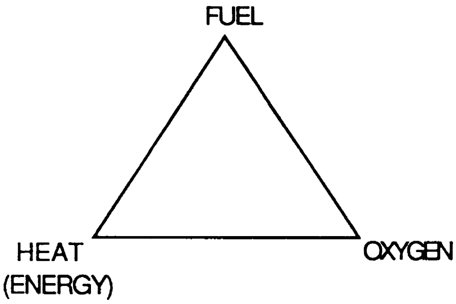


Fig. 2.10 The fire triangle

The fire triangle: The well-known "fire triangle" (see Fig. 2.10) is used to represent the three conditions necessary for a fire:

1. Fuel
2. Oxidizer: oxygen or other gaseous oxidizer such as chlorine; or liquid oxidizer such as bromine; or solid oxidizer such as sodium bromate
3. Energy, usually in the form of heat

If one of the conditions in the fire triangle is missing, fire does not occur; if one is removed, fire is extinguished. Usually a fire occurs when a source of heat contacts a combustible material in air, and then the heat is supplied by the combustion process itself.

The fire triangle indicates how fires may be fought or prevented:

1. Cut off or remove the fuel.
2. Remove the heat or energy—usually by putting water on the fire.
3. Remove the supply of oxygen—usually by foam or inert gas.

Flammability: No single factor defines flammability, but some relevant parameters include:

1. Flash point—often considered the main index of flammability; low flash points usually mean increased flammability.
2. Flammability limits; wide limits mean increased flammability.
3. Autoignition temperature; low temperature means increased flammability.
4. Ignition energy; low ignition energy means increased flammability.
5. Burning velocity; high velocity means increased flammability.

A combustion process is an exothermic reaction initiated by a source of ignition that produces more energy than it consumes. The speed at which the reaction proceeds through the mixture of reactants depends on the concentration of the flammable gas or vapor. This speed is lower at higher ("rich") as well as at lower ("lean") concentrations of the flammable gas than at the stoichiometric mixture. There are lower and upper limits beyond which the reaction cannot propagate through the gas mixture on its own. Some definitions follow:

1. *Flammability limits*: The range of flammable vapor–air or gas–air mixtures between the upper and lower flammable limits. Flammability limits are usually expressed in volume percent. Flammability limits are affected by pressure, temperature, direction of flame propagation, oxygen content, type of inerts, and other factors. The precise values depend on the test method.
2. *Upper flammability limit*: The maximum concentration of vapor or gas in air above which propagation of flame does not occur on contact with a source of ignition.
3. *Lower flammability limit*: The minimum concentration of vapor or gas in air or oxygen below which propagation of flame does not occur with a source of ignition.
   The concentrations at the lower and upper flammability limits are roughly 50 and 200–400%, respectively, of the stoichiometric mixture. The maximum flammability usually (not always) occurs at the stoichiometric mixture for combustion [17, 19].
4. *Flammable limits for mixtures of flammable gases and vapors*: For mixtures of several flammable gases and vapors, the flammable limits can be estimated by application of Le Chatelier's equation, if the flammable limits of the components are known: [19]

$$L = \frac{1}{\sum_{i=1}^{n} (y_i/L_i)} \qquad U = \frac{1}{\sum_{i=1}^{n} (y_i/U_i)}$$

   where

   L = lower flammability limit of the fuel mixture, vol.%.
   L_i = lower flammability limit of fuel component i, vol.%.
   U = upper flammability limit of the fuel mixture, vol.%.
   U_i = upper flammability limit of fuel component i, vol.%.
   y_i = concentration of fuel component i, vol.%.

   This equation is empirical and is not universally applicable, but is useful and a reasonable approximation when actual mixture data are not available.

   It is possible for a mixture to be flammable even though the concentration of each constituent is less than its lower limit.
5. *Methods of measurement*: Flammability limits are determined by measuring the volume percent of a flammable gas in an oxidizing gas that will form a flammable mixture, thus identifying the lower and upper flammable limits as well as the critical oxygen concentration (the minimum oxidizer concentration that can be used to support combustion).
6. *Uniformity of lower limits on a mass basis*: Concentrations of vapors and gases usually are reported in volume percent. As molecular weight increases, the lower limit usually decreases. On a mass basis, the

lower limits for hydrocarbons are fairly uniform at about 45 mg/L at 0°C and 1 atm. Many alcohols and oxygen-containing com pounds have higher values; for example, on a mass basis, ethyl alcohol in air has a lower limit of 70 mg/L [17].

7. *Effect of temperature on flammable limits*: The higher the temperature at the moment of ignition, the more easily the combustion reaction will propagate. Therefore, the reference temperature (initial temperature) of the flammable mixture must be stated when flammable limits are quoted. There are not a lot of data for flammable limits under different conditions of initial temperature. The behavior of a particular mixture under different conditions of initial temperature usually must be determined by tests.
8. *Burning in atmospheres enriched with oxygen*: The flammability of a substance depends strongly on the partial pressure of oxygen in the atmosphere. Increasing oxygen content affects the lower flammability limit only slightly, but it has a large effect on the upper flammability limit. Increasing oxygen content has a marked effect on the ignition temperature (reduces it) and the burning velocity (increases it).

   At the lower explosive limits of gas–air mixtures, there is already an excess of oxygen for the combustion process. Replacing nitrogen by additional oxygen will influence this limit very little [7].
9. *Burning in chlorine*: Chemically, oxygen is not the only oxidizing agent, though it is the most widely recognized and has been studied the most. Halogens are examples of oxidants that can react exothermically with conventional fuels and show combustion behavior. The applicability of flammability limits applies to substances that burn in chlorine. Chlorination reactions have many similarities to oxidation reactions. They tend not to be limited to thermodynamic equilibrium and often go to complete chlorination. The reactions are often highly exothermic. Chlorine, like oxygen, forms flammable mixtures with organic compounds. As an example: a chlorine–iron fire occurred in a chlorine pipeline, causing a chlorine gas release. Chlorine had liquefied in the lines because of the very cold weather, and the low spot was steam-traced. Steam had been taken from the wrong steam line, using 400 psig steam instead of 30 psig steam. The 400 psig steam was hot enough to initiate the reaction. This serves as a reminder that steel and chlorine can react. The allowable temperature for safe use depends upon the state of subdivision of the iron.
10. *Burning in other oxidizable atmospheres*: Flames can propagate in mixtures of oxide of nitrogen and other oxidizable substances. For example, Bodurtha [17] reports that the flammability limits for butane in nitric oxide are 7.5% (lower) and 12.5% (upper).
11. *Flame quenching*: Flame propagation is suppressed if the flammable mixture is held in a narrow space. There is a minimum diameter for apparatus used for determination of flammability limits. Below this diameter the flammable range measurements are narrower and inaccurate.

    If the space is sufficiently narrow, flame propagation is suppressed completely. The largest diameter at which flame propagation is suppressed is known as the quenching diameter. For an aperture of slotlike cross section there is critical slot width.

    The term “quenching distance” sometimes is used as a general term covering both quenching diameter and critical slot width, and sometimes it means only the latter.

    There is a maximum safe gap measured experimentally that will prevent the transmission of an explosion occurring within a container to a flammable mixture outside the container. These data refer to a stationary flame. If the gas flow is in the direction of the flame propagation, a smaller gap is needed to quench the flame. If the gas flow is in the opposite direction, a larger gap will provide quenching. If the gas velocity is high enough, the flame can stabilize at the constriction and cause local overheating. These quenching effects are important in the design of flame arrestors.
12. *Heterogeneous mixtures* [16]. In industry, heterogeneous (poorly mixed) gas phase mixtures can lead to fires that normally would be totally unexpected. It is important to recognize that heterogeneous mixtures can ignite at concentrations that normally would be nonflammable if the mixture were homogeneous. For example, 1 L of methane can form a flammable mixture with air at the top of a 100-L container although the mixture only would contain 1.0% methane by volume if complete mixing occurred at room temperature, and the mixture would not be flammable. This is an important concept because “layering” can occur with any combustible gas or vapor in both stationary and flowing mixtures.

    Heterogeneous gas phase mixtures can lead to unexpected fires if a relatively small amount of flammable gas is placed in contact with a large amount of air without adequate mixing, even though the average concentration of flammable gas in the mixture is below the flammable limit. *Heterogeneous mixtures are always formed at least for a short time when two gases or vapors are first brought together*.
13. *Effect of pressure*: Flammability is affected by initial pressure. Normal variations in atmospheric pressure do not have any appreciable effect on flammability limits.

    A decrease in pressure below atmospheric usually narrows the flammable range. When the pressure is reduced low enough, a flame or an explosion can no longer be propagated throughout the mixture.

An increase in pressure above atmospheric usually (not always) widens the flammability range, especially the upper limit.

14. *Explosions in the absence of air*: Gases with positive heats of formation can be decomposed explosively in the absence of air. Ethylene reacts explosively at elevated pressure, and acetylene reacts explosively at atmospheric pressure in large-diameter piping. Heats of formation for these materials are +52.3 and +227 kJ/g/mol, respectively. Explosion prevention can be practiced by mixing decomposable gases with more stable diluents. For example, acetylene can be made nonexplosive at a pressure of 100 atm by including 14.5% water vapor and 8% butane.

Ethylene oxide vapor will decompose explosively in the absence of oxygen or air under certain conditions when exposed to common sources of ignition if heated to high enough temperatures. One way to prevent the decomposition reaction is to use methane gas to blanket the ethylene oxide liquid. It has also been found that liquid ethylene oxide will undergo a deflagration in the absence of oxygen with a very rapid pressure increase if ignited at a temperature and pressure above a certain level. Fortunately, the conditions required for propagation of the decomposition of liquid phase ethylene oxide are outside the current normal handling and processing ranges for the pure liquid. Propagation has not been observed below 80°C at from 14 to 100 atm pressure [27]. Ethylene oxide also can undergo explosive condensation when catalyzed by a small amount of caustic [28].

Mists and Foams: If the temperature of a liquid is below its flash point, flammable concentrations of vapor cannot exist, but conditions still can exist for flammability if mists or foams are formed. A suspension of finely divided drops of a flammable liquid in air has many of the characteristics of a flammable gas–air mixture and can burn or explode. A mist may be produced by condensation of a saturated vapor or by mechanical atomization. Normally, the diameter of drops in a condensed mist is less than 0.01 mm, whereas in a mechanical spray it usually is greater than 0.1 mm.

The commonly accepted fallacy that liquids at temperatures below their flash points cannot give rise to flammable mixtures in air has led to numerous accidents. Flash points are measured under stagnant conditions in carefully controlled laboratory experiments, but in the real world one works with a wide variety of dynamic conditions that can produce mists and foams.

Flammable mist–vapor–air mixtures may occur as the foam on a flammable liquid collapses [16]. Thus, when ignited, many foams can propagate flame. An additional hazard can arise from the production of foams by oxygen-enriched air at reduced pressures. Air confined over a liquid can become oxygen enriched as pressure is reduced because oxygen is more soluble than nitrogen in most liquids. Thus, the presence of foams on combustible liquids is a potential explosion hazard.

The lower flammability limit for fine mists (<0.01 mm diameter) of hydrocarbons below their flash point, plus accompanying vapor, is about 48 g of mist/m^3 of air at 0°C and 1 atm. Mist can occur in agitated vessels under some conditions, especially when an agitator blade is at or near the liquid–vapor interface in the vessel.

Work on condensed oil mists (drop diameter mostly less than 0.01 mm) has demonstrated that they have flammability characteristics similar to those the mixture would have if it were wholly in the vapor phase at the higher temperature necessary for vaporization. The flammability characteristics are affected by drop size. For larger drop sizes (above 0.01 mm) the lower limit of flammability decreases as drop diameter increases. For mists, the amount of inert gas needed to suppress flammability is about the same as that needed to suppress an equivalent vapor–air mixture of the same material if it were vaporized at a somewhat higher temperature.

A useful rule is that mists of flammable or combustible liquids in air can burn or explode at temperatures below their flash points.

Ignition: Flammable gases and vapors can be ignited by many sources. In the design and operation of processes, it is best not to base fire and explosion safety on the presumption that ignition sources have been excluded. Bodurtha [22] reported that of 318 natural gas fires and explosions, the sources of ignition of 28% were unknown. All reasonable measures should be taken to eliminate possible sources of ignition in areas in which flammable materials are handled.

Autoignition: If the temperature of a flammable gas–air mixture is raised in a uniformly heated apparatus, it eventually reaches a value at which combustion occurs in the bulk gas. This temperature is defined as the spontaneous ignition temperature (SIT) or autoignition temperature (AIT). The gas–air mixture that has the lowest ignition temperature is called by various names, such as the minimum AIT, the minimum spontaneous ignition temperature, and the self-ignition temperature [17]. Usually the AIT reported in the literature is the minimum AIT.

The AIT of a substance depends on many factors, such as:

- Ignition delay
- Energy of ignition source
- Pressure
- Flow effects
- Surfaces
- Concentration of vapors
- Volume of container
- Oxygen content

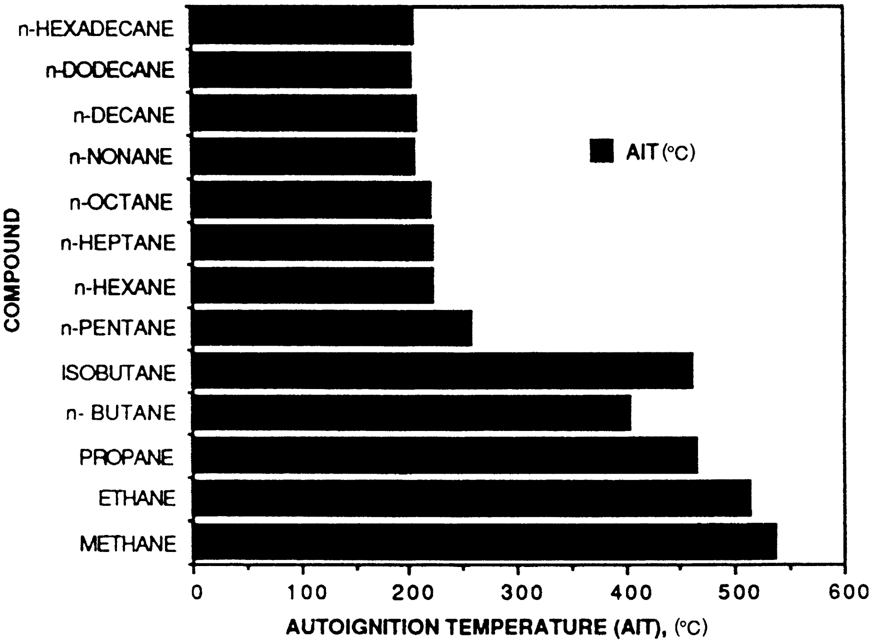


Fig. 2.11 Autoignition temperatures of paraffin hydrocarbons at 1 atm. (Data from Bodurtha [17])

- Catalytic materials
- Flow conditions

Thus, a specific AIT applies only to the experimental conditions employed in its determination. Usually the values quoted are obtained in clean laboratory equipment.

The AIT of a substance may be reduced below ideal laboratory conditions by as much as 100–200°C for surfaces that are insulated with certain types of insulation, or are contaminated by dust.

Mixtures that are fuel-rich or fuel-lean ignite at higher temperatures than do those of intermediate compositions. Also, in a homologous series of organic compounds, the AIT decreases with increasing molecular weight, as shown in Fig. 2.11.

Ignition delay: Ignition of a flammable mixture raised to or above the temperature at which spontaneous combustion occurs is not instantaneous; the time delay between the moment of exposure to high temperature and visible combustion is called the ignition delay. This time delay decreases as the ignition temperature increases. The time delay may be as little as a fraction of a second at higher temperatures, or several minutes close to the AIT.

Environmental effects: It has been found that the AIT becomes lower with increasing vessel size in the range of 35–12 L. An increase in pressure usually decreases AITs, and a decrease in pressure raises AITs. Usually oxygen enrichment of the air tends to decrease the minimum AIT, and a decrease in oxygen content increases the minimum AIT. Low-temperature oxidation can result in "cool flames," which may grow into ignition.

Catalytic materials: Ignition may occur where the temperature is less than the minimum AIT. Catalytic materials, such as metal oxides, can promote oxidation on their surfaces, leading to a high local temperature and subsequent ignition. There is a recorded reactive chemical case [29] in which a solvent at 80°C was being pressurized with a gas phase consisting of a high oxygen concentration. The solvent has a flash point in oxygen of greater than 130°C and normally is considered not to be a flammability hazard. There was an ignition, causing the vessel to rupture its main gasket with major damage to the facility. It was found that a mist had been formed in the vessel by the agitator, and that the source of ignition probably was a trace of palladium catalyst remaining from a previous run.

From this incident, several important lessons can be learned:

1. Ignition of a flammable mixture can result from totally unexpected contamination by trace amounts of catalysts if the oxidizer and fuel are present.
2. Mists of oxidizable liquids may form that can burn or explode at temperatures outside the "normal" flammable range.
3. It can be dangerous to perform experiments with pure oxygen, or air enriched with oxygen, especially under pressure and at elevated temperatures, when oxidizable materials are present.
4. The real criterion regarding flammable mixtures in air should be whether a flammable atmosphere can exist

under the given process conditions, rather than whether a flammable liquid is at a temperature below its flash point.

Cleaning up spills of flammable or combustible liquids: It is customary to clean up small spills of many liquid materials with sand or other noncombustible absorbent material. Some absorbing agents, such as untreated clays and micas, will cause an exothermic reaction with some liquids, especially monomers, which might ignite the liquid if it is flammable or combustible. Before any material is provided to be used to soak up spills of oxidizable material, tests should be made to determine if the material can cause fires with potential spills.

Ignition caused by insulation: Ignition of combustible materials that have been absorbed into commonly used insulating materials is possible at temperatures lower than the AIT for nonabsorbed material. All oxidizable materials oxidize to some extent in air at ambient temperatures, usually at a very low rate. When an absorbent material is absorbed into insulation, it is "spread" over a large area, increasing its access to oxygen. Because the absorbent is an insulator, heat from oxidation is retained rather than dissipated, and the temperature will rise if the heat is produced faster than it can be dissipated. The rate of oxidation increases as the material temperature increases, which produces more heat, compounding the hazard. If the temperature rises enough, the material will ignite ("spontaneous combustion"). This is similar to the classic oily rag and wet haystack phenomenon, which has caused many fires in homes and on farms. In the wet haystack phenomenon, fermentation by microorganisms will create heat. Some air is necessary; too much air will remove too much heat to allow the combustion temperature to be reached. For equipment operating above about 200°C containing combustible liquids with high boiling points, insulation should be impervious to the material handled. To date, only a closed cell foamed glass provides the required degree of protection where oxidizable liquid materials are used above 200°C. Insulation based on glass fiber, silicate, or alumina materials is known to cause hazardous situations and should not be used in this service.

Laboratory tests and actual fires show that Dowtherm A® (a heat-transfer fluid consisting of a eutectic mixture of biphenyl oxide and biphenol) can be ignited if it is soaked in glass fiber insulation and in contact with air at temperatures considerably below the normal AIT. This is also true for stearic acid soaked in glass fiber insulation. Table 2.1 shows the reduction in AIT of Dowtherm A® and of stearic acid soaked in glass fiber insulation.

Ignition of this type generally occurs only with materials having a high boiling point. Usually materials with low boiling points will vaporize and cannot remain soaked in hot insulation. There are exceptions. For example, ethylene oxide has a fairly low boiling point, but if it leaks into insulation, a polymer can be formed that has a high boiling point and can autoignite insulation at low temperatures.

Table 2.1 Reduction in AITs caused by liquids soaking in glass fiber insulation

Material	Normal AIT (°C)	Ignition temperature in glass fiber insulation (°C)
Dowtherm A®	621	260–290
Stearic acid	395	260–290

Ignition caused by impact: Solids and liquids can be ignited by impact. Impact tests are made by having a weight fall freely through a known distance and impacting the sample. Impact can occur, for example, if containers are accidentally dropped. The interpretation of the data from impact tests can be difficult.

Ignition caused by compression of liquids: Liquids can be ignited by sudden compression. This can happen when there is water hammer caused by the pressure surge from quick-acting valves and from the compression in liquid pumps. Sudden compression can occur with liquids, for example if a tank car is bumped rapidly and the liquid goes to one end very quickly, possibly trapping some vapor bubbles that compress and create local hot spots that can cause ignition.

Ignition caused by rubbing friction: Solids can be ignited by frictional sources when rubbed against each other or against another material. The frictional heat produced may be enough to ignite other materials, such as lubricants, that are nearby. A common example of this occurs when bearings run hot, causing oil or grease to vaporize and possibly ignite.

Ignition caused by glancing blows: Friction can cause ignition in other ways. Sparks may occur when two hard materials come in contact with each other in a glancing blow (the blows must be glancing to produce friction sparks). These kinds of sparks are not directly related to frictional impact. Hand and mechanical tools are the most likely sources of friction sparks that occur outside of equipment. The need for nonsparking tools is somewhat controversial; Bodurtha [17] states that it is extremely unlikely that anyone would be using tools in a flammable atmosphere, and it is usually more prudent to control the atmosphere than the tools. Sparkproof tools are not really sparkproof in all situations.

Ignition caused by static electricity: Static electricity is a potential source of ignition wherever there is a flammable mixture of dusts or gases (see previous section).

Ignition caused by compression of gases: If a gas is compressed rapidly, its temperature will increase. Autoignition

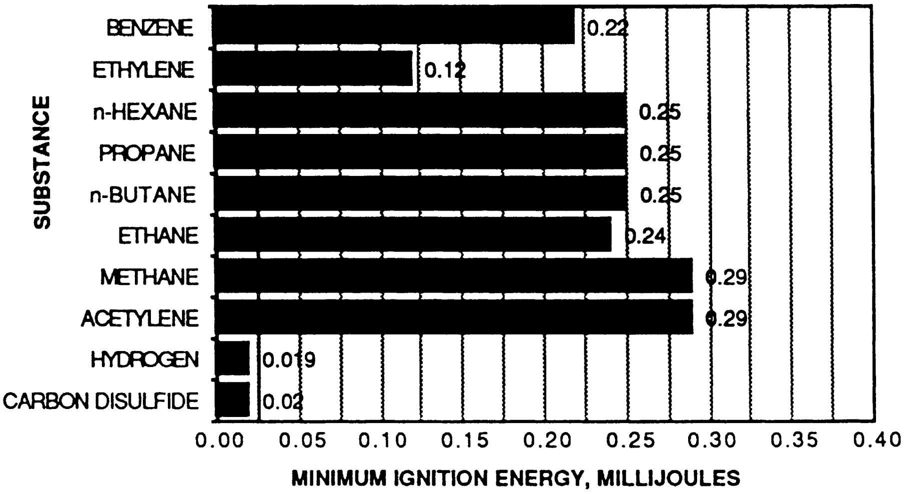


Fig. 2.12 Minimum ignition energy for selected substances (Less [19])

may occur if the temperature of the gas becomes high enough (this is more or less the principle of diesel engines).

An advancing piston of high-pressure gas can compress and heat trapped gas ahead of it. For a perfect gas, the temperature rise due to adiabatic compression is given by

$$\frac{T_2}{T_1} = \left(\frac{P_2}{P_1}\right)^{(k-1)/k}$$

where T_1 and T_2 are the initial and final gas absolute temperatures, P_1 and P_2 are the initial and final absolute pressures, and k is the ratio of heat capacity at constant pressure to the heat capacity at constant volume. For air and many other diatomic gases, $k = 1.4$. Many hydrocarbons have k values of between 1.1 and 1.2. The value of k is a function of temperature and pressure.

Energy levels for ignition: If a flammable gas mixture is to be ignited by a local source of ignition, there is a minimum volume of mixture required to cause a continuing flame throughout the mixture.

For example, to ignite a methane–air mixture in a cold container, a hot patch of 18 mm^2 at 1,000–1,100°C is required in order to heat enough volume of gas to produce a continuing flame [19], even though the auto-oxidation temperature for methane is 540°C. Ignition of a flammable gas–air mixture by electrical discharge can occur only if the electrical discharge is of sufficient energy.

Minimum ignition energy: There is a minimum ignition energy, which usually occurs near the stoichiometric mixture. The minimum ignition energy for some representative substances in air is shown in Fig. 2.12 [19]. The energy required to cause ignition frequently is reported in millijoules (mJ). One joule is 0.24 cal, so 1 mJ is 0.00024 cal, which is a very small amount of energy.

Table 2.2 Comparison of initiation energies of some common substances in air and pure Oxygen [47]

	Relative amount of energy required to initiate combustion (mJ)	
Flammable substance	In air	In pure oxygen
Methane	0.3	0.003
Hydrogen	0.019	0.0012
Acetone	1.15	0.0024
Diethyl ether	0.2	0.0013

A person typically has capacitance of 200 pF (pF), and if charged to 15 kilovolts (kV) could initiate a discharge of 22.5 mJ. This is enough to ignite many flammable mixtures. The energy in ordinary spark plugs is 20–30 mJ.

The hazard of an explosion should be minimized by avoiding flammable gas–air or dust–air mixtures in a plant. It is bad practice to rely solely on elimination of sources of ignition, as it is nearly impossible to ensure this.

Effect of oxygen-enriched atmospheres: The minimum spark energy to cause ignition varies greatly with the amount of oxygen in oxygen-enriched air. Stull [30] showed that with a composition of 10% methane in air, about 0.5 mJ of spark energy is required to initiate a reaction at the lower flammable limit. If the air is enriched with oxygen, the minimum spark energy decreases. If the flammable material is combined with 100% oxygen, the spark energy required is only about 1% of the required energy in air at 21% oxygen! This demonstrates the extremely small amount of energy required to initiate the reaction, as well as the additional ease with which oxygen-enriched atmospheres are initiated. Table 2.2 compares initiation energies of some common substances in air and in pure oxygen.

Effect of pressure: An increase in pressure decreases the amount of energy required to cause ignition. In a mixture of propane, oxygen, and nitrogen, doubling the pressure decreases the minimum energy required to cause ignition by a factor of about 5.

If no other data are available for determination of hazards, minimum ignition energies at ambient temperatures and pressures should be considered as approximately:

- 0.1 mJ for vapors in air
- 1.0 mJ for mists in air
- 10.0 mJ for dusts in air

Explosions

Development of Pressure

Exothermic reactions can lead to high temperatures and in the case of large fires to large loss of property and severe damage from radiant energy. However, in many plant accidents it is the sudden generation of pressure that leads to severe damage, injury, and deaths. Hence, it can be stated that "pressure blows up plants, not temperature." Of course, temperature and pressure are closely related, but it is the pressure effect that is of concern in this section.

The word "deflagration" can be defined in several ways. One definition is "a reaction that propagates to the unreacted material at a speed less than the speed of sound in the unreacted substance." [17] Another definition of deflagration is from Latin meaning "to burn down, or to burn rapidly with intense heat and sparks given off." [28] A deflagration may be an explosion, but not all deflagrations are explosions (a violently burning fire may be a deflagration, but that is not an explosion). On the other hand, not all explosions are deflagrations (a steam boiler may explode, but that is not a deflagration).

An explosion is a sudden and violent release of energy. Usually it is the result, not the cause, of a sudden release of gas under high pressure. The presence of a gas is not necessary for an explosion. An explosion may occur from a physical or mechanical change, as in the explosion of a steam boiler, or from a chemical reaction. The explosion of a flammable mixture in a process vessel may be either a deflagration or a detonation, which differs fundamentally. Both can be very destructive. Detonations are particularly destructive, but are unlikely to occur in vessels.

A detonation is a reaction that propagates to unreacted material at a speed greater than the speed of sound in the unreacted material; it is accompanied by a shock wave and extremely high pressures for a very short time. It is debatable whether the flammable range is the same as the detonable range. Detonation limits normally are reported to be within the flammable limits, but the view is widely held that separate detonation limits do not exist.

Unconfined vapor clouds can both deflagrate and detonate, with a deflagration being much more likely. A detonation is more destructive, but a deflagration also can produce a damaging pressure wave. A deflagration can undergo transition to a detonation in a pipeline, but this is most likely in vessels.

If a flammable mixture may be present in process equipment, precautions should be taken to eliminate ignition sources. However, it is prudent to assume that, despite these efforts, a source of ignition will at some time occur.

Deflagration

The conditions for a deflagration to occur are that the gas mixture is within the flammable range and that there is a source of ignition or that the mixture is heated to its AIT.

For the burning of hydrocarbon–air mixtures:

$$\frac{P_{2\text{MAX}}}{P_1} = \frac{N_2 T_2}{N_1 T_1} = \frac{M_1 T_2}{M_2 T_1}$$

where

T = absolute temperature

M = molecular weight of gas mixture

N = number of moles in gas mixture

P = absolute pressure

1,2 = initial and final states

2MAX = final state maximum value

The maximum pressure rise for a deflagration of flammable mixtures is approximately as follows for initial absolute pressures of 1–40 bar, for initial temperatures of 0–300°C, and for relatively small volumes of a few cubic meters:

$$\frac{P_2}{P_1} = \text{approximately 8 for hydrocarbon} - \text{air mixtures}$$

$$\frac{P_2}{P_1} = \text{approximately 16 for hydrocarbon} - \text{oxygen mixtures}$$

For conventionally designed pressure vessels:

$$\frac{P_b}{P_1} = \text{approximately } 4 - 5$$

where P_b = vessel bursting pressureP_1 = normal design pressureP_2 = pressure caused by deflagration

Therefore, in the absence of explosion relief, the deflagration explosion of a hydrocarbon–air mixture is easily capable of bursting a vessel if it is operating near its design pressure when the deflagration takes place. For reactions operating at or near atmospheric pressure, such as many drying and solids

Table 2.3 Overpressure from Detonations [39]

	Pressure (MPa)	Pressure (lb/in.^{2})
Incident overpressure	3.5	510
Maximum reflected pressure (wave strikes surface head-on)	28	4,100
Load the structure feels (due to acceleration)	56	8,100

MPa means pressure in megapascals

processing operations, it may be practical to construct facilities that will withstand the maximum explosion pressure of most dust–air and flammable gas–air mixtures.

Detonations

Detonation of a gas–air mixture may occur by direct initiation of detonation by a powerful ignition source or by transition from deflagration. This transition occurs in pipelines but is most unlikely in vessels. Two useful rules are:

1. Almost any gas mixture that is flammable is detonable if initiated with a sufficiently energetic source.
2. Detonation of a gas–air mixture is possible in pipelines but is unlikely in vessels.

Bartknecht [31] states that the range of detonability is narrower than the range of flammability. For example, the range of detonability of hydrogen in air is 18–59 vol.%, compared with the flammability of 4–75 vol.%. With flammable gases in air, if the length-to-diameter ratio of a pipe or vessel is more than about 10:1, and the pipe diameter is above a critical diameter, 12–25 mm, a detonation is possible.

Detonation pressure: In the case of the burning of a flammable mixture of gases in a pipe with one end closed, a series of pressure waves traveling at the speed of sound moves through the unburned gas. Later waves traveling through the unburned gas, which has been heated by compression from the earlier waves, speed up because of the higher temperature and overtake the first wave, and a shock wave develops. Flame follows the shock wave and catches up with it, forming a detonation wave. A stable detonation wave may develop, which moves with supersonic speed relative to the unburned mixture, and peak incident (side-on) pressures are of the order of *30 times* the initial absolute pressure.

Reflected pressure: Reflected pressure increases the pressure on a rigid surface if the shock wave impinges on the surface at an angle to the direction of the propagation of the wave. The maximum ratio of reflected pressure to incident (side-on) pressure when a strong shock wave strikes a flat surface head-on is 8:1. Furthermore, acceleration from a suddenly applied force of the detonation wave can double the load that a structure "feels." Table 2.3 shows overpressure that can be expected from typical detonations [32].

Thus, the stable detonation wave may cause enormously high pressures at closed ends of pipes, bends, and tees, where the greatest destruction from a gaseous detonation may occur.

Geometry: The following are some factors to consider when detonation is possible:

1. Large length-to-diameter ratios promote the development of detonations; vessels should be designed with the lowest length-to-diameter ratio practicable if a detonation is possible.
2. Equipment such as tanks (not including pipelines) designed to withstand 3.5 MPa (about 500 psig) usually will be adequate to contain a detonation, with other safeguards, for flammable gases in air at atmospheric pressure.
3. Dished heads survive detonations better than do flat heads because of the more unfavorable incidence of flat heads.
4. If turns in a process line are necessary, two 45° bends or a long sweep elbow will greatly reduce reflected pressure compared with a single 90° elbow.
5. Restrictions such as orifices in pipelines may intensify a detonation by promoting pressure piling, which results when there are interconnected spaces such that the pressure rise in one space causes a pressure rise in a connected space. The enhanced pressure in the latter then becomes the starting pressure for a further explosion.
6. Detonation may be extinguished when it enters a wider pipe from a smaller one, but the detonation may be regenerated somewhere along the longer pipe.
7. Flame arresters, if properly designed, can arrest detonations.

Explosion Violence

The *rate of pressure rise* is a measure of the violence of an explosion. The maximum rate of pressure rise for confined explosions is greatly affected by the volume of the vessel, with the influence of vessel volume on the rate of pressure rise being given by the following equation:

$$(\mathrm{d}p/\mathrm{d}t)_{\max}(V^{1/3}) = \text{a constant } = K_G$$

where$(\mathrm{d}p/\mathrm{d}t)_{\max}$ = maximum rate of pressure rise, bar/sV = vessel volume, m^3K_G = a specific material constant (bar)(m)(s)$^{-1}$This is the *cubic law*, which states that for a given

Table 2.4 K_G values of gases, spark-ignited with zero turbulence, ignition energy ~10 J, $P_{\max} = 7.4$ bar[8]

Flammable gas	K_G (bar)(m)/s
Methane	55
Propane	75
Hydrogen	550

flammable gas, the product of the maximum pressure rise and the cube root of the vessel volume is a specific material constant, K_G.

The cubic law allows the prediction of the course of an explosion of a flammable gas or vapor in a large vessel, based on laboratory tests. It is valid only for the following conditions:

- The same optimum concentration of the gas–air mixture
- Same shape of reaction vessel
- The same degree of turbulence
- The same ignition source

Thus, to characterize an explosion, it is not enough to quote the maximum rate of pressure rise: the volume, vessel geometry [31], turbulence, and ignition energy must also be stated. Table 2.4 lists the K_G values for some common flammable gases measured under laboratory conditions.

It can be seen that the violence of an explosion with propane is about 1.5 times higher than one with methane, and one with hydrogen is about 10 times higher than one with methane. The explosive behavior of propane is representative of many flammable organic vapors in air. Some important relationships among pressure, temperature, turbulence, and vessel shape are discussed below.

1. *Explosion pressure is primarily the result of temperature reached during combustion, not a change in moles.* With complete combustion of propane in air there is a negligible change in moles of gas:

$$C_3H_8 + 5O_2 + \overset{\{...\text{air}...\}}{18.8N_2} = 3CO_2 + 4H_2O + 18.8N_2$$

   Number of moles at start = 24.8. Number of moles after complete combustion = 25.8. *Therefore, explosion pressure usually develops principally from an increase in temperature, not an increase in gas moles, during the combustion process of many materials.*Peak explosion pressure at constant volume occurs near the stoichiometric concentration in air. If only a small part of the total volume of a container is filled by an explosive gas–air mixture at atmospheric pressure, and the remainder of the vessel contains air, an explosion of this mixture can create enough pressure to severely damage containers that are designed to withstand only slight pressure—such as buildings and low-pressure storage tanks.

2. *Initial pressure affects maximum explosion pressure and rate of pressure rise.* If the initial pressure is increased above atmospheric pressure, there will be a proportional increase in the maximum explosion pressure and in the rate of pressure rise. Reducing the initial pressure will cause a corresponding decrease in maximum explosion pressure until finally an explosion reaction can no longer be propagated through the gas mixture.
3. *Initial temperature affects maximum explosion pressure and rate of pressure rise*: The maximum explosion pressure decreases when the starting temperature increases at the same starting pressure because of the lower density and thus smaller mass of material within a confined volume at higher temperatures. The maximum rate of pressure rise, $(dp/dt)_{\max}$, increases as the initial temperature rises because the burning velocity increases with an increase in initial temperature.
4. *Initial turbulence increases the rate of pressure rise*: Initial turbulence greatly increases the rates of explosion-pressure rise [17, 31]. It has been found that with pentane and methane mixtures in air, $(dp/dt)_{\max}$ can be five to nine times more with high initial turbulence than with no turbulence. The maximum explosion pressure is raised by about 20%. The course of explosions of flammable gases with a low normal speed of combustion, such as methane, is influenced by turbulence to a much higher degree than is the course of explosions with a high speed of combustion, such as hydrogen. Test data usually are obtained in equipment with a high degree of turbulence.
5. *Effect of vessel shape and increased initial pressure*: The maximum explosion pressure in confined vessels is not significantly affected by the volume or shape of the vessel in confined explosions for vessels that approximate the "cubic shape," that is, with a ratio of diameter to length (or vice versa) of about 1:1–1:1.5. In closed elongated vessels with central ignition, spherical ignition of the flame front will cause the flame to proceed swiftly in an axial direction. In the process, it compresses the unburned gases ahead of it, causing the violence of the explosion to increase, and pressure oscillations may occur.

Losses from Dust Explosions

Most organic solids, most metals, and some combustible inorganic salts can form explosive dust clouds. In order to have a dust explosion, it is necessary to satisfy certain conditions:

- Suitably sized dust particles
- Sufficient source of ignition energy
- Dust concentration within explosive limits
- Explosible dust
- Oxidizer must be present

If an explosive dust in air that meets the above criteria occurs in a process, an explosion should be considered as inevitable. The process designer should take into account the possibility of dust explosions and design accordingly.

In dust explosions the combustion process is very rapid. The flame speed is high compared with that in gas deflagrations. Detonations normally do not occur in dust explosions in industrial plants.

The sequence of events in a serious industrial dust explosion is often as follows:

1. A primary explosion occurs in part of a plant, causing an air disturbance.
2. The air disturbance disperses dust and causes a secondary explosion, which is often more destructive than the primary explosion.

If the occurrence of a flammable (explosive) dust is inevitable in a particular process, several design alternatives or combinations of alternatives are available:

- Containment (maximum pressure of a dust explosion is usually below 120–150 psig)
- Explosion venting to a safe place
- Inerting (most organic dusts are non flammable in atmospheres containing less than about 10% oxygen)
- Suppression

A fundamental solution to the dust explosion problem is to use a wet process so that dust suspensions do not occur. However, the process must be wet enough to be effective. Some dusts with a high moisture content can still be ignited.

Dust concentrations in major equipment may be designed to be below the lower flammable limit, but this often cannot be depended on in actual operation. Dust concentrations cannot be safely designed to be above an upper flammable limit because such a limit is ill-defined [19].

For a large number of flammable dusts, the lower explosion limit lies between 0.02 and 0.06 kg/m^3. The upper explosion limit is in the range of 2–6 kg/m^3, but this number is of limited importance.

A small amount of flammable gas or vapor mixed in with a flammable dust can cause an explosive mixture to be formed even if both are at concentrations below the explosive range by themselves. These mixtures are called “hybrid” mixtures. The ignition energy to ignite a hybrid mixture is often less than that required for the flammable dust by itself.

Venting is only suitable if there is a safe discharge for the material vented. Whenever an explosion relief venting device is activated, it may be expected that a tongue of flame containing some unburned dust will first be ejected. The unburned dust will be ignited as it flows out of the vent and can produce a large fireball that will extend outward, upward, and downward from the vent. It is essential for protection of personnel that venting is to an open place not used by people. If a duct must be used, the explosion pressure in the enclosure will be increased considerably. Therefore, particular attention must be paid to the design of the enclosure in which the explosion could take place.

The NFPA 68 guide issued in 1998 [33] has nomographs, which can be used to select relief areas required for combustible dusts when test data on the dusts are available. The nomographs in NFPA 68 are considered by many to be the preferred way to design dust explosion relief devices.

Relief venting to reduce dust explosion pressure requires the equipment to be protected to have a certain minimum strength. If the enclosure strength is too low, the enclosure will be damaged or destroyed before the explosion relief device can function. NFPA 68 [33] states that the strength of the enclosure should exceed the vent relief pressure by at least 0.35 psi. For industrial equipment such as dryers and baghouses, it is often desirable to have considerably more strength built into the structure to reduce the size of the vent area required. Also, the supporting structure for the enclosure must be strong enough to withstand any reaction forces developed as a result of operation of the vent.

Inerting is a very good preventive measure against dust explosions. The maximum oxygen concentration at which dust explosions are “just not possible” cannot be predicted accurately, as it depends on the nature of the combustible material; testing is usually required. It has been found that in an atmosphere of 10% oxygen and 90% nitrogen, most combustible organic dusts are no longer explosive. To allow a safety margin, it is good industrial practice to maintain oxygen concentrations below 8%. For metal dusts, the allowable oxygen content is about 4% [6].

Inerting leads to the possibility of asphyxiation by operating personnel if they were exposed to the inert gas. Strict precautions must be taken to prevent exposure of personnel to inerting atmospheres.

Explosion suppression systems are designed to prevent the creation of unacceptably high pressure by explosions within enclosures that are not designed to withstand the maximum explosion pressure [31]. They can protect process plants against damage and also protect operating personnel in the area. Explosion suppression systems restrict and confine the flames in a very early stage of the explosion. Suppression systems require more maintenance than do relief venting devices. Explosion suppression systems are made by only a few manufacturers and are quite expensive. This may be the reason why this type of safe-guard has not been as widely used in industry as one might expect, although its effectiveness has been proved by much practical experience.

Explosion suppression is a proven technology and should be considered as a candidate for explosion protection. The NFPA has published a standard reference on explosion-suppression protection [8]. Manufacturers should be consulted on design, installation, and maintenance.

Even with explosion suppression, it is common for the explosion pressure to reach one atmosphere before it is suppressed. The added pressure surge from the injection of the suppressing agent must also be considered. Therefore, sufficient mechanical strength is always required for enclosures protected by explosion suppression.

Boiling Liquid Expanding Vapor Explosions: Among the most damaging of accidents is a boiling liquid expanding vapor explosion (BLEVE, pronounced BLEV-ee). This occurs when a pressure vessel containing liquid is heated so that the metal loses strength and ruptures. Typically, this happens when the vessel failure results from overheating upon exposure to fire. The failure usually is in the metal contacting the vapor phase; the metal in this area heats to a higher temperature because there is no liquid heat sink to keep the metal temperature from rising rapidly, as there is where metal contacts a liquid phase. A BLEVE can occur with both flammable materials and nonflammable materials, such as water. In all cases the initial explosion may generate a blast wave and missiles. If the material is flammable, it may cause a fire or may form a vapor cloud that then gives rise to a secondary explosion and fireball. Kletz states that BLEVEs can cause as many casualties as can unconfined vapor cloud explosions (UVCEs) [19].

The best known type of BLEVE involves liquefied petroleum gas (LPG). Once a fire impinges on the shell above the liquid level, the vessel usually fails within 10–20 min. In the case of a BLEVE involving a flammable material, the major consequences are, in order of decreasing importance:

- Thermal radiation from the resultant fireball
- Fragments produced when the vessel fails
- Blast wave produced by the expanding vapor/liquid

For example, a BLEVE of a propane sphere with a diameter of 50 ft, holding about 630,000 gal, could cause damage as far away as 13,600 ft, and radiation damage and fragmentation damage would each extend to about 3,000 ft.

In a fire, a tank containing liquid is most vulnerable in the shell at the vapor space because very little heat can be absorbed by the vapor, and the metal in the vapor space can heat up rapidly to a temperature where it will weaken rapidly. The metal contacting the liquid phase will heat up much less rapidly because the liquid can absorb significant amounts of heat, keeping the shell temperature down in that area for a significant amount of time. Thus, there is a dilemma: a partly full vessel may BLEVE sooner than will a full vessel, but a full vessel will have more fuel for the resulting fireball and fire than will a partly empty vessel.

Significant equipment and building damage from radiation is possible from a BLEVE. Wooden structures may be ignited if the radiant heat density at the structure's location exceeds the threshold value for ignition of wood. Severe damage from fragmentation can be expected in the area where 50% or more of the fragments may fall, or typically about 300 ft from the vessel.

A BLEVE can lead to shock waves, projectiles, and thermal radiation. The effects of a shock wave and projectiles were dealt with earlier; by far the most serious consequence of a BLEVE is the radiation received from the fireball. The following calculational procedure is used to determine thermal impact (details are available in CPQRA [34]):

Damage Estimates [17]: Damage estimates deal with the consequences of explosions and thermal radiation to both people and property. Physical models for explosions and thermal radiation generate a variety of incident outcomes: shock wave overpressure estimates, fragment velocities, and radiant flux. These models rely on the general principle that severity of outcome is a function of distance from the source of release. In addition to estimating the damage resulting from an explosion, it is also necessary to estimate how the consequences of these incident outcomes depend on the object of the study. To assess effects on human beings, damage estimates may be expressed as deaths or injuries. If physical property is the object, the damage estimates may be expressed as monetary losses.

Explosion Consequences: A principal parameter characterizing an explosion is the overpressure. Explosion effect modeling generally is based on TNT explosions to calculate the overpressure as a function of distance. Although the effect of a TNT explosion differs from that of a physical or a chemical explosion (particularly in the near-field), the TNT model is the most popular because a large data base exists for TNT explosions.

More recently the explosion multi-energy method was developed by an international group in Europe. This model is the best representation of overpressures at far distances. The mass of vapor in congested volume participates in explosion. This approach requires intricate volume calculation and estimate of congestion. The TNT model overpredicts pressures at the source and underpredicts at longer distances. However, when you have overpressures of 5 psi or so it really doesn't matter as most structures and equipment at that level are destroyed [2].

Several kinds of energy may be released in an explosion; three basic types are: (1) physical energy, (2) chemical energy, and (3) nuclear energy. Nuclear energy is not considered here. Physical energy may take such forms as pressure energy in gases, strain energy in metals, or electrical energy. Chemical energy derives from a chemical reaction. Examples of explosions involving chemical energy are runaway exothermic reactions, including decomposition and polymerization.

Table 2.5 Effect of explosion overpressure on structures (Copyright 1989 by the American Institute of Chemical Engineers, Reproduced by permission of the Center for Chemical Process Safety of AIChE [34])

Pressure (psi)	Damage
0.02	Annoying noise (137 dB if of low, 10–15 Hz frequency)
0.03	Breaking of large glass windows under strain
0.04	Loud noise (143 dB), sonic boom, glass failure
0.10	Breakage of small glass windows under strain
0.15	Typical pressure for glass breakage
0.30	"Safe distance" (probability 0.95 of no serious damage below this value); projectile limits; some damage to house ceilings; 10 % window glass broken
0.40	Limited minor structural damage
0.5–1.0	Large and small windows usually shattered; occasional damage to window frames
0.70	Minor damage to house structures
1.00	Partial demolition of houses; houses made uninhabitable
1–2.00	Corrugated asbestos shattered; corrugated steel or aluminum panels, fastenings fail, followed by buckling; wood panel fastenings of standard housing fail; panels blown in
1.30	Steel frames of clad buildings slightly distorted
2.00	Partial collapse of walls and roofs of houses
2–3.00	Concrete or cinder blocks shattered if not reinforced
2.30	Lower limit of serious structural damage
2.50	50 % destruction of brickwork of houses
3.00	Heavy machines (300 lb), industrial buildings suffered little damage; steel frame buildings distorted and pulled away from foundation
3–4.00	Frameless, self-framing steel panel building demolished; rupture of oil storage tanks
4.00	Cladding of light industrial buildings ruptured
5.00	Wooden utility poles snapped
5–7.00	Nearly complete destruction of houses
7.00	Loaded railcars overturned
7–8.00	Brick panels, 8–12 in. thick, not reinforced, fail by shearing or flexure
9.00	Loaded train boxcars completely demolished
10.00	Probable total destruction of buildings; heavy machine tools (7,000 lb) moved and badly damaged; very heavy machine tools (12,000 lb) survive
300.00	Limit of crater lip

Table 2.5 summarizes the effects of explosion overpressure on structures. With respect to human casualties, heavy building damage usually is equated to a fatal effect, as the people inside the buildings probably would be crushed. People outside of buildings or structures are susceptible to direct blast injury (blast overpressure) and indirect blast injury (missiles or whole body translation).

Relatively high blast overpressures (>15 psig) are necessary to produce a human fatality from a direct blast. Instead, the major threat is produced by missiles or by whole body translation. Fatalities arising from whole body translation are mainly due to head injury from decelerative impact. Injury to people due to fragments usually results from either penetration by small fragments or blunt trauma from large fragments. TNO [22] suggested that projectiles with a kinetic energy of 100 J can cause fatalities. Table 2.6 shows damage to people (physiological damage) as a function of overpressure.

Table 2.6 Physiological damage as a result of overpressure

Effect	Peak overpressure (psi)
Knock down	1.0
Ear drum damage	5.0
Lung damage	15
Threshold for fatalities	35
50 % fatalities	50
99 % fatalities	65

Radiation Consequences: The effect of thermal radiation on people and objects is determined by one of two approaches:

1. Simple tabulations based on experimental results
2. Theoretical models based on the physiology of the skin burn response

Data on time to pain threshold [34] are summarized in Table 2.7. For comparison, solar radiation intensity on a clear, hot summer day is about 320 Btu/h/ft^2 (1 kW/m^2).

Table 2.7 Time to pain threshold for varying levels of radiation [41] (Courtesy American Petroleum Institute)

Radiation intensity (Btu/h/ft^2)	Radiation intensity(kW/m^2)	Time to pain threshold (s)
500	1.74	60
740	2.33	40
920	2.90	30
1,500	4.73	16
2,200	6.94	9
3,000	9.46	6
3,700	11.67	4
6,300	19.87	2

Table 2.8 Effects of thermal radiation (Copyright American Institute of Chemical Engineers, reproduced by permission of the Center for Chemical Process Safety of AIChE [34].)

Radiation intensity (kW/m^2)	Observed effect
37.5	Sufficient to cause damage to process equipment
25.0	Minimum energy required to ignite wood at indefinitely long exposures
12.5	Minimum energy required for piloted ignition of wood, melting of plastic tubing
9.5	Pain threshold reached after 6 s; second-degree burns after 20 s
4.0	Sufficient to cause pain to personnel if unable to reach cover within 20 s; however, blistering of the skin (second degree burns) is likely; 0% lethality
1.6	Will cause no discomfort for long exposure

Other criteria for thermal radiation damage are shown in Table 2.8 [34].

The effect of thermal radiation on structures depends on whether they are combustible or not, and the nature and duration of the exposure. Thus, wooden materials will fail because of combustion, whereas steel will fail because of thermal lowering of the yield stress.

Unconfined Vapor Cloud Explosions: When a large amount of volatile material is released rapidly to the atmosphere, a vapor cloud forms and disperses. If the cloud is ignited before it is diluted below its lower flammability limit, an uncontrolled vapor cloud explosion will occur. This is one of the most serious hazards in the process industries. Both shock waves and thermal radiation will result from the explosion, with the shock waves usually the more important damage producers. UVCEs usually are modeled by using the TNT model [34]. The energy of the blast wave generally is only a small fraction of the energy available from the combustion of all the material that constitutes the cloud; the ratio of the actual energy released to that available frequently is referred to as the “explosion efficiency.” Therefore, the TNT weight equivalent of a UVCE includes an explosion efficiency term, which typically is an empirical factor ranging from 1 to 10%. The explosion effects of a TNT charge are well documented.

Physical Explosions: A physical explosion usually results from the production of large volumes of gases by non-chemical means. The gases necessary for a physical explosion may be those already existing, such as compressed nitrogen released suddenly from a ruptured cylinder, or steam released explosively from a crack in a steam drum.

The following are some settings and situations in which physical explosions have been known to take place:

- Steam boilers
- Hydraulic overfill of tanks or pipes with external applied pressure (as in pressure testing)
- Compressed air tanks
- Deadheaded pumps
- Thermal expansion of tanks or pipes
- Liquid cryogenic fluids on water (such as liquid methane on water)
- Water suddenly mixed with sulfuric acid (also may cause a chemical explosion)
- BLEVE with superheated liquid (flammable or nonflammable) (see next section)
- Explosion of grinding wheel at too high a speed
- Liquid water in molten mgcl$_2$ solution at high temperatures
- Implosions due to vacuum
- Overpressured refrigerant systems
- Molten metals exploding violently on contact with water
- Some molten metals exploding when mixed with each other
- The mixing of two immiscible liquids whose boiling points are not widely separated

Steam boilers are commonly used in power plants and industries of all kinds. They generally are taken for granted now, but in the second half of the nineteenth century boilers blew up with alarming regularity. Records indicate that from 1870 to 1910 there were at least 10,000 boiler explosions in the United States and adjacent areas of Canada and Mexico; that is, more than one recorded explosion every 36 h! By 1910, the rate had jumped to between 1,300 and 1,400 per year. On October 8, 1894, in the Henry Clay Mine in Shamokin, Pennsylvania, 27 boilers disintegrated almost simultaneously! Mainly because of the incorporation of the ASME Boiler Code into laws, boiler explosions have decreased dramatically [35].

When a pressurized vessel ruptures, the resulting stored energy is released. This energy can cause a shock wave and accelerate vessel fragments. If the contents are flammable, ignition of the released gas could produce additional effects. There is a maximum amount of energy in a bursting vessel that can be released, and it is released in the following proportions: [32]

	Distribution of energy when vessel ruptures		
Type of failure	Strain energy	Kinetic energy of fragments	Shock wave energy
Brittle failure	<10%	~20%	up to 80%
Plug ejection	small	up to 60–80%	20–40%

The relative distribution of these energy components will change over the course of the explosion, but most of the energy is carried by the shock wave with the remainder going to fragment kinetic energy. To estimate the damage resulting from the shock wave from a physical explosion, the TNT model is used widely. To determine the TNT equivalent of a physical explosion, the total energy in the system must be estimated. For a physical explosion, if the expansion occurs isothermally, and ideal gas laws apply, then the TNT equivalent of the explosion can be calculated. This energy then can be used to estimate overpressure at any distance from the explosion. The analogy of the explosion of a container of pressurized gas to a point source explosion of TNT is not appropriate in the near-field. Prugh [36] suggests a correction method using a virtual distance R_v from an explosion center.

In addition to shock wave effects, a major hazard of a ruptured gas-filled vessel is from projectiles. To estimate damage from projectiles, both the initial velocity and the range are required. A simplified method for calculating the initial velocity uses the following equation: [37]

$$u = 2.05\left(\frac{PD^3}{W}\right)^{0.5}$$

where
u = initial velocity, ft/s.
P = rupture pressure, psig.
D = fragment diameter, in.
W = weight of fragments, lb.

Clancey [38] gives the following values for initial velocity for the majority of fragments from a TNT explosion:

- Thin case: 8,000 ft/s
- Medium case: 6,000 ft/s
- Thick case: 4,000 ft/s

Once the initial velocity has been determined, the maximum range of the fragment, ignoring air resistance, can be estimated from

$$R_{\max} = \frac{u^2}{g}$$

Where $R_{\max}$ is the maximum range of fragments and g is the acceleration of gravity.

If the above values for typical velocity are substituted into the above equation, a maximum range of 5×10^5 ft is possible. Therefore, it is clearly necessary to include air resistance. To include air resistance, a value of C_D, the drag coefficient, must be estimated. The drag coefficient ranges from 0.48 for a sphere to 2 for flow perpendicular to a flat strip, and for most fragments ranges from 1.5 to 2.0.

If one knows the air density, drag coefficient, exposed area of the fragment, mass of the fragment, and the initial velocity, the maximum range R can be calculated with the aid of Fig. 2.13 [18]. Although this technique gives the maxim-um range, most fragments do not travel the maximum distance but fall at distances between 0.3 and 0.8 of the maximum.

The energy required to impart an initial velocity of u to a fragment is

$$E = \frac{1}{2}mu^2$$

where m is the mass of the fragment (lb)u the initial velocity (ft/s)

Example: A high pressure vessel containing air at 600 bar has ruptured, leading to 15 fragments of approximately equal mass (85 lb), one of which was found as far as 400 ft from the vessel. This fragment has a drag coefficient of 1.5 and an exposed area of 3 ft^2. Assuming that 20% of the explosion energy went to energy of the fragments, estimate the energy of the explosion in weight equivalent TNT. The air density is 0.081 lb/ft^3.

Procedure: Assuming that the fragment found at 400 ft is at the maximum range for the fragments, the scaled fragment range R_s can be calculated:

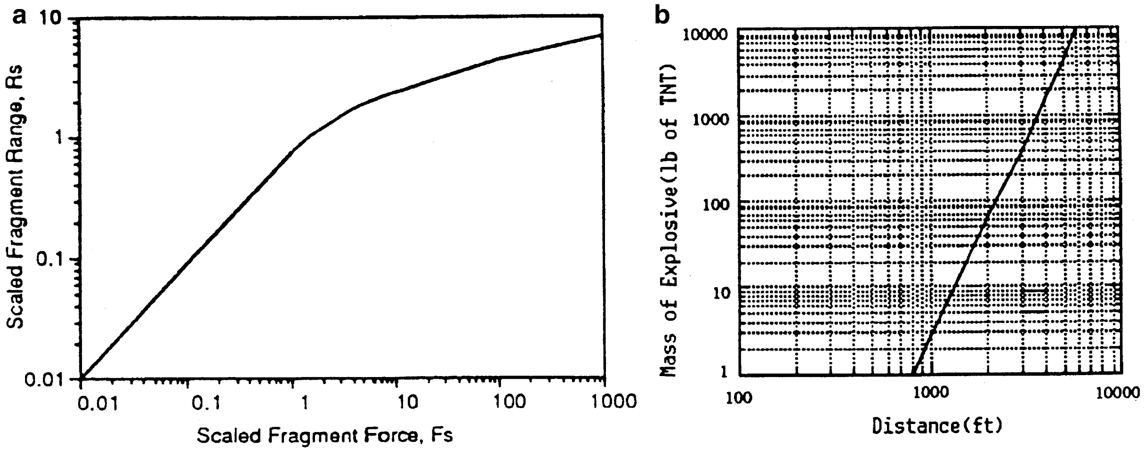
From Fig 2.13, we obtain a scaled force (F_s) of approximately 5. The initial velocity of the fragment then can be calculated as

$$u = \sqrt{\frac{MgF_s}{r_0 C_D A_D}} = \sqrt{\frac{85 \times 32.17 \times 5}{0.081 \times 1.5 \times 3}} = 194\,\text{ft/s}$$

The energy required to give the fragment this initial velocity is

$$E = \frac{1}{2}(85)(194)^2 = 1.6 \times 106\,\text{lb ft}^2/\text{s}^2 = 64\,\text{BTU}$$

Since there were 15 fragments, the total energy of the explosion that went into fragment kinetic energy is 15 × 635.8 BTU = 9,537 BTU. If only 20% of the explosion energy went into fragment kinetic energy, then the total



where: $F_s = \frac{r_o C_D A_D u^2}{Mg}$ $R_s = \frac{r_o C_D A_D R}{M}$

and: r_o = density of air lb/cu ft C_D = drag coefficient

A_D = exposed area of fragment, ft2 R = max range of fragment, ft

M = mass of fragment, lb U = fragment initial velocity, fps

$g = 32.17\ \text{ft. lb.}_m/\text{lb}_f\text{s}^2$

Fig. 2.13 (**a**) Scaled fragment range vs. scaled force. (Baker et al. [18]) (**b**) Maximum horizontal range of blast fragments. (Clancey [38])

explosion energy is 47,680 BTU, which is the equivalent of 23.8 lb of TNT. Using the method of Clancey [38], 2.4 lb TNT can provide a maximum range of 950 ft for projectiles (Fig. 2.13b).

Mechanical Heat

Mechanical motion in fluids becomes kinetic energy and may become heat in devices with rotating parts. Mechanical heat input from rotating agitators, pump impellers, and other mechanical equipment must be taken into account in the design of process equipment, particularly in systems containing reactive chemicals. This section will provide some guidelines for the analysis of individual cases involving pumps and agitated tanks [20, 23, 39].

Some useful rules are as follows.

1. A deadheaded pump is a pump operating full of liquid and with inlet and outlet valves closed.
2. Almost all deadheaded centrifugal pumps with motors of three horsepower or larger are headed for trouble if left deadheaded. (Depending on the horse power, a few minutes may be too long.)
3. The heat input from the rotating impeller in a deadheaded centrifugal pump is always a large value relative to the heat sink of the fluid and the pump.
4. It is not necessary for there to be a chemical reaction in a pump for an explosion to take place. Deadheaded pumps containing only water or brine have blown up.
5. An agitator or a circulating pump left on in a vessel of a reactive chemical may heat up the contents enough to cause a runaway reaction.
6. All centrifugal pumps with motors larger than 3 hp should be protected in some way to prevent deadheading.
7. A temperature alarm in the casing is a minimum form of protection. A better way may be to have the high-temperature alarm wired to the process control computer, to both alarm and shut off the pump. Other systems are available and may be used; they may include (but are not limited to) a relief valve on the pump, a minimum flow valve, and a flow orifice in the recirculating line. A relief valve on a pump relieving back to the pump inlet may not eliminate the problem of heat buildup in a deadheaded pump and usually should be avoided unless other protective measures are used such as a high-temperature device.
8. An ammeter on the pump motor usually is not a reliable means of detecting deadheaded conditions. The low power factors often experienced with pump motors, and the nature of pump curves, often make it difficult to distinguish between normal running and dead headed conditions using an ammeter.
9. For mechanical heat equivalent, the following are recommended: (a) For pumps, use 50% of the connected motor horsepower for centrifugal pumps that are deadheaded, unless better information is available. (b) For agitators, use 100% of the vendor rated shaft input

horsepower for the input shaft (total power less drive and bearing inefficiencies) for the actual material in the vessel.

Vacuum [39]

Ask any chemical engineers who have had some plant experience what they know about vacuum, and they probably will smile and tell a tale about some piece of equipment that tried to turn itself inside out. Usually no one was hurt, and often there is no massive leakage—but not always!

The design for the internal pressure condition of vessels usually is straightforward and well understood. Under vacuum conditions, equipment is subject to external pressure from the atmosphere; and the design for external pressures is more difficult than that for internal pressures. The devious ways in which external pressure can be applied often may be overlooked.

The following are some obvious causes of vacuum collapse:

- Liquid withdrawal by pump or gravity draining
- Removal of gas or vapor by withdrawing with a blower, fan, or jet
- Siphoning of liquids

Less obvious causes include:

- Condensation of vapor
- Cooling of hot gas
- Combination of cooling and condensation of a mixture of gas and condensable vapor

Sometimes obscure causes of vacuum collapse include:

1. Absorption of a gas in a liquid; for example, ammonia in water, carbon dioxide in water, hydrogen chloride in water.
2. Reaction of two or more gases to make a liquid or solid; for example, ammonia plus hydrogen bromide to form ammonium bromide.
3. Reaction of a gas and a solid to form a solid; for example, corrosion in a tank, air plus Fe or FeO forming Fe_2O_3 in the presence of water.
4. Reaction of a gas and a liquid to give a liquid; for example, chlorination, hydrogenation, ethylation.
5. Sudden dropping of finely divided solids in a silo, creating a momentary vacuum that can suck in the sides of the silo.
6. Flame arrestors plugging; for example:
   (a) In styrene service, vapor may condense in flame arrestors, and the liquid formed is low in inhibitor; the liquid may polymerize and plug off the arrestor. Possible solutions: clean the arrestor frequently or use a PVRV (pressure-vacuum-relief valve).
   (b) Liquid service in cold weather: vapor may condense in a flame arrestor and the liquid formed may freeze and plug the arrestor. Possible solution: heat and insulate the arrestor to prevent condensation.
7. Maintenance and testing. It is not a good idea to apply vacuum on a vessel during maintenance or testing without full knowledge of the external pressure rating unless a suitable vacuum relief device is in place and operable.

Protective Measures for Equipment

If equipment may be subject to vacuum, consideration should be given to designing the equipment for full vacuum. This may eliminate the need for complicated devices such as vacuum relief valves and instruments; if they are used, designing the equipment for full vacuum will prevent collapse of the vessel if the instruments or relief valves fail or plug.

A disadvantage of this approach is that it usually is expensive. However, when the total cost of a suitably instrumented vessel not designed for vacuum is compared with the cost of a vessel designed for vacuum but without the extra equipment, the difference may be small or negligible, and the vessel designed for vacuum will be inherently safer. If a vessel is designed for vacuum, precautions should be taken to ensure that internal or external corrosion will not destroy the integrity of the vessel.

Reactivity Hazards [2]

As a nation, we continue to have chemical reactivity incidents that cause harm to people, property and the environment. The Chemical Safety Board's report analyzed 167 incidents from 1980 to 2001 that resulted in a total of 108 fatalities and significant property damage. While this number may seem small in comparison to, say the number of automobile related fatalities annually in the US, it is significant because the data used are admittedly incomplete, leading to the expectation that the "true" impact of chemical reactivity incidents is much higher.

By way of definition, the CCPS Concept Book intentionally uses the term "chemical reactivity hazard" rather than "reactive hazard," "reactive chemical hazard," or "chemical reaction hazard."

A chemical reactivity hazard is defined as a situation with the potential for an uncontrolled chemical reaction that can result in serious harm or loss.

According to the CSB report, in a vast majority of these cases, the information needed to properly assess (and therefore, control) these hazards was known prior to the incident. In 90% + of all incidents studied, the information necessary to have prevented the incident was documented and publicly available.

As the Chemical Safety Board has concluded in its Hazard Investigation, the problem is not reactive chemicals but managing reactive chemicals.

Reactivity of chemicals provides us with much of the materials necessary for modern civilization, but the hazards associated with reactive chemicals must be controlled.

- It's not reactive chemicals, it's reactive chemistry—and the management of its hazards.

- Many reactive chemistry incidents have occurred in operations where there was no intended chemical reaction (storage, blending, distillation, etc.).
- Reactive chemicals can be reactants, intermediates and products.

An approach to reactivity hazard evaluation used by a major chemical company that is a leader in reactive chemical safety includes the following steps.

- Identify all chemicals used in the process (raw materials, intermediates and products).
- Obtain reactivity information and data on the above including inadvertent mixing—testing may be required.
- Use a team approach to identify what can go wrong.
- Determine the consequences of all possible deviations.
- Calculate the worst case scenario.
- Identify and implement appropriate safe guards.

Many reactive chemical incidents take place when there is no chemical reaction intended. Such an incident occurred April 21, 1995 at Napp Technologies in Lodi, New Jersey. The Napp incident involved inadvertent mixing of water with a water reactive chemical (aluminum powder and sodium hydrosulfite) during a blending operation. There were five fatalities, evacuation of 300 people and major property damage and loss of business. Neither aluminum powder or sodium hydrosulfide are included in the OSHA PSM standard.

The first step in managing a reactive hazard is to identify that there is a hazard. This can be done by literature surveys, energy of reaction, chemical structures and bonds, and interaction matrixes. It helps to have someone is your organization that has special expertise on reactive hazards.

It is important to have enough data to describe the hazard and to provide control measures. Special expertise is required to do this right.

There is plenty of literature, vendor information and other resources to help. General types of reactivity testing:

– DOT/UN tests
– Screening tests
– Reaction calorimetry
– Detailed hazard testing
– Special studies

There are numerous testing methods that can be used depending on the hazard information required. If more detailed information is required for engineering calculations, such as relief valve sizing, different methods should be used.

Reactive Chemistry References:

- CCPS, Guidelines for Chemical Reactivity Evaluation and Application to Process Design, 1995.
- CCPS (2003), Essential Practices for Managing Chemical Reactivity Hazards, AIChE, NY.
- Bretherick, Handbook of Reactive Chemical Hazards Vols. 1 and 2, Elsevier 2007.
- Sax, Dangerous Properties of Industrial Materials (Lewis and Irving, 2001).
- CHRIS Chemical Hazards Response Information Systems (US Coast Guard).
- NFPA Stds. 49, 325, 432, 491.
- ASTM CHETAH (Balaraju et al. 2002).

There are standards that can help you construct binary chemical compatibility charts. These charts are useful when you handle a variety of chemicals in an area and are concerned with inadvertently mixing any two chemicals. Factors that should be considered when specifying the mixing scenario include material quantities and temperatures, degree of confinement, atmosphere, and the maximum time the materials may be in contact. Do this for all chemicals including warehousing, cleaning chemicals, etc.

- ASTM 2012–00 Standard Guide for Preparation of a Binary Compatibility Chart.

Chemical Reactivity Worksheet, Version 2.0.2, NOAA/CCPS, last updated August 2010. http://response.restoration.noaa.gov/chemaids/react.html.

It is important to have enough data to describe the hazard and to provide control measures. Special expertise is required to do this right.

- Sufficient data to fully characterize reactive hazards is sometimes available from supplier or other sources.
- Work with someone who knows:
  – The various tests and their limitations
  – When they should be conducted
  – How to interpret and use the results

To summarize, if you handle chemicals, you need a reactive chemicals program. Some important aspects include . . .

- Ownership of the reactive chemistry
- Reactive chemical (life cycle) reviews
- Screening and testing protocols
- Screening and testing facilities
- Capture of reactivity hazard data
- Availability of experts to participate in reviews, be available for consultation

Toxicity Hazards [2]

Toxicity is a life safety issue and there are different ways to express the threat. A particular toxin can represent hazards in more than one type of exposure. Fluorine has both serious skin contact and respiratory issues. Respiratory rate can also contribute to the extent of the exposure. Toxicity is a measure of harm from direct exposure to certain chemical substances. It measures potential life threatening exposures including oral, skin contact and respiratory. Life health risk is a function of type of contact, toxicity and duration of exposure. Reaction of individuals to a specific toxin varies. Toxicity data do not reflect chronic health effects. Examples of highly toxic chemicals include:

- Acetic anhydride
- Acrylamide
- Aniline

- Arsenic compounds
- Benzoyl peroxide
- Barium salts
- Fluorine, Chlorine, Bromine
- Formaldehyde
- Hydrazine
- Hydrogen sulfide
- Mercury compounds
- Nitrobenzene
- Nickel carbonyl
- Oxalic acid
- Phenol
- Phosgene
- Pthalic anhydride
- Propylamines
- Silver salts
- Tetrachlorethane
- Sulfur dioxide

Toxicity includes:

- Acute exposure
  - A single exposure to a toxic substance which may result in severe biological harm or death; acute exposures are usually characterized as lasting no longer than a day and
- Chronic exposure
  - Continuous exposure to a toxin over an extended period of time, often measured in months or years can cause irreversible side effects

Process safety is primarily concerned with acute exposure. Some common terms used to describe toxicity include:

- LC 50—Concentration of a material in air that will kill 50%of the test subjects (typically animals) when administered as a single exposure (typically 1–4 h)
- LC Lo—Lowest concentration reported to have killed animals or humans
- LD 50—Dose required to kill half the members of a tested population
- LD 50—Dose at which 50% of a tested population are killed

There are many ways to define and present acute exposure levels. These definitions and exposure levels have been created with help from organizations like the American Conference of Governmental Hygienists (ACGIH). The data is normally hard to find for acute effects and an even more extensive amount of data, say compared to an IDLH, is necessary to produce a probit relationship which will map out all the combinations of concentration and time for a particular probability of fatality. Probits have application for detailed QRAs.

Some other terms to describe acute toxicity include:

- ERPG—Emergency Response Planning Guidelines
  ERPG 1—1 h exposure mild health effects
  ERPG 2—1 h exposure w/o/ irreversible health effects
  ERPG 3—1 h exposure w/o life threatening health effects
- IDLH—Immediately Dangerous to Life and Health
  Maximum airborne concentration to which a healthy male worker can be exposed up to 30 min and be able to escape without loss of life or irreversible organ damage
- Probit—Dose/Response Algorithm
- EEGL/SPEGL—Emergency Exposure Guidance Levels

Asphyxiants are a special class of toxic gases. Asphyxiants are normally inert gases such as nitrogen, argon, carbon dioxide and others. Nitrogen is the most prevalent in the chemical industry. There are multiple deaths because someone tries to be a hero in a rescue attempt. In some cases where they suspect the cause they think they can hold their breath long enough. Breathing air is normally produced from compressed air. However, air is sometimes manufactured as a mixture of nitrogen and oxygen. These mixtures are normally not acceptable for breathing air but if it is used that way there is always a chance that the ratio of nitrogen to oxygen is not correct. There are approximately 8 deaths/year from N_2 asphyxiation alone in the United States. Contributing causal factors for these incidents include in or near a confined space, inadequate monitoring, mix up of N_2 and breathing air and attempted rescue.

Corrosivity Hazard [2]

Rust is the most simple form of corrosion. Rust can result in a damaged hose clamp that cannot be tightened or removed due to rust deposition. The single biggest concern with corrosion is inside the equipment but external corrosion can also be a major concern. A broken hose clamp can cause a hose to slip off a connection. Corrosion is a chemical reaction between a metal and its environment. Common corrosion (rust) requires air and moisture. The corrosive layer can weakened a pipe or equipment structurally and thus initiate a failure. Corrosion or erosion often occurs in pipe elbows where high velocity can scour corrosion products exposing additional metal to corrosion. Common corrosion rates in pipe wall are 0.1–0.2 mm/year, but corrosion rates may increase tenfold in highly corrosive or erosive service.

The most common example of metal oxidation is rust but other metal oxides can also oxidize (e.g. aluminum). Insulation can absorb moisture and act to dramatically increase the rate of rust formation. High temperature corrosion does not require the presence of a liquid electrolyte such as water. Oxidation is the major type of high temperature corrosion but you can also have sulfidation and carburization. Alloys often rely on the oxidation reaction to produce a protective scale. Galvanic corrosion is an electrochemical process in which one metal corrodes preferentially when in electrical contact with another metal and both metals are immersed in an electrolyte. The galvanic couple is set up and ions move from the anode to the cathode. The presence of electrolyte and a conducting path between the metals may cause corrosion. Underground piping can undergo this problem and sometimes sacrificial anodes such as zinc, magnesium or aluminum are used. One form of microbial corrosion is caused by two types of bacteria one that eats sulfates in the absence of oxygen and nitrates to form hydrogen sulfide (sulfate reducing bacteria) and another that eats the hydrogen sulfide to form sulfuric acid. This can be a problem when city wastewater is used in the plant. The city of Phoenix had problems a few years ago and firewater systems in plants in that area had severe corrosion using that water. The addition of as little as 3 ppm of chlorine eliminated that problem. Corrosion in passivated materials can produce localized pitting if the coating is not completely uniform. Be aware of whether materials passivated by a layer of oxidized material such as aluminum oxide are not in an environment where that passivated layer is continuously scraped away. Corrosion rates are normally in the range of mils/year. Clad vessels are sometimes used if normal materials have too high a corrosion rate for the service and alternative materials are too costly for the thickness needed. Additional thickness needed to compensate for an expected corrosion rate over the lifetime of the equipment Corrosion rates are dependent on piping or vessel materials of construction and chemical conditions including flow, concentration of corrosive chemical(s), temperature, and pressure.

Hazard Identification and Risk Analysis

The other element included in the pillar Understanding Hazards and Risk is Hazard Identification and Risk Analysis (HIRA). HIRA encompasses all activities involved in identifying hazards and evaluating risks to employees, the public and the environment at facilities, throughout the facility's life cycle, to control the risk within the organization's risk tolerance. HIRA addresses three questions:

- What can go wrong?
- How bad is it?
- How likely is it to occur?

HIRA Logic Diagram

The logic diagram (Fig. 2.14) explains the process of how to uncover hazards and how to analyze and address the risks. Note there is a recycle loop at the end. If risks are too high as judged by a company then ways for risk reduction must be sought and the final risk accepted. Risk is a function of the probability and consequences of an undesirable event which could occur as a result of the presence of a hazard. Another way to express risk is to say it is some function of the combination of probability that something might happen and the expected consequences if it does.

HIRA includes the following topics

- Hazard identification
- Qualitative hazard evaluation methodologies
- Quantitative hazard evaluation methodologies
- Consequence analysis
- Probability analysis

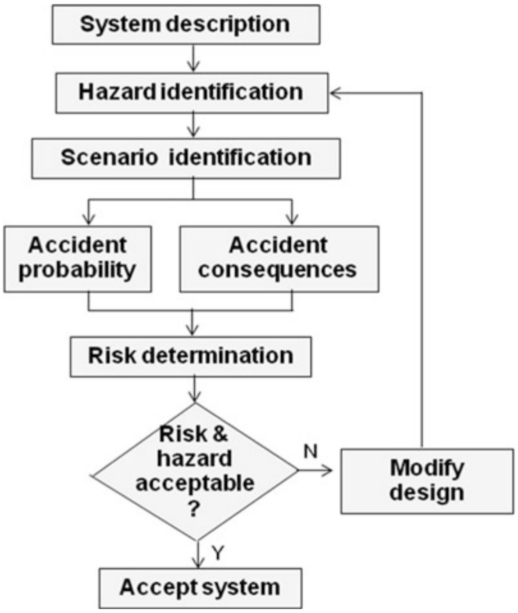


Fig. 2.14 HIRA logic diagram. Copyright 2010 Center for Chemical Process Safety (CCPS) of the American Institute of Chemical Engineers (AIChE), www.aiche.org/ccps and used with permission"

Hazard Identification

Hazard identification answers the question "What can go wrong?"

OSHA requires that Process Hazard Analyses (PHAs) be conducted on covered processes. PHAs use qualitative hazard evaluation methodologies.

PHAs include both the identification and evaluation of the hazards (refer to section on "Regulations"). Commonly methodologies used to conduct PHAs include:

- What if
- Checklist
- What if/Checklist
- FMEA (Failure Modes and Effects Analysis)
- FTA (Fault Tree Analysis)
- HAZOP

HAZOP stands for "Hazard and Operability Studies," a set of formal hazard identification and elimination procedures designed to identify hazards to people, processes, plants, and the environment. The techniques aims to stimulate the imagination of designers and operations people in a systematic way so they can identify potential hazards. In effect, HAZOP studies assume that there will be mistakes, and provide a systematic search for these mistakes. In some European countries, HAZOP studies are mandatory and attended by observers from regulatory authorities to ensure that the studies are carried out correctly. The examination of accidents [40] during 1988 at a large US chemical company revealed that the accidents could be classified as follows:

- Spills: 52%
- Emissions: 30%
- Fires: 18%

Of the fires, about 50% occurred during construction, 25% were due to pump seal failure, and the remaining 25% resulted from engineering and operational oversights that a HAZOP study possibly could have prevented.

Of the emissions, 37% were due to piping failure, with lined pipe being the largest contributor. Operational and procedural issues accounted for 53% of the remainder.

Of the spills, 11% were due to equipment failures. Piping failures (especially lined pipe and gaskets) accounted for 30, and 56% were caused by various types of operational errors, noncompliance with procedures, or nonexistent procedures. Material handling was a factor in many spills and emissions. The most frequent type of operational error was a valve being left in an improper position, either open or closed. HAZOP studies probably could have reduced the number and seriousness of the problems experienced. Some investigations have shown that a HAZOP study will result in recommendations that are 40% safety-related and 60% operability-related. Thus, HAZOP is far more than a safety tool; a good HAZOP study also results in improved operability of the process or plant, which can mean greater profitability.

The HAZOP technique can be used to identify human error potential. From a practical point of view, human error and its consequences can occur at all levels of a management structure as well as in the operation of a particular plant or process. Carried out correctly, Technica [41] states that a HAZOP study will identify at least 70–75% of potential operational and safety problems associated with a particular design process, including human error.

The HAZOP technique also can be used for the evaluation of procedures. Procedures may be regarded as a "system" designed to "program" an operator to carry out a sequence of correct actions. Deviations from intent are developed, with the emphasis on "operator action deviation" rather than "physical property deviation." It is the procedure, not the hardware, that is the object of study, but hardware modifications may be recommended to cover potential problems identified from procedure deviations.

Some Tools for Evaluating Risks and Hazards

Dow Fire and Explosion Index: The Dow Fire and Explosion Index (F&EI), developed by The Dow Chemical Company, is an objective evaluation of the potential of a facility for a fire, an explosion, or a reactive chemical accident. Its purpose is to quantify damage from incidents, identify equipment that could contribute to an incident, and suggest ways to mitigate the incident; it also is a way to communicate to management the quantitative hazard potential. It is intended for facilities handling flammable, combustible, or reactive materials whether stored, handled, or processed. The goal of the F&EI evaluation is to become aware of the loss potential and to identify ways to reduce the potential severity in a cost-effective manner. It does not address frequency (risk) except in a general way. The number is useful mainly for comparisons and for calculations of damage radius, maximum probable property damage, and business interruption loss, and to establish frequency of reviews. The method of carrying out an F&EI evaluation is available to the public from the American Institute of Chemical EngineersNew York, NY.

Failure modes and effects analysis (FMEA): FMEA is a systematic, tabular method for evaluating the causes and effects of component failures. It represents a "bottom–up" approach, in contrast with a fault tree, where the approach is "top–down." In large part, HAZOP is a well-developed form of FMEA [19].

Fault tree: A fault tree is a logical model that graphically portrays the combinations of failures that can lead to a particular main failure or accident of interest. A fault tree starts with a top event, which is usually a hazard of some kind. The possibility of the hazard must be foreseen before the fault tree can be constructed. A fault tree helps reveal the possible causes of the hazard, some of which may not have been foreseen [19].

Safety audit: A safety audit is a method of reviewing the actual construction and operation of a facility. Often, safety audits are conducted by a small interdisciplinary team. At least some of the members of the team are not connected with the plant. The audit may be carried out before startup and also is repeated later at intervals of, typically, 1–5 years.

Chemical exposure index: The Chemical Exposure Index is a technique for estimating the relative toxic hazards of chemicals, developed by The Dow Chemical Company. It provides for the relative ranking of toxic chemicals in a given facility, including factors relating to toxicity, quantity volatilized, distance to an area of concern, and physical properties. A description of the method can be found in *Guidelines for Safe Storage and Handling of High Toxic Hazard Materials*, Center for Chemical Process Safety, American Institute of Chemical Engineers [42].

The methods above are described in detail in *Guidelines for Hazard Evaluation Procedures, Third Edition,* CCPS, 2008 and will not be discussed in any more detail in this chapter.

Consequences and Impacts

The next question is "How bad is it?" Evaluating consequences is a very technical subject requiring special training and expertise. There are several good books on the topic in the references. Some general comments will be made here. The first step in evaluating consequences is to select a release scenario. How to choose a scenario is not always well defined. There is a requirement in the Seveso regulations that major hazards must be addressed. The top event of a fault tree may be another starting point. Possible sources of potential release incidents include:

- Hazard evaluation process
  - Scenario based methodologies most useful (e.g., HAZOP)
- Regulatory requirements
- Fault tree analysis
- Company initiative

The next step involves choosing a source model. Source models describe how a material escapes from a process. Use of source models should be referred to experts. If potential release is flammable and/or toxic the following are possible consequences:

- Fires
- Explosions
- Toxic Releases
- Environmental Pollution

The above consequences have already been discussed in the section on “Chemical Hazards”.

Probability

The last HIRA question is “How likely is it to occur?” Generating the frequency or probability of an event requires both technical competence in the calculations and experience. Evaluation methodologies to estimate frequency or probability include:

- Fault tree analysis (FTA)
- Event tree analysis (ETA)
- Level of protection analysis (LOPA)
- Cause–consequence analysis
- Human reliability analysis (HRA)

Detailed description on the methodologies can be found in *Guidelines for Hazard Evaluation Procedures, Third Edition*, CCPS, 2008. Again, there will not be a detailed description of these methodologies in this chapter.

There are always questions about sources of failure data. Experienced analysts can help work through this maze. In some cases frequency information can be obtained from the plant history (e.g., number of times a relief valve has gone off in anger) which comes in at a higher point in the tree. Predicted data relates to establishing a failure rate from the sum of all the individual component failures. PERD is an AICHE CCPS organization (Process Equipment Reliability Data) where companies contribute data and then have access to others data.

Common cause failures must be accounted for when assigning failure rates. See the references at the end of this chapter for other sources of frequency or probability data.

Risk Analysis

Risk is a combination of the consequence if someone is impacted by a hazard combined with the expected frequency of being impacted by that hazard. For instance if a person steps in a hole in the floor what would the consequence be? It depends on a lot of things—including how deep the hole is, the sharpness of the edges of the hole, etc. . . . Risk is often expressed in terms of probability/likelihood and impact/consequence.

Risk Understanding

As Fig. 2.15 illustrates, risk can be managed by managing the probability and the consequence of occurrence. For probability the question asked is—How likely is it? For consequence the question is—What can go wrong. . .and if it does what are the impacts? Answers are often based on our personal past experiences, what has been learned from others experiences, and/or by using some analytical (and perhaps technical) methodologies. Risk estimates are based on exposure to impact, magnitude of impact and probability. The risk estimate must be related to base level of risk. What principal factor determines the magnitude of the risk? If risk is a function of the probability and consequences then risk can be reduced by lowering the probability an incident will occur, reducing the consequences of occurrence, or a combination of both. Lowering the probability of occurrence is the prevention approach. Reducing the consequences of occurrence is the mitigation approach. It is best to try and prevent before trying to mitigate. Risk analysis followed by risk assessment is usually needed to support either approach.

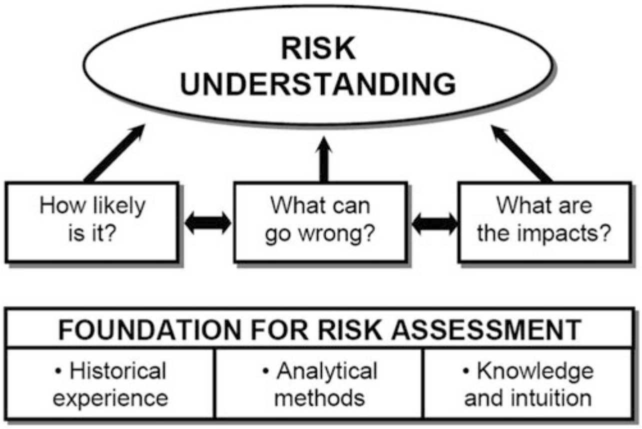


Fig. 2.15 Risk understanding (Copyright 2008 Center for Chemical Process Safety (CCPS) of the American Institute of Chemical Engineers (AIChE), www.aiche.org/ccps, and used with permission”)

Good risk analysis has the following attributes:

- Level of effort commensurate with the risk
- Experienced analysts
- Uncertainties defined
- Options for risk reduction identified
- Risk analysis documented clearly and understandably
- Risk analysis is both defendable and repeatable

There are several approaches to risk analysis including qualitative risk ranking, semi—quantitative analyses using risk matrices, layer of protection analysis (LOPA), and quantitative analysis (QRA). These approaches will not be discussed in detail here. See the references at the end of the chapter.

Quantitative Risk Analysis

Quantitative risk analysis (QRA) models the events, incidents, consequences, and risks, and produces numerical estimates of some or all of the frequencies, probabilities, consequences, or risks [34, 43]. QRA can be done at a preliminary level or a detailed level, and in all cases may or may not quantify all events, incidents, consequences, or risks [36]. QRA is the art and science of developing and understanding numerical estimates of the risk associated with a facility or an operation. It uses highly sophisticated but approximate tools for acquiring risk understanding.

QRA can be used to investigate many types of risks associated with chemical process facilities, such as the risk of economic losses or the risk of exposure of members of the public to toxic vapors. In health and safety applications, the use of QRA can be classified into two categories:

1. Estimating the long-term risk to workers or the public from chronic exposure to potentially harmful substances or activities.
2. Estimating the risk to workers or the public from episodic events involving a one-time exposure, which may be acute, to potentially harmful substances or activities.

QRA is fundamentally different from many other chemical engineering activities (e.g., chemistry, heat transfer), whose basic property data are capable of being theoretically and empirically determined and often established experimentally. Some of the basic "property data" used to calculate risk estimates are probabilistic variables with no fixed values, and some of the key elements of risk must be established by using these probabilistic variables. QRA is an approach for estimating the risk of chemical operations by using the probabilistic approach; it is a fundamentally different approach from those used in many other engineering activities because interpreting the results of QRA requires an increased sensitivity to uncertainties that arise primarily from the probabilistic character of the data.

Safety Risk Criteria [44]

Risk is something that exists every day in our business and private lives. Each one of us has a tolerance level, very often based on our personal experiences. While that may be acceptable for us personally, we need to have a logical and documented way of making risk based decisions in the work place. Let's look at a standard diagram (Fig. 2.7) of how risk based decisions should be made in the work place.

1. First the system that is being analyzed must be understood. So, a system description is required and understood. In some sense, this is compiling the process safety knowledge.
2. Second, hazard identification process must be completed. PHAs are one route to hazards identification. Another way of identifying hazards is to review the process safety information and/or to do a walkthrough of the area.
3. Now, the hazard must be converted to a risk. Remember that risk is a combination of probability that something might happen combined with the consequences if it does.
4. With that in hand, an estimate of the risk can be made.
5. Now it needs to be determined if that risk can be tolerated or if risk must be reduced. So, the risk must be compared to the risk tolerance criteria.
6. If it passes the tolerance test, no changes are needed.
7. If it does not pass the tolerance test, then something must be changed. And that something is either the probability or the consequence.

There are a variety of ways to set the risk tolerance. It can be just a guess based on what is known or believed but that isn't very scientific and is certainly not repeatable or defendable in a court of law. Perhaps it can be copied from someone or some company that we know. That might work, but it may not consider all the things in our company that makes it unique. Or, we can make our own list and criteria based on company culture, beliefs, resources, etc. To do that we probably want to break our decision making process down into small pieces so that each decision has a relatively small impact on the final product. Often performance is evaluated in three areas—safety, environmental, and financial. Failure to pass the tolerance test in any of these areas causes the risk to be unacceptable. There may be other topics that need to be to detail out when making risk tolerance decisions. Whatever they are, write them down and get them universally accepted in your company. They will be the standard that a company will use for critical decisions. Company criteria must be developed so that it is defendable and repeatable. So, what is acceptable (and we really should use *Tolerable* instead of *Acceptable* since Acceptable seems to be an inflammatory word in general public use) This brings us back to our tolerance criteria. The CCPS book entitled *Guidelines for*

Developing Quantitative Safety Risk Criteria, CCPS, 2009 cover this topic in detail.

The benefits of risk management include that it identifies key exposures minimizes surprises, provides an objective basis for allocating resources, improves culture and puts risk to competitive advantage.

Manage Risk

The third pillar of RBPS is *Manage Risk*. To manage risk facilities should focus on:

- Developing written operating procedures
- Implementing an integrated suite of safe work policies, procedures, permits and practices to control maintenance and other non routine work.
- Executing work activities to ensure that equipment is fabricated and installed in accordance with specifications and that it remains fit for duty for service over its entire life cycle.
- Managing contractors, and evaluating work performed by contractors.
- Providing training.
- Recognizing and managing changes.
- Ensuring that units, and the people who operate them, are properly prepared for start ups.
- Maintaining a very high level of human performance.
- Preparing for and managing emergencies.

Operating Procedures

Operating procedures can be written (or electronic) documents that list the steps for a given task and describe the manner in which steps are to be performed. Procedures describe the process, the hazards, tools, protective equipment, and controls; provide instructions for troubleshooting, emergency shutdown, and special situations; describe the tasks necessary to safely start up, operate, and shutdown processes, including emergency shutdown; and provide formatted instructions.

Operating procedures are important because without written operating procedures a facility can have no assurance that the intended procedures and methods are used by each operator or even that an individual operator will consistently execute a particular task in the intended manner. Operating procedures are also a regulatory requirement for PSM covered facilities in the United States (see section on "Regulations"). The implementation of operating procedures requires the identification of when operating procedures are needed, the development of procedures, the use of procedures to improve human performance and the assurance that procedures are maintained.

Safe Work Practices

Safe work practices help control hazards and manage non-routine work. A non-routine activity is any activity that is not fully described in an operating procedure. Safe work practices typically control hot work, stored energy (lock-out/tag out), opening process vessels or lines, confined space entry, and similar operations as well as other routine highly hazardous operations.

Safe work practices are important because non-routine work increases risk and can lead to conditions that make a catastrophic event more likely. Some examples include:

- Piper Alpha (removal of a pressure safety valve for recertification).
- Nitrogen asphyxiation during confined space entry resulted in 80 fatalities from 1992 to 2002.

There are also regulatory requirements regarding safe work practices in the US.

Asset Integrity and Reliability

Asset integrity and reliability is the systematic implementation of activities, such as inspections and tests necessary to ensure equipment will be suitable for its intended application throughout its life. Specifically, work activities focused on preventing catastrophic release of a hazardous material or sudden release of energy and ensuring high availability or (dependability) of critical safety or utility systems that prevent or mitigate the effects of these types of events.

Asset integrity and reliability is important because designing and maintaining equipment fit for its purpose and functions when needed is paramount in maintaining containment of hazardous materials and ensuring that safety systems work when needed. These are two primary responsibilities for any facility. It is a PSM regulatory requirement under the element mechanical integrity. Mechanical integrity requirements emphasize the safety aspect rather than the reliability aspect, but both are important in managing risk.

Contractor (Safety) Management

Contractor (safety) management is a system of controls to ensure contracted services support both facility operations and the company's process safety and personal safety performance goals. It does not address the procurement of goods and supplies or offsite equipment fabrication functions that are covered by the asset integrity quality assurance function. It involves workers located closest to process hazards and more routine tasks such as janitorial and ground keeping services.

It is important because it facilitates the company in achieving the goals of accessing specialize expertise that is not continuously or routinely required, supplementing limited company resources during periods of unusual demand and providing staff increases without the overhead cost of direct hire employees. Companies and contractors must work together to provide a safe workplace that protects the workforce, the community, and the environment, as well as the welfare and interest of the company.

Training and Performance Assurance

Training is practical instruction in job and task requirements and methods. Performance assurance is the means by which workers demonstrate that they understood the training and can apply it in practical situations.

This element is important because a high level of performance is a critical aspect of any process safety program. A less than adequate level of human performance will adversely affect all aspects of operations. Without an adequate training and performance assurance program, a facility can have no confidence that work tasks will consistently be completed to minimum acceptable standards and in accordance with accepted procedures and practices. The implementation of training and performance assurance involves the principles to identify what training (and retraining) is needed, to provide effective training, and to monitor worker performance.

Management of Change

A change is any change (modification) to process chemicals, technology, equipment, or procedures and changes to facilities that affect a covered process except replacement in kind. A replacement which satisfies the design specification is not a change. Management of change (MOC) helps to ensure that changes to a process do not inadvertently introduce new hazards or unknowingly increase the risk of existing hazards. MOC includes a review and authorization process for evaluating proposed adjustments to facility design, operations, and organizations. It is a system to ensure that all introduced changes are thoroughly scrutinized prior to implementation. More than 80% of large losses are related to change. In a MOC system all changes are evaluated, communicated and coordinated prior to execution. A rational basis is required to initiate the process. MOC applies to physical equipment, products, operating conditions, staffing and including organizational changes.

It is important because if a proposed modification is made to a hazardous process without appropriate review, the risk of a process safety incident could increase significantly. The principles of MOC are to identify potential change situations, evaluate possible impacts, decide whether to allow change, and complete follow-up activities.

Operational Readiness

Operational readiness ensures that shut down processes are verified to be in a safe condition before re-start. It is defined more broadly than OSHA PSM prestart up safety review element because it addresses start up from all shut conditions, not only from those resulting from new or changed processes.

Operational readiness is important because experience has shown that the frequency of incidents is higher during process transitions such as startups.

It is important that the process be verified as safe to start. The principles of operational readiness include the following: to conduct appropriate readiness reviews as needed, make startup decisions based on the readiness results, and to follow through on decisions, actions, and the use of readiness results.

Conduct of Operations

Conduct of operations involves the execution of operational and management tasks in a deliberate and structured manner. It institutionalizes the pursuit of excellence in the performance of every task and minimizes variations in performance. Some companies call this Operating Discipline (walk the talk).

It is important for several reasons. A consistently high level of human performance is a critical aspect of any process safety management program. A less than adequate level of human performance will adversely impact all aspects of operations. As operational activities become more complex, an increase in the formality of operations must also occur to ensure safe, and consistent performance of critical tasks. The principles of conduct of operations include the control operational activities, control the status of systems and equipment, development of required skills/behaviors and the monitoring of organizational performance.

Emergency Management

Emergency management includes:

- Planning for possible emergencies
- Providing resources to execute the plan
- Practicing and continuously improving the plan
- Training or informing employees, contractors, neighbors, and local authorities on what to do, how they will be notified, and how to report an emergency
- Effectively communicating with stakeholders in the event an incident does occur

It is important the consequences of any particular incident can be significantly reduced with effective emergency planning and response. Effective emergency management saves lives and protects property and the environment. It also helps reassure stakeholders that, in spite of the incident, the facility is well managed and should be allowed to continue to operate. It is a PSM Regulatory requirement. The principles of emergency management are to prepare for emergencies, and to periodically test the adequacy of plans and level of preparedness. Chap. 3 , Managing an Emergency

Preparedness Program has an in-depth discussion on this topic.

Learn from Experience

The fourth pillar of process safety management is *Learn from experience*.

To learn from experience facilities should focus on:

- Investigating incidents.
- Applying lessons from incidents that occur at other facilities in the company and the industry.
- Measuring performance and striving to continuously improve in areas of significant risk.
- Auditing RBPS management systems and work activities.
- Holding periodic management review to see if things are working and helping to manage risk.

Incident Investigation

Incident investigation is a process for reporting, tracking, and investigating incidents that includes:

- The trending of incident and incident data to identify reoccurring incidents.
- A formal process for investigating incidents including staffing, performing, documenting, and tracking investigations of process safety incidents.
- Managing the resolution and documentation of recommendations generated by the investigations.

Incident investigation is important for the following reasons. It is a way to learn from incidents and communicate lessons learned to internal personnel and other stakeholders. Feedback can apply to the specific incident or a group of incidents sharing similar root causes. Determination of root causes of equipment failures and personnel errors can result in solutions that reduce the frequency and/or consequences of entire categories of incidents. In the United States it is required for facilities covered by the PSM regulation.

The key principles of incident investigation are to:

- Identify potential incidents for investigation.
- Use appropriate techniques to investigate incidents.
- Document incident investigation results.
- Follow through on the results of investigations.
- Trend data to identify repeat incidents that warrant investigation.

Measurement and Metrics

Measurement and metrics establishes performance and efficiency indicators to monitor the near-time effectiveness of RBPS and address which indicators to consider (leading and lagging), how often to collect data, and what to do to ensure effective RBPS. It is important for several reasons. Facilities should monitor the real-time performance managements systems rather than wait for incidents or for infrequent audits to identify management system failures. Performance monitoring allows problems to be identified and corrective actions taken before a serious incident occurs.

The principles of measurement and metrics are to conducts metric acquisition (determine what measurements are needed and collect them) and to use metrics to make corrective action decisions.

Auditing

Auditing is a systematic, independent review to verify conformance with prescribed standards of care. Auditing employs a well-defined review process to ensure consistency and to allow the auditor to reach defensible conclusions. An RBPS management system audit is the systematic review of RBPS management systems to verify suitability and effective, consistent implementation.

Auditing is important because it evaluates RBPS management systems to ensure they are in place and functioning in a manner that protects employees, customers, communities, and physical assets against process safety risk. Audits are an important control mechanism within the overall management of process safety.

The principles of auditing are to conduct the necessary work activities and to use audits to enhance RBPS effectiveness.

Management Review and Continuous Improvement

Management review and continuous improvement include the routine evaluation of whether management systems are performing as intended and producing the desired results as efficiently as possible. It is important because it provides regular checkups on the health of the process safety management systems in order to identify and correct current or incipient deficiencies before they may be revealed by an audit or incident. In other words, if you are management you can expect what you inspect.

The principles of management review and continuous improvement are to conduct review activities and to monitor organizational performance.

Process Safety in Bioprocess Manufacturing Facilities [45]

CCPS defines bioprocess as “A process that makes use of microorganisms, cells in culture, or enzymes to manufacture products or complete a chemical transformation.” Chapters 30, 31 and 32 in this book discuss these processes in some detail. This section discusses the process safety issues associated with these technologies.

Bioprocessing has been used by humans since prehistoric times. Examples include making bread, making cheese, and fermenting alcoholic beverages. Recent advances include the commercialization of recombinant DNA and the production of a variety of protein based therapeutic drugs. Emerging industries and technologies include the production of biofuels from renewable biomass feed stocks such as ethanol biodiesel and polymeric materials. Other emergency technologies include stem cells, gene therapy vectors, and new vaccines.

Bioprocess have many of the same process safety hazards as chemical manufacturing along with other hazards specific to bioprocess. Biohazards can represent extremely low risk (e.g., most recombinant mammalian cell lines used for large scale production of antibody and protein drugs). However in some cases where infectious organisms are used, or where the culture may be susceptible to adventitious contamination (e.g., contamination of human cell lines with a virus), the hazard may be much more significant and the risk to the workers or the public from an accidental release considerably higher. These hazards require the same risk based process safety management systems already discussed.

In addition to the toxic, flammable, or explosive process safety risk that may be present in a bioprocessing facility, risk based process safely management systems must account for biohazard or potentially biohazard materials including the following:

- Biological agent:
  - Pathogenicity
  - Infectious dose
  - Virulence (primary or secondary communicability)
  - Host factors (immunocompetence, pregnancy, underlying medical conditions, extreme age, or immunity
  - Sensitization reactions (allergies, toxins, or biologically active compounds)
  - Incidents of laboratory acquired infections (LAI)
  - Availability of vaccine and/or prophylactic treatment
  - Environmental impact (agent stability—sensitivity to chemical or physical inactivation—survivability and dissemination in the environment)
- Routes/Modes of transmission in the workplace:
  - Respiratory: inhaling of contaminated particles
  - Mucous membrane: splashing, spraying, or droplets in the eyes or mouth
  - Parenteral: penetration through the skin such as cuts, needle sticks, or abrasions
  - Non-intact skin: contact with skin affected with dermatitis, chaffing, hangnails, abrasions, acne, or other conditions that can alter the barrier properties of the skin)
  - Ingestion: swallowing contaminated material
  - Adsorption: adhesion to a surface
- Environmental factors:
  - Climate
  - Geography
  - Proximity to the public
- Procedural and facility factors:
  - Ventilation and laboratory design: directional air, pressure gradients, separation of laboratories from offices, interlocking autoclave and airlock doors
  - Laboratory procedures: use of inherently safer engineered sharps, containment of aerosols, and other means
  - Containment equipment: Class II and III biological safety cabinets, sealed centrifuges, cups and rotors, gasket seals and unbreakable tubes
  - PPE: gloves, safety glasses, lab coats, face masks, respirators or gowns
  - Training: standard microbiological practice, aseptic practices, decontamination, spill cleanup, and handling of accidents
  - Facility sanitation: decontamination, housekeeping, routine cleaning and disinfection, pest and rodent control program
  - Medical surveillance: as dictated by the risk present in the bioprocessing facility

1. For a complete discussion of the topic refer [45].

Regulations[1]

Regulations are a major consideration in the design and operation of chemical facilities. This section provides a description of the significant process requirements. Details of the regulations are available on the Internet or from government agencies, such as the US Department of Labor, or from publications such as those produced by the Thompson Publishing Group and by Primatech, Inc.

Abbreviations used in Government Regulations information:

CFR	Code of Federal Regulations
EPA	Environmental Protection Agency
EPCRA	Emergency Planning and Community Right To Know Act
HAZWHOPER	Hazardous Waste Operations and Response Emergency
HHC	Highly Hazardous Chemicals
MSDS	Material Safety Data Sheet

(continued)

---

[1] This section was prepared with the help of William Carmody, Midland, Michigan. Carmody has had more than 30 years experience in chemical and manufacturing operations for The Dow Chemical Company, Midland, Michigan and six years in Safety and Loss Consulting for Midland Engineering Limited, Midland, Michigan. He has developed entire PSM programs and has conducted many Process Hazard Analyses.

NIOSH	National Institute for Occupational Safety and Health
OSHA	Occupational Safety and Health Administration
PHA	Process Hazard Analysis
PPA	Pollution Prevention Act
PSM	Process Safety Management
RCRA	Resource Conservation and Recovery Act
RMP	Risk Management Plans
SARA	Superfund Amendments and Reauthorization Act
TRI	Toxics Release Inventory

Process Safety Management

On February 24, 1992, the US Department of Labor, Occupational Safety and Health Administration (OSHA) promulgated a final rule, 29 CFR Part 1910.119, "Process Safety Management of Highly Hazardous Chemicals."

OSHA administrates regulations whose objectives are primarily involved with protecting workers. This can be regarded as "inside the fence line." This is a safety issue and is addressed in this section. The rule requires employers to effectively manage the process hazards associated with chemical processes to which the rule applies. OSHA is responsible for the Process Safety Management (PSM) program that is used to prevent or minimize the consequences of catastrophic releases of toxic, reactive, flammable, or explosive chemicals. Standard Number CFR 1910.119 contains requirements for preventing or minimizing the consequences of catastrophic releases of toxic, reactive, flammable, or explosive chemicals. It establishes procedures for PSM that will protect employees by preventing or minimizing the consequences of chemical accidents involving highly hazardous chemicals. The requirements in this standard are intended to eliminate or mitigate the consequences of such releases.

PSM applies to a process involving a chemical at or above the specified threshold quantities listed in 1910.119, Appendix A, and also listed in Table 2.9. The requirements of the rule are also applicable to processes that involve a flammable liquid or gas on-site, in one location, in a quantity of 10,000 lb or more, except for hydrocarbon fuels used solely for workplace consumption as a fuel, or flammable liquids stored in atmospheric pressure tanks.

Process means any activity involving a highly hazardous chemical including any use, storage, manufacturing, handling, or the on-site movement of such chemicals, or combination of these activities. For purposes of this definition, any group of vessels that are interconnected and separate vessels which are located such that a highly hazardous chemical could be involved in potential release shall be considered a single process.

The PSM elements required by 29 CFR Part 1910.119 are briefly described in the following.

Employee participation: Employers must develop a written plan of action for how they will implement employee participation requirements. Employers must consult with employees, affected contractors, and their representatives on the conduct and development of process hazard analyses and on other elements of the standard. They must have access to information developed from the standard, including process hazard analyses.

Process safety information: Employers must compile considerable documented process safety information on the hazards of chemicals used in a covered process as well as information on the process technology and equipment before conducting the process hazard analyses required by the standard.

Process hazard analysis (*PHA*): Employers must perform an analysis to identify, evaluate, and control hazards on processes covered by this standard. The PHA shall be appropriate to the complexity of the process and shall identify, evaluate, and control the hazards involved in the process. The OSHA standard specifies a number of issues that the analysis must address, as well as requirements for who must conduct the analysis, how often it must be performed, and response to its findings. Methodologies that are appropriate include:

- What-if
- Checklists
- What-if/checklist
- HAZOP Study
- Failure Mode and Effects Analysis (FMEA)
- Fault tree analysis

The selection of a PHA methodology or technique will be influenced by many factors including the amount of existing knowledge about the process. All PHA methodologies are subject to certain limitations. The team conducting the PHA needs to understand the methodology that is going to be used. A PHA team can vary in size from two people to a number of people with varied operational and technical backgrounds. Some team members may only be a part of the team for a limited time. The team leader needs to be fully knowledgeable in the proper implementation of the PHA methodology that is to be used and should be impartial in the evaluation. The other full- or part-time team members need to provide the team with expertise in areas such as process technology, process design, operating procedures, and practices.

Standard Number: 1910.119 Appendix A (on the Internet).

This is a listing of toxic and reactive highly hazardous chemicals that present a potential for a catastrophic event at or above the threshold quantity.

Operating procedures: Employers must develop and implement written operating instructions for safely conducting

Table 2.9 List of highly hazardous chemicals, toxics, and reactive chemicals (mandatory)

Chemical name	CAS[a]	Threshold quantity
Acetaldehyde	75–07–0	2,500
Acrolein (2-Popenal)	107–02–8	150
Acrylyl chloride	814–68-6	250
Allyl chloride	107-05-1	1,000
Allylamine	107–11–9	1,000
Alkylaluminum	Varies	5,000
Ammonia, anhydrous	7664–41–7	10,000
Ammonia solutions (greater than 44% ammonia by weight)	7664–41–7	15,000
Ammonium perchlorate	7790–98–9	7,500
Ammonium permanganate	7787–36–2	7,500
Arsine (also called arsenic hydride)	7784–42–1	100
Bis(chloromethyl) ether	542–88–1	100
Boron trichloride	10294–34–5	2,500
Boron trifluoride	7637–07–2	250
Bromine	7726–95–6	1,500
Bromine chloride	13863–41–7	1,500
Bromine pentafluoride	7789–30–2	2,500
Bromine trifluoride	7787–71–5	15,000
3-Bromopropyne (also called propargyl bromide)	106–96–7	100
Butyl hydroperoxide (tertiary)	75–91–2	5,000
Butyl perbenzoate (tertiary)	614–45–9	7,500
Carbonyl chloride (see Phosgene)	75–44–5	100
Carbonyl fluoride	353–50–4	2,500
Cellulose nitrate (concentration greater than 12.6% nitrogen)	9004–70–0	2,500
Chlorine	7782–50–5	1,500
Chlorine dioxide	10049–04–4	1,000
Chlorine pentrafluoride	13637–63–3	1,000
Chlorine trifluoride	7790–91–2	1,000
Chlorodiethylaluminum(also called diethylaluminum chloride)	96–10–6	5,000
1-chloro-2, 4-dinitrobenzene	97–00–7	5,000
Chloromethyl methyl ether	107–30–2	500
Chloropicrin	76–06–2	500
Chloropicrin and methyl Bromide mixture	None	1,500
Chloropicrin and methyl Chloride mixture	None	1,500
Commune hydroperoxide	80–15–9	5,000
Cyanogen	460–19–5	2,500
Cyanogen chloride	506–77–4	500
Cyanuric fluoride	675–14–9	100
Diacetyl peroxide (concentration greater than 70%)	110–22–5	5,000
Diazomethane	334–88–3	500
Dibenzoyl peroxide	94–36–0	7,500
Diborane	19287–45–7	100
Dibutyl peroxide (tertiary)	110–05–04	5,000
Dichloro acetylene	7572–29–4	250
Dichlorosilane	4109–96–0	2,500
Diethylzinc	557–20–0	10,000
Diisopropyl peroxydicarbonate	105–64–6	7,500
Dilauroyl peroxide	105–74–8	7,500
Dimethyldichlorosilane	75–78–5	1,000
Dimethylhydrazine, 1,1-	57–14–7	1,000
Dimethylamine, anhydrous	124–40–3	2,500
2,4-dinitroaniline	97–02–9	5,000
Ethyl methyl ketone peroxide (also methyl ethyl ketone peroxide; concentration greater than 60%)	1338–23–4	5,000

(continued)

Table 2.9 (continued)

Chemical name	CAS[a]	Threshold quantity
Ethyl nitrite	109–95–5	5,000
Ethylamine	75–04–7	7,500
Ethylene fluorohydrin	371–62–0	100
Ethylene oxide	75–21–8	5,000
Ethyleneimine	151–56–4	1,000
Fluorine	7782–41–4	1,000
Formaldehyde (formalin)	50–00–0	1,000
Furan	110–00–9	500
Hexafluoroacetone	684–16–2	5,000
Hydrochloric acid, anhydrous	7647–01–0	5,000
Hydrofluoric acid, anhydrous	7664–39–3	1,000
Hydrogen bromide	10035–10–6	5,000
Hydrogen chloride	7647–01–0	5,000
Hydrogen cyanide, anhydrous	74–90–8	1,000
Hydrogen fluoride	7664–39–3	1,000
Hydrogen peroxide (52% by weight or greater)	7722–84–1	7,500
Hydrogen selenide	7783–07–5	150
Hydrogen sulfide	7783–06–4	1,500
Hydroxylamine	7803–49–8	2,500
Iron, pentacarbonyl	13463–40–6	250
Isopropylamine	75–31–0	5,000
Ketene	463–51–4	100
Methacrylaldehyde	78–85–3	1,000
Methacryloyl chloride	920–46–7	150
Methacryloyloxyethyl isocyanate	30674–80–7	100
Methyl acrylonitrile	126–98–7	250
Methylamine, anhydrous	74–89–5	1,000
Methyl bromide	74–83–9	2,500
Methyl chloride	74–87–3	15,000
Methyl chloroformate	79–22–1	500
Methyl ethyl ketone peroxide (concentration greater than 60%)	1338–23–4	5,000
Methyl fluoroacetate	453–18–9	100
Methyl fluorosulfate	421–20–5	100
Methyl hydrazine	60–34–4	100
Methyl iodide	74–88–4	7,500
Methyl isocyanate	624–83–9	250
Methyl mercaptan	74–93–1	5,000
Methyl vinyl ketone	79–84–4	100
Methyltrichlorosilane	75–79–6	500
Nickel carbonly (nickel tetracarbonyl)	13463–39–3	150
Nitric acid (94.5% by weight or greater)	7697–37–2	500
Nitric oxide	10102–43–9	250
Nitroaniline (para)		
Nitroaniline	100–01–6	5,000
Nitromethane	75–52–5	2,500
Nitrogen dioxide	10102–44–0	250
Nitrogen oxides (NO; NO(2); N_2O_4; N_2O_3)	10102–44–0	250
Nitrogen tetroxide (also called nitrogen peroxide)	10544–72–6	250
Nitrogen trifluoride	7783–54–2	5,000
Nitrogen trioxide	10544–73–7	250
Oleum (65–80% by weight; also called fuming sulfuric acid)	8014–94–7	1,000
Osmium tetroxide	20816–12–0	100
Oxygen difluoride (fluorine monoxide)	7783–41–7	100

(continued)

Table 2.9 (continued)

Chemical name	CAS[a]	Threshold quantity
Ozone	10028–15–6	100
Pentaborane	19624–22–7	100
Peracetic acid (concentration greater 60% acetic acid; also called peroxyacetic acid)	79–21–0	1,000
Perchloric acid (concentration greater than 60% by weight)	7601–90–3	5,000
Perchloromethyl mercaptan	594–42–3	150
Perchloryl fluoride	7616–94–6	5,000
Peroxyacetic acid (concentration greater than 60% acetic acid; also called peracetic acid)	79–21–0	1,000
Phosgene (also called carbonyl chloride)	75–44–5	100
Phosphine (Hydrogen phosphide)	7803–51–2	100
Phosphorus oxychloride (also called phosphoryl chloride)	10025–87–3	1,000
Phosphorus trichloride	7719–12–2	1,000
Phosphoryl chloride (also called phosphorus oxychloride)	10025–87–3	1,000
Propargyl bromide	106–96–7	100
Propyl nitrate	627–3–	2,500
Sarin	107–44–8	100
Selenium hexafluoride	7783–79–1	1,000
Stibine (antimony hydride)	7803–52–3	500
Sulfur dioxide (liquid)	7446–09–5	1,000
Sulfur pentafluoride	5714–22–7	250
Sulfur tetrafluoride	7783–60–0	250
Sulfur trioxide (also called sulfuric anhydride)	7446–11–9	1,000
Sulfuric anhydride (also called sulfur trioxide)	7446–11–9	1,000
Tellurium hexafluoride	7783–80–4	250
Tetrafluoroethylene	116–14–3	5,000
Tetrafluorohydrazine	10036–47–2	5,000
Tetramethyl lead	75–74–1	1,000
Thionyl chloride	7719–09–7	250
Trichloro (chloromethyl) silane	1558–25–	100
Trichloro (dichlorophenyl) silane	27137–85–5	2,500
Trichlorosilane	10025–78–2	5,000
Trifluorochloroethylene	79–38–9	10,000
Trimethyoxysilane	2487–90–3	1,500

[a]Chemical abstract service number

[b]Threshold quality in pounds (amount necessary to be covered by this standard)

activities involved in each covered process consistent with the process safety information. The written procedures must address steps for each operating phase, operating limits, safety and health considerations, and safety systems and their functions. Included must be normal operation, startup, shutdown, emergency operations, and other operating parameters.

Training: The proposal requires training for employees involved in covered processes. Initial training requires all employees currently involved in each process, and all employees newly assigned, be trained in an overview of the process and its operating procedures. Refresher training shall be provided at least every 3 years, and more often if necessary, to each employee involved in the process. After training, employees must ascertain that workers have received and understood the training.

Contractors: Employers must inform contract employees prior to the initiation of the contractor's work of the known potential fire, explosion, or toxic release hazards related to the contractor's work and the process. Contract employees and host employers must ensure that contract workers are trained in the work practices necessary to perform their jobs safely and are informed of any applicable safety rules of the facility work and the process.

Pre-startup safety review: Employers must perform a pre-startup safety review for new facilities and for modified facilities when the modification is significant enough to require a change in the process safety information. The safety review shall confirm that prior to the introduction of highly hazardous chemicals to a process:

1. Construction and equipment is in accordance with design specifications.

2. Safety, operating, maintenance, and emergency procedures are in place and are adequate.
3. For new facilities, a PHA has been performed and recommendations have been resolved or implemented before startup.
4. Modified facilities meet the requirements contained in management of change.

Mechanical integrity: Employers must ensure the initial and on-going integrity of process equipment by determining that the equipment is designed, installed, and maintained properly. The standard requires testing and inspection of equipment, quality assurance checks of equipment, spare parts and maintenance materials, and correction of deficiencies. The following process equipment is targeted in this proposal: pressure vessels and storage tanks; piping systems (including valves); relief and vent systems and devices; emergency shutdown systems; controls, and pumps.

Hot work permit: Employers must have a hot work program in place and issue a permit for all hot work operations conducted on or near a covered process.

Management of change: Employers must establish and implement written procedures to manage changes (except for "replacements in kind") to process chemicals, technology, equipment, and procedures; and, changes to facilities that affect a covered process. Employees involved in operating a process and maintenance and contract employees whose tasks will be affected by a change in the process shall be informed of, and trained in, the change prior to startup of the process or affected part of the process. The procedures shall ensure that the necessary time period for the change and authorization requirements for the proposed change are addressed.

Incident investigation: Employers must investigate each incident that resulted in, or could reasonably have resulted in a catastrophic release of highly hazardous chemical in the workplace. An incident investigation shall be initiated as promptly as possible, but not later than 48 h following the incident. A report shall be prepared at the conclusion of the investigation.

Although not stressed by the regulations, the objective of the incident investigation should be the development and implementation of recommendations to ensure the incident is not repeated. This objective should apply not only to the process involved, but also to all similar situations having the same potential. In major incidents, the Chemical Safety Board's investigation of reports serves as a vehicle to communicate to a much broader audience than the organizations that had the incident.

Emergency planning and response: Employers must establish and implement an emergency action plan for the entire plant in accordance with the provisions of OSHA's emergency action plan to meet the minimum requirements for emergency planning. This is the only element of PSM that must be carried out beyond the boundaries of a covered process.

Compliance audits: Employers must certify that they have evaluated compliance with the provisions of this section at least every 3 years to verify that procedures and practices developed under the standard are adequate and are being followed. The compliance audit shall be conducted by at least one person knowledgeable in the process. The employer shall determine and document an appropriate response to each of the findings of the compliance audit, and document that deficiencies have been addressed.

Trade secrets: Employers must make all information necessary to comply with the requirements of this section available to those persons responsible for compiling the process safety information, developing process hazard analyses, developing the operating procedures, those involved in incident investigations, emergency planning, and response and compliance audits without regard to possible trade secret status of such information. Nothing in this paragraph shall preclude the employer from requiring the persons to whom the information is made available to enter into confidentiality agreements not to disclose the information.

The above elements outline the programs required by PSM. These programs are performance-type standards. They spell out programs and choices and are not limited to specific details. These elements have served to organize and guide the process safety programs of all who are covered by it. They have served to bring direction to training and publications involving process safety. The AIChE's Center for Chemical Process Safety has publications and training programs to support most of these elements.

Risk Management Plans

The EPA is charged primarily with the responsibility to protect the public and the environment. One could regard this as "outside the fence line." Risk management plans (RMPs) are required by the Environmental Protection Agency (EPA). Since protecting the public and the environment is mainly an environmental issue rather than a safety issue, this subject will be covered only briefly in this section.

Congress enacted Section 112(r) of the 1990 Clean Air Act (CAA) to address the threat of catastrophic releases of chemicals that might cause immediate deaths or injuries in communities. It requires owners and operators of covered facilities to submit RMPs to the EPA. The final RMP rule was published in 40 CFR 68 in the Federal Register on June 20, 1996. RMPs must summarize the potential threat of sudden, large releases of certain dangerous chemicals and facilities' plans to prevent such releases and mitigate any damage.

Operators of facilities that are subject to the EPA's RMP must perform offsite consequence analyses to determine whether accidental releases from their processes could put nearby populations at risk. In performing a consequence analysis it is assumed that all or part of a hazardous substance escapes from a process at a given facility. It is then estimated how far downwind hazardous gas concentrations may extend.

Facilities that must prepare and submit RMPs must estimate the offsite consequences of accidental releases. This can be done using tables (such as those provided in CAA 112 (r) Offsite Consequence Analysis) or a computerized model. There are a number of commercially available computer models. Submitters are expected to choose a tool that is appropriate for their facility.

The owners and operators of stationary sources producing, processing, handling, or storing of extremely hazardous substances have a general duty to identify hazards that may result from an accidental release This includes agents that may or may not be identified by any government agency which may cause death, injury, or property damage. In other words, just because a substance is not listed is not an excuse to fail to consider its hazards.

This section with its emphasis on Process Safety does not cover the considerable other safety, design, and operating requirements of other chemical-related regulations. Many of these requirements also include national codes as guidelines or as adopted regulations. Examples of these are in the American Society of Mechanical Engineers (ASME) *2001 Boiler Pressure Vessel Code*, the National Fire Protection Association (NFPA) which covers a wide range of fire safety issues and the American Petroleum Institute (API) *Recommended Practice 520, Sizing, Selection, and Installation of Pressure Relieving Devices in Refineries.*

An extremely hazardous substance is any agent that may or may not be listed by any government agency which, as the result of short-term exposures associated with releases to the air, cause death, injury, or property damage due to its toxicity, reactivity, flammability, volatility, or corrosivity.

Toxics Release Inventory

Two statutes, the Emergency Planning and Community Right-to-Know Act (EPCRA) and Section 6607 of the Pollution Prevention Act (PPA), mandate that a publicly accessible toxic chemical database be developed and maintained by the US EPA. This database, known as the Toxics Release Inventory (TRI), contains information concerning waste management activities and the release of toxic chemicals by facilities that manufacture, process, or otherwise use these materials. The TRI of 1999 is a publicly available database containing information on toxic chemical releases and other waste management activities that are reported annually by manufacturing facilities and facilities in certain other industry sectors, as well as by federal facilities. The TRI program is now under the EPA's Office of Environmental Information. This inventory was established under the EPCRA of 1986 which was enacted to promote emergency planning, to minimize the effects of chemical accidents, and to provide the public. As of November 2001, there were 667 toxic chemicals and chemical compounds on the list.

Hazardous Waste Operations and Emergency Response Standard

The Hazardous Waste Operations and Emergency Response (HAZWOPER) standard, 29 CFR Part 1910.120, applies to five distinct groups of employers and their employees. This includes any employees who are exposed or potentially exposed to hazardous substances—including hazardous waste—and who are engaged in one of the following operations as specified by 1910.120:

- Clean-up operations
- Corrective actions
- Voluntary clean-up operations
- Operations involving hazardous wastes
- Emergency response operations for releases of, or substantial threats of release of, hazardous substances regard less of the location of the hazard.

In addition, with the passage of the Pollution Prevention Act (PPA) in 1991, facilities must report other waste management amounts including the quantities of TRI chemicals recycled, combusted for energy recovery, and treated on- and offsite.

More Information

For more information on Regulations, the books, magazine articles, and Internet references in the reference section can be very helpful. Following the requirements of the many aspects of Regulations can be quite complicated and involve a lot of detail. There is a considerable amount of good assistance available which help can make the subject manageable.

The Principal Reason for Most Chemical Process Accidents

Ask any group of people experienced in chemical plant operations what causes most chemical process accidents, and you will get a variety of answers including: operator error, equipment failure, poor design, act of God, and bad luck. However, in the opinion of representatives of many of the large chemical and oil companies in the United States, these answers are generally incorrect. The Center for Chemical Process Safety, an organization sponsored by the American Institute of Chemical Engineers, includes representatives of many of the largest chemical and oil companies in the United States and the world, and states that "It is an axiom that process safety incidents are the result of *management system failure*." Invariably, some aspect of a process safety management system can be found that, had it functioned properly, could have prevented an incident (or reduced the seriousness of it). "It is a rare situation where an "Act of God" or other uncontrollable event is the sole cause of an incident. Much more common is the situation where an incident is the result of multiple causes, including management system failures. Therefore, it is more appropriate to presume that management system failures underlie every incident so that we may act to uncover such failures and then modify the appropriate management systems, rather than presume that if an "Act of God" appears to be the immediate cause, investigation should cease because there is nothing that can be done to prevent such future incidents" [17].

For example, consider a case where a small amount of hazardous material is spilled while a sample is being taken from a process line. It is not enough to look into the situation and conclude that this is an example of an operator error where procedures were not followed, and then simply to recommend that the employee be instructed to follow procedures in the future. Further investigation may reveal deficiencies in the training system or in the equipment. Still more investigation may reveal deficiencies in the management system that plans resources for training or that provides for proper equipment for sampling. It then may be appropriate to change the management system to prevent repetition of the incident.

Levels of Causes

There are several levels of causes of accidents, usually (1) the immediate cause, (2) contributing causes to the accident or to the severity of the accident, and (3) the "root cause." The root cause is what really caused the accident, and when this is determined, it may be possibly to *prevent* future similar accidents. With the 20–20 hindsight that is available after an accident, the root cause usually can be found. The purpose of the discussion in the next section is to illustrate how knowledge about the root causes of some important accidents can help to keep them from happening again. It will be noted that the root cause is rarely the fault of one person, but instead is the result of a management system that does not function properly.

Following are brief analyses of several case histories that have been of landmark importance in the industrial world, and that have affected the chemical industry all over the world.

Case Histories

Flixborough, England 1974 [19]

On June 1, 1974, an accident occurred in the Nypro plant in Flixborough, England, in a process where cyclohexane was oxidized to cyclohexanone for the manufacture of caprolactam, the basic raw material for the production of Nylon 6. The process consisted of six reactors in series at 155°C and 8.8 bar (130 psig) containing a total of 120 t of cyclohexane and a small amount of cyclohexanone. The final reactor in the process contained 94% cyclohexane. There was a massive leak followed by a large UVCE and fire that killed 26 people, injured 36 people, destroyed 1,821 houses, and damaged 167 shops. It was estimated that 30 t of cyclohexane was involved in the explosion. The accident occurred on Saturday; on a working day, casualties would have been much higher.

The accident happened when the plant had to replace one of six reactors and rushed to refit the plant to bypass the disabled reactor. Scaffolding was jerry-rigged to support a 20-in. pipe connecting reactor four with reactor six, which violated industry and the manufacturer's recommendations. The reactor that failed showed stress crack corrosion. The only drawings for the repair were in chalk on the machine shop floor. Both ends of the 20-in. pipe had expansion joints where they attached to the reactors. The pipe was supported on scaffolding-type supports and was offset with a "dog-leg" to fit the reactors, which were at different levels to promote gravity flow. The safety reviews, if any, were insufficient.

Immediate cause: A pipe replacing a failed reactor failed, releasing large quantities of hot cyclohexane forming a vapor cloud that ignited.

Contributing causes to the accident and the severity of the accident:

1. The reactor failed without an adequate check on why (metallurgical failure).

2. The pipe was connected without an adequate check on its strength and on inadequate supports.
3. Expansion joints (bellows) were used on each end of pipe in a "dog-leg" without adequate support, contrary to the recommendations of the manufacturer.
4. There was a large inventory of hot cyclohexane under pressure.
5. The accident occurred during startup.
6. The control room was not built with adequate strength, and was poorly sited.
7. The previous works engineer had left and had not been replaced. According to the Flixborough Report, "There was no mechanical engineer on site of sufficient qualification, status or authority to deal with complex and novel engineering problems and insist on necessary measures being taken."
8. The plant did not have a sufficient complement of experienced people, and individuals tended to be overworked and liable to error.

Root cause: Management systems deficiencies resulted in:

1. A lack of experienced and qualified people
2. Inadequate procedures involving plant modifications
3. Regulations on pressure vessels that dealt mainly with steam and air and did not adequately address hazardous materials
4. A process with a very large amount of hot hydrocarbons under pressure and well above its flash point installed in an area that could expose many people to a severe hazard

This accident resulted in significant changes in England and the rest of the world in the manner in which chemical process safety is managed by industry and government. One of the conclusions reached as a result of this accident, which has had a wide effect in the chemical industry, is that "limitations of inventory (or flammable materials) should be taken as specific design objectives in major hazard installations."

The use of expansion joints (bellows, in this case) which were improperly installed may have been a principal reason for the accident. This provides additional reasons not to use expansion joints (except in special exceptional circumstances).

Bhopal, 1985 (C&EN Feb 11, 1985; Technica 1989 [41])

On December 3 and 4, 1985, a chemical release causing a massive toxic gas cloud occurred at the Union Carbide India, Ltd, plant in Bhopal, India. (Union Carbide is now a part of The Dow Chemical Company.) The process involved used methyl isocyanate (MIC), an extremely toxic chemical, to make Sevin, a pesticide. According to various authoritative reports, about 1,700–2,700 (possibly more) people were killed, 50,000 people were affected seriously, and 1,000,000 people were affected in some way. The final settlement may involve billions of dollars. It was one of the worst industrial accidents in history. The accident occurred when about 120–240 gal of water were allowed to contaminate an MIC storage tank. The MIC hydrolyzed, causing heat and pressure, which in turn caused the tank rupture disk to burst.

Equipment designed to handle an MIC release included a recirculating caustic soda scrubber tower and a flare system designed for 10,000 lb/h, which would be moderate flows from process vents. It was not designed to handle runaway reactions from storage. The design was based on the assumption that full cooling would be provided by the refrigeration system. The actual release was estimated to be 27,000 lb over 2 h, with the tank at 43°C. At the time of the release the refrigeration had been turned off. The flare tower was shut down for repairs. A system of pressurized sprinklers that was supposed to form a water curtain over the escaping gases was deficient, in that water pressure was too low for water to reach the height of the escaping gas.

There have been conflicting stories of how the water got into the tank, including operator error, contamination, and sabotage.

Immediate cause: The immediate cause was hydrolysis of MIC due to water contamination. The exact source of the water has not been determined.

Contributing causes

1. Flare tower was shut down for repair.
2. Scrubber was inadequate to handle a large release.
3. Chilling system was turned off. (It also was too small.)
4. MIC tank was not equipped with adequate instrumentation.
5. Operating personnel lacked knowledge and training.
6. The inventory of MIC was large.
7. There was a lack of automatic devices and warning systems; it has been reported that safety systems had to be turned on manually.
8. When the plant was built, over 20 years before the accident, there were very few people near it. At the time of the accident, a shanty town had grown up near the plant with a density of 100 people per acre, greatly increasing the potential exposure of people to toxic releases. There was no emergency action plan to notify neighbors of the potential for toxic releases or of what to do if there was a release, nor was there a functioning alarm system.

Root cause: The root cause of the accident appears to be a management system that did not adequately respond to the potential hazards of MIC. There was probably a greater inventory of MIC than was needed. The main process expertise was in the United States. Local management does not appear to have understood the process or the consequences of changes made. This includes plant design, maintenance and operations, backup systems, and community responsibility. (Union Carbide has provided legal arguments alleging that sabotage caused the release; there appears to be enough blame to go around for all those involved in any way in the plant, including government units.)

This accident has become widely known. It is an objective of many chemical process safety programs and government actions to "avoid another Bhopal"—that is, to avoid a severe release of toxic chemicals (usually referring to toxic chemicals in the air). Almost every chemical company in the world has been affected by this incident in one way or another, in the design and operation of chemical plants, in community action programs, and in the activities of such organizations as the American Institute of Chemical Engineers, the Chemical Manufacturers Association (now the American Chemistry Council), and many governmental units.

Phillips Explosion, 1989 [37]

On October 23, 1989, at approximately 1,300, an explosion and fire ripped through the Phillips 66 Company's Houston Chemical Complex in Pasadena, Texas. At the site, 23 workers were killed, and more than 130 were injured. Property damage was nearly $750 million. Business interruption cost is not available but is probably a very large figure.

The release occurred during maintenance operations on a polyethylene reactor. Two of the six workers on the maintenance crews in the immediate vicinity of the reactor leg where the release occurred were killed, together with 21 other employees of the facility. Debris from the plant was found 6 miles from the explosion site. Structural steel beams were twisted like pretzels by the extreme heat. Two polyethylene production plants covering an area of 16 acres were completely destroyed.

The Phillips complex produces high-density polyethylene, which is used to make milk bottles and other containers. Prior to the accident, the facility produced approximately 1.5 billion pounds of the material per year. It employed 905 company employees and approximately 600 daily contract employees. The contract employees were engaged primarily in regular maintenance activities and new plant construction.

The accident resulted from a release of extremely flammable process gases that occurred during regular maintenance operations on one of the plant's polyethylene reactors. It is estimated that within 90–120 s more than 85,000 lb of flammable gases were released through an open valve. A huge flammable vapor cloud was formed that came into contact with an ignition source and exploded with the energy of 4,800 lb of TNT. The initial explosion was equivalent to an earthquake with a magnitude of 3.5 on the Richter scale. A second explosion occurred 10–15 min later when two isobutane tanks exploded. Each explosion damaged other units, creating a chain reaction of explosions. One witness reported hearing ten separate explosions over a 2-h period.

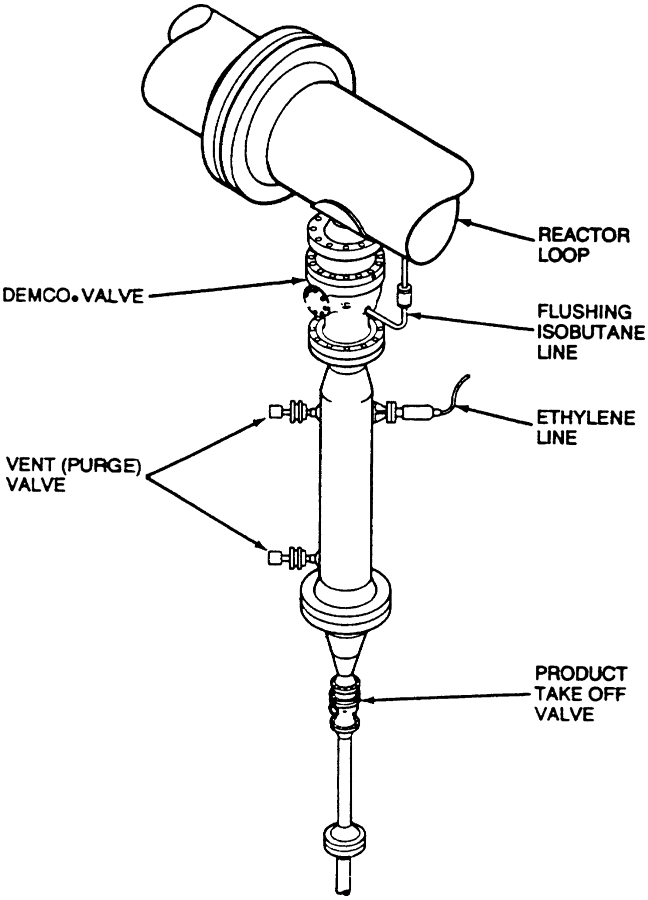


Fig. 2.16 Typical piping settling leg arrangement

In the process used by Phillips at this site to produce high-density polyethylene, ethylene gas is dissolved in isobutane and, with various other chemicals added, is reacted in long pipes under elevated pressure and temperature. The dissolved ethylene reacts with itself to form polyethylene particles that gradually come to rest in settling legs, where they are eventually removed through valves at the bottom. At the top of the legs there is a single ball valve (DEMCO® brand) where the legs join with other reactor pipes. The DEMCO valve is kept open during production so that the polyethylene particles can settle into the leg. A typical piping settling leg arrangement is shown in Fig. 2.16.

In the Phillips reactor, the plastic material frequently clogged the settling legs. When this happened, the DEMCO valve for the blocked leg was closed, the leg disassembled, and the block removed. During this particular maintenance process, the reactor settling leg was disassembled and the block of polymer removed. While this maintenance process was going on, the reaction continued, and the product settled in the legs that remained in place. If the DEMCO valve were to open during a cleaning-out operation, there would be nothing to prevent the escape of the gas to the atmosphere.

After the explosion it was found that the DEMCO valve was open at the time of the release. The leg to be cleaned had been prepared by a Phillips operator. The air hoses that operated the DEMCO valve were improperly connected in a reversed position such that a closed DEMCO valve would be opened when the actuator was in the closed position. In addition, the following unsafe conditions existed:

1. The DEMCO valve did not have its lock out device in place.
2. The hoses supplied to the valve actuator mechanism could be connected at any time even though the Phillips operating procedure stipulated that the hoses should never be connected during maintenance.
3. The air hoses connecting the open and closed sides of the valve were identical, thus allowing the hoses to be cross-connected and permitting the valve to be opened when the operator intended to close it.
4. The air supply valves for the actuator mechanism air hoses were in the open position so that air would flow and cause the actuator to rotate the DEMCO valve when the hoses were connected.
5. The DEMCO valve was capable of being physically locked in the open position as well as in the closed position. The valve lockout system was inadequate to prevent someone from inadvertently opening the DEMCO valve during a maintenance procedure.

Established Phillips corporate safety procedures and standard industry practice require back-up protection in the form of a double valve or blind flange insert whenever a process or chemical line in hydrocarbon service is opened. According to OSHA, Phillips had implemented a special procedure for this maintenance operation that did not incorporate the required backup. Consequently, none was used on October 23.

The consequences of the accident were exacerbated by the lack of a water system dedicated to fire fighting, and by deficiencies in the shared system. When the process water system was extensively damaged by the explo-sion, the plant's water supply for fighting fires was also disrupted. The water pressure was inadequate for fire fighting. The force of the explosion ruptured water-lines and adjacent vessels containing flammable and combustible materials. The ruptured water lines could not be isolated to restore water pressure because the valves to do so were engulfed in flames. Of the three backup diesel pumps, one had been taken out of service and was unavailable, and another soon ran out of fuel. It was necessary to lay hoses to remote sites—settling ponds, a cooling tower, a water treatment plant, and so on. Electric cables supplying power to regular service pumps were damaged by fire, and those pumps were rendered inoperable. Even so, the fire was brought under control within 10 h.

In the months preceding the explosion, according to testimony, there had been several small fires, and the alarm had sounded as many as four or five times a day. There had been a fatality at the same plant doing a similar operation about 3 months before this incident. Some of the employees in the area where the release occurred may not have heard the siren because of the ambient noise level, and may not have known of the impending disaster. Employees in the immediate area of the release began running as soon as they realized the gas was escaping.

The large number of fatalities was due in part to the inadequate separation between buildings in the complex. The site layout and the proximity of normally high-occupancy structures, such as the control and finishing building, to large-capacity reactors and hydrocarbon vessels contributed to the severity of the event.

The distances between process equipment were in violation of accepted engineering practices and did not allow personnel to leave the polyethylene plants safely during the initial vapor release; nor was there sufficient separation between reactors and the control room to carry out emergency shutdown procedures. The control room, in fact, was destroyed by the initial explosion. Of the 22 victims' bodies recovered at the scene, all were located within 250 ft of the vapor release point.

OSHA's investigation revealed that a number of company audits had identified unsafe conditions but largely had been ignored. Thus, a citation for willful violations of the OSHA "general duty" clause was issued to Phillips with proposed penalties of $5,660,000. In addition, proposed penalties of $6,200 were issued for other serious violations. A citation for willful violations with proposed penalties of $724,000 was issued to Fish Engineering and Construction, a Phillips maintenance contractor. Other financial penalties have been proposed. In the investigation it became apparent that Fish had become accustomed to tolerating safety and health violations at the site by its personnel and Phillips personnel, as well as participating in those violations by knowing about them and not taking direct positive action to protect its employees.

Since 1972, OSHA has conducted 92 inspections in the Dallas region at various Phillips locations; 24 were in response to a fatality or a serious accident. OSHA determined that Phillips had not acted upon reports by its own safety personnel and outside consultants who had pointed out unsafe conditions. OSHA also had conducted 44 inspections of the Fish Company, seven of them in response to a fatality or a serious accident.

One of the major findings by OSHA was that Phillips had not conducted a PHA or equivalent (such as HAZOP) in its polyethylene plants.

Immediate cause: There was a release of flammable process gases during the unplugging of Number 4 Reactor Leg on Reactor 6 while undergoing a regular maintenance procedure by contractor employees. The unconfined flammable vapor cloud was ignited and exploded with devastating results.

The immediate cause of the leak was that a process valve was opened by mistake while the line was open. The valve was open to the atmosphere without a second line of defense such as another valve or a blind flange.

Contributing causes to the accident and the severity of the accident.

1. Procedures to require backup protection in the form of a double valve or a blind flange insert were not used. The lockout system was inadequate.
2. Air hoses were improperly connected in the reversed position.
3. The air hoses for the open and closed side of the valve were identical, allowing the hoses to be cross-connected.
4. The DEMCO valve actuator mechanism did not have its lockout device in place.
5. There was not a water system dedicated to fire fighting, and there were deficiencies in the system shared with the process.
6. The site layout and proximity of high-occupancy structures contributed to the severity.
7. There was inadequate separation of buildings within the complex. Especially, there was inadequate spacing between the reactors and the control room.

Root causes: The root causes of the accident and its extreme severity appear to be failures of the management system, as shown by the following: [37]

1. According to OSHA, Phillips had not conducted a PHA or equivalent (such as HAZOP) in its polyethylene plants.
2. It was found by OSHA that the contractor, Fish Engineering, had a history of serious and willful violations of safety standard, which Phillips had not acted upon. The same contractor also had been involved in a fatal accident at the same facility 3 months earlier.
3. A report by OSHA stated that Phillips had not acted upon reports issued previously by the company's own safety personnel and outside consultants. Phillips had numerous citations from OSHA since July 1972. OSHA discovered internal Phillips documents that called for corrective action but which were largely ignored.
4. Safe operating procedures were not required for opening lines in hazardous service.
5. An effective safety permit system was not enforced with Phillips or contractor employees, especially line opening and hot work permits.
6. Buildings containing personnel were not separated from process units in accordance with accepted engineering principles, or designed with enough resistance to fire and explosion.
7. The fire protection system was not maintained in a state of readiness:
    (a) One of the three diesel-powered water pumps had been taken out of service.
    (b) Another of the three diesel-powered water pumps was not fully fueled, and it ran out of fuel during the fire fighting.
    (c) Electric cables supplying power to regular service fire pumps were not located underground and were exposed to blast and fire damage.

Summary

As the tragic case histories illustrate, the importance of the risk based process safety practices presented in this chapter to prevent and mitigate potential catastrophic process safety incidents in the future becomes alarmingly apparent, and the necessity for risk based process design is clear. The case histories also reveal significant flaws in the management systems necessary to ensure that good process safety practices are followed. Even a process designed according to principals of process safety can be transformed into one with a high potential for disaster if risk based process safety management systems are not in place to ensure that good process safety practices are followed throughout the life cycle of the process.

Because there is always risk when equipment, instrumentation, and human activity are involved, there is no way to make a plant completely safe. However, facilities can be made risk tolerant by careful examination of all aspects of design and management, using modern techniques that are now available. If the process safety performance and public image of the chemical processing industries is to improve, risk based process design coupled with risk based process management is imperative.

In addition to the information presented in this chapter and in the publications it has cited, references listed below are recommended as appropriate source material.

Acknowledgment The author wishes to acknowledge that earlier versions of this chapter, which appeared in the eleventh and preceding editions of the Handbook, were authored by Stan Englund of Midland, Michigan, and that significant content from the previous editions is included herein.

References

1. M & M Protection Consultants (1989) One hundred largest losses: a thirty year review of property damage in the hydrocarbon-chemical industry, 12th edn. M & M Protection Consultants, Riverside Plaza
2. CCPS (2010) Process safety boot camp (CCPS Training Course). Center for Chemical Process Safety, American Institute of Chemical Engineers, New York
3. CCPS (2008) Guidelines for hazard evaluation procedures. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York
4. CCPS (2008) Guidelines for safe and reliable instrumented protective system. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York
5. Grinwis D (1986) Process Systems Associate; Larsen, Paul, Associate Instrument Engineering Consultant; and Schrock, Luther, Process Consultant, The Dow Chemical Co., Midland, MI and Freeport, personal-communication, Oct 1986
6. Fisher H (1985) Chem Eng Prog 3
7. Huff JE (1977) A general approach to the sizing of emergency pressure relief systems. Reprints of International symposium on loss prevention and safety promotion in the process industries, Heidelberg. DECHEMA, Frankfurt, p. IV 223
8. Huff JE (1984) Institute of chemical engineers symposium series, No. 85, p. 109
9. Huff JE (1984) Emergency venting requirements for gassy reactions from closed system tests. Plant/Operations Progress 3 (1):50–59
10. Huff JE (1987) The role of pressure relief in reactive chemical safety. In: International symposium on preventing major chemical accidents, Washington, DC (Sponsored by the Center for Chemical Process Safety of the American Institute of Chemical Engineers, The United States Environmental Protection Agency, and the World Bank), Feb 3, 1987
11. Fauske, Hans K, Leung J (1985) New experimental technique for characterizing runaway chemical reactions. Chem Eng Progr 39
12. CCPS (2010) Guidelines for process safety metrics. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York
13. CCPS (2007) Guidelines for risk based process safety. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York
14. CCPS (2009) Inherently safer chemical processes, a life cycle approach. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York
15. Kletz TA (1991) Plant design for safety—a user-friendly approach. Hemisphere, New York
16. Zabetakis MG (1965) Flammability characteristics of combustible gases and vapors. In: Bulletin 627. U.S. Dept. of the Interior, Bureau of Mines, Washington, DC
17. Bodurtha FT (1987) Industrial explosion control course. Center for Professional Advancement, Chicago, Sept 14–16, 1987
18. Baker WE, Cox PA, Westine PS, Kulesz JJ, Strehlow RA (1983) Explosion hazards and evaluation. Elsevier, New York
19. Lees FP (1980) Loss prevention in the process industries. Butterworths, London
20. Ludwig EE (1977) Applied process design for chemical and petrochemical plants, vol 1, 2nd edn. Gulf, Houston
21. Eichel FG (1967) Electrostatics. Chem Eng 154–167
22. TNO (1979) Methods for the calculation of the physical effects of the escape of dangerous materials: liquids and gases. In: The yellow book. TNO, Apeldoorn
23. Brasie WC (1983) Michigan Division, Process Engineering, The Dow Chemical Co., Midland, MI, personal communication, March 9, 1983
24. Kletz T (1985) Cheaper, safer plants or wealth and safety at work. Institution of Chemical Engineers, Rugby
25. Cloud MJ (1990) Fire, the most tolerable third party. Mich Nat Res 18
26. FPA (1988) Fire safety data. Fire Protection Association, London
27. Cawse JN, Pesetsky B, Vyn WT (n.d.) The liquid phase decomposition of ethylene oxide. Union Carbide Corporation., Technical Center, North Charleston
28. Stull DR (1976) Fundamentals of fire and explosion, Corporate Safety and Loss Dept., The Dow Chemical Co., Midland. American Institute of Chemical Engineers, New York, p 50
29. Kohlbrand (1990) Case history of a deflagration involving an organic solvent/oxygen system below its flash point. In: 24th Annual loss prevention symposium. Sponsored by the American Institute of Chemical Engineers, San Diego, Aug 19–22, 1990
30. Prugh RW (1988) Quantitative evaluation of BLEVE hazards. In: AICHE loss prevention symposium, Paper No. 74e, AICHE Spring National Meeting, New Orleans, March 6–10, 1988
31. Bartknecht W (1981) Explosions course prevention protection. Springer, Berlin, p 108
32. NFPA 77 (1986) National Fire Protection Association, Batterymarch Park, Quincy
33. NFPA 68, National Fire Protection Association, Batterymarch Park, Quincy, MA, 1998.
34. CCPS (1989) CPQRA (chemical process quantitative risk analysis) Tables 2.12, 2.13. Center for Chemical Process Safety, American Institute of Chemical Engineers, p. 161, 165
35. Walters S (1984) The beginnings. Mech Eng 4:38–46
36. Burk, Art., Principal Safety Consultant. Du Pont, Newark, DE, personal communication (Feb. 20, 1990).
37. OSHA (Occupational Safety and Health Administration) (1990) U. S. Department of Labor, The Phillips Company Houston Chemical Complex explosion and fire, Apr 1990
38. Clancey VJ (1972) Diagnostic features of explosion damage. In: Sixth Int. Meeting of Forensic Sciences, Edinburgh
39. Brasie WC (1982) Michigan Division, Process Engineering, The Dow Chemical Co., Midland, MI, personal communication, Oct 6, 1982
40. NUS Corp. (1989) HAZOP Study Team training manual. Predictive hazard identification techniques for Dow Corning facilities. NUS Corp., Gaithersburg
41. Technica, Inc. (1989) HAZOP leaders course. Technica, Inc., Columbus, OH, Nov 6–10 (course leaders David Slater and Frederick Dyke)
42. CCPS (1988) Guidelines for safe storage and handling of high toxic hazard materials. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York
43. API RP 521 (1982) 2nd ed, American Petroleum Institute, Washington, DC
44. CCPS (2009) Guidelines for developing quantitative safety risk criteria. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York
45. CCPS (2011) Guidelines for process safety in bioprocess manufacturing facilities. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York
46. CCPS (1993) Guidelines for safe automation of chemical processes. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York
47. Mackenzie J (1990) Hydrogen peroxide without accidents. Chem Eng 84ff

Additional Reading References

American Petroleum Institute, *API Recommended Practice 520, Sizing, Selection, and Installation of Pressure Relieving Devices in Refineries, Part I, Sizing and Selection*, American Petroleum Institute, 1995–2000

American Society of Mechanical Engineers (2001) *2001 Boiler Pressure Vessel Code*, ASME International, Fairfield, NJ

Bartknecht W (1993) Explosions course: prevention, protection. Springer, Berlin

Bretherick L (1995) Handbook of reactive chemical hazards, 5th edn. Butterworths, London

Crowl D, Bollinger R (1997) Inherently safer chemical processes: a life cycle approach. Center for Chemical Process Safety (CCPS), American Institute of Chemical Engineers, New York

DIERS (Design Institute for Emergency Relief Systems), American Institute of Chemical Engineers, 3 Park Ave, New York. http://www.diers.net/

Englund SM (1991) Design and operate plants for inherent safety. *Chemical Engineering Progress*, 85–91 (Part 1); 79–86 (Part 2)

Englund SM (1993) Process and design options for inherently safer plants. In: Fthenakis VM (ed) Prevention and control of accidental releases of hazardous gases. Van Nostrand Reinhold, New York

Englund SM (1997) Chemical process safety. In: Green DW, Green DW (eds) Perry's chemical engineers' handbook, 7th edn. McGraw-Hill, New York

Hendershot D (2000) *Smaller is safer—simplifying chemical plant safety*, Safe Workplace, National Council on Compensation Insurance, Boca Raton (Hendershot is a senior technical fellow in the Process Hazard Assessment Department of the Rohm and Haas Company, Bristol, PA, and has written extensively on Process Safety)

Hendershot D (1994) Chemistry—the key to inherently safer manufacturing processes. Presented before the Division of Environmental Chemistry, American Chemical Society, Washington, DC, 21 August 1994

Kletz T (1998) *What went wrong?*: *case histories of process plant disasters*. Gulf Publishing Company, Houston, May 1998 (Kletz is well known for his many publications and for bringing the term, "Inherently Safer Plants," into popular usage.)

Kletz T (1998) Process plants. In: A handbook of inherently safer design. Taylor & Francis, Philadelphia

Lees F (1996) Loss prevention in the process industries: hazard identification, assessment, and control. Butterworths, London

Loss Prevention Committee, Safety, Health Division, AIChE (1995) Proceedings of the 29th Annual Loss Prevention Symposium—(Serial), December, 1995

Publications by National Fire Protection Association (NFPA), 1 Batterymarch Park, Quincy, MA 02269. For a more complete list, see http://www.nfpa.org/Codes/CodesAndStandards.asp

NFPA 30 (2000) Flammable and combustible liquids code

NFPA 69 (1997) Standard on explosion prevention systems

NFPA 68 (1998) Guide for venting of deflagrations

NFPA 325 Guide to fire hazard properties of flammable liquids

PHA Software, "PHA Works," PSM Source (reference tool for OSHA's 1910.119) Primatech Inc., Columbus. http://www.primatech.com

Smith KE, Whittle DK (2001) Six steps to effectively update and revalidate PHAs. Chem Eng Progr 70–77

Thompson (2006) Chemical process safety report, Washington, DC

Internet References and WEB pages

American Institute of Chemical Engineers, Center for Chemical Process Safety. http://www.aiche.org/ccps/

American Society of Mechanical Engineers (2001) Boiler pressure vessel code. ASME International, Fairfield. http://www.asme.org/

CCPS (Center for Chemical Process Safety). American Institute of Chemical Engineers, New York

Chemical Safety Board (Incident Reports). http://www.acusafe.com/Incidents/frame-incident.htm

Manufacturers Chemical Association: http://es.epa.gov/techinfo/facts/cma/cma.html

Manufacturers Chemical Association (Responsible Care): http://es.epa.gov/techinfo/facts/cma/cmacommo.html OSHA Regulations & Compliance Links http://www.osha.gov/comp-links.html

OSHA Regulations (Standards—29 CFR) http://www.osha-slc.gov_OshStd_toc/OSHA_Std_toc.html

RMP Regulations http://www.epa.gov/swercepp/acc-pre.html. CCPS (Center for Chemical Process Safety)

American Institute of Chemical Engineers, New York (This is not a complete list. For a complete list, see on the Internet. http://www.aiche.org/pubcat/seadtl.asp?Act=C&Category=Sect4&Min=30)

CCPS (2009) Guidelines for chemical process quantitative risk analysis, 2nd edn. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York

CCPS (2011) Guidelines for engineering design for process safety, 2nd edn. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York

CCPS (2011) Guidelines for auditing process safety management systems, 2nd edn. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York

CCPS (1995) Guidelines for chemical reactivity evaluation and application to process design, 2nd edn. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York

CCPS (2010) Guidelines for evaluating the characteristics of vapor cloud explosions, 2nd edn. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York

CCPS (2003) Guidelines for investigating chemical process incidents, 2nd edn. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York

CCPS (1995) Guidelines for safe storage and handling of reactive materials. Center for Chemical Process Safety, American Institute of Chemical Engineers, New York

Managing an Emergency Preparedness Program

3

Thaddeus H. Spencer and James W. Bowman

Introduction

Prevention, Prediction, and Preparation

The preceding chapter explored many technical aspects of chemical process safety and some safety management systems that form the foundation of a comprehensive emergency preparedness program. Clearly, the first step in preparing for emergencies is to identify and mitigate the conditions that might cause them. This process starts early in the design phase of a chemical facility, and continues throughout its life. The objective is to *prevent* emergencies by eliminating hazards wherever possible.

Although hazard elimination is the goal, experience has taught us that guaranteed, failure-free designs and devices have so far eluded human kind, despite astonishing advances in knowledge and technology. Even the most "inherently safe" chemical facility must prepare to control potentially hazardous events that are caused by human or mechanical failure, or by natural forces such as storms or earthquakes.

The process of careful, structured analysis and evaluation used to eliminate hazards during design and construction will also allow a chemical facility to accurately *predict* unplanned events that may create emergencies, and to effectively *prepare* to manage them should they occur. A comprehensive emergency preparedness program has all of these elements: *prevention, prediction*, and *preparation*.

The fundamental need to predict and prepare for a failure of some kind is familiar to everyone. Fortunately, most of the failures that we encounter create little more than inconveniences in our lives. Others have much more serious potential. Such a failure can trigger an emergency, a term that Webster defines as "an unforeseen combination of circumstances or the resulting state that calls for immediate action." If the immediate action is ineffective, the emergency will escalate to a full-blown crisis.

Certainly most if not all of us in our personal experiences have had many opportunities to reconfirm the wisdom of the admonition, "Plan for the best, but prepare for the worst." As a result, we prepare ourselves for human and mechanical failure in a variety of ways, some so simple and familiar that we scarcely are aware that we are managing a personal emergency preparedness program.

To varying degrees, each of us has assessed our personal vulnerability to a specific emergency and the potential consequences to family and property. In some cases the required immediate action may be minor, but elsewhere the same combination of circumstances will demand significant resources to avoid a crisis of major proportions. The loss of household electric power in an urban condominium may only turn out the light, but on a farm, where electricity pumps the well water, milks the cows, refrigerates the produce, and irrigates the land, loss of power can bring disaster to the unprepared.

In an industrial environment, the consequences of human or mechanical failure can be far greater, even threatening the lives of employees and neighbors. Therefore, emergency prevention and preparedness efforts must have high priority, receiving continuous attention from every employee, including those at every level of management and supervision.

In the chemical industry, emergency preparedness programs have long been recognized as vital elements in protecting people, property, and the environment from harm. Few chemical facilities are without an emergency response plan (ERP). Still, when the alarms sound and the emergency is real, the response often does not proceed as planned. Too often, preparation for the unlikely event has been inadequate. As a result, many chemical facility managers are taking a fresh and critical look at their existing emergency preparedness programs in cooperation with their

---

T.H. Spencer (✉)
Safety and Environmental Management Services, E. I. du Pont de Nemours and Company, Wilmington, DE, USA

J.W. Bowman
Safety and Fire Protection Engineering, E. I. DuPont de Nemours & Company, Wilmington, DE, USA

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_3, © Springer Science+Business Media New York 2012

communities, and placing a still higher priority on being truly prepared for emergencies. This is consistent with the aggressive approach to other aspects of safety management that has made the chemical industry one of the safest industries.

Although the commitment to *prevent* the events that lead to injuries and emergencies of all kinds remains the first priority in safety management, we must "prepare for the worst." In this chapter we will explore how chemical facilities and their neighbors can better *predict* and *prepare* for unplanned events that threaten lives, property, and the environment—events that call for immediate and coordinated action.

Need for Emergency Preparedness Programs

The need for more effective management of emergency preparedness programs by chemical facilities and their host communities became painfully clear to the world in the 1980s—a decade marred by tragic events linked to the manufacture and distribution of chemical products. These events revealed serious deficiencies, not only in training people to react effectively during an emergency, but also in managing the systems employed to identify, evaluate, and mitigate hazards that may cause an emergency.

The chemical industry has vigorously responded to these problems in a variety of ways. In some cases well before thecrisis of the 1980s, individual chemical units took aggressive action to increase the reliability of their operations, to communicate these actions to their neighbors, and to involve the communities in the process of emergency response planning. This proactive interaction at the local level proved fruitful in promoting the active partnership needed to solve mutual problems and to respond to public concerns. In 1986, these individual initiatives were institutionalized by the industry in the Chemical Manufacturers Association's (now the American Chemistry Council [ACC]) "Community Awareness/Emergency Response (CAER)" program, which has in turn become a major element of the more recent "Responsible Care" program.

The proven success of these activities provides a powerful reason for giving high priority to open and cooperative management of an emergency preparedness program. By implementing such a program, a chemical facility builds important bridges to its neighbors and fosters positive perceptions. Involving the host community provides an opportunity to demonstrate a sincere commitment to the protection of people and the environment, and a route to the mutually profitable solution of real problems through meaningful two-way communication with the public.

Much of what had been voluntary prior to 1986 became law in the United States with the enactment of the Superfund Amendments and Reauthorization Act (SARA). The Emergency Planning and Community Right-to-Know portions of the Act (known as Title III, or EPCRA) place specific organizational, planning, communication, and training responsibilities on the public and private sectors, as do the accidental release provisions of the Clean Air Act (CAA) amendments enacted in 1990. Other laws, such as OSHA's 29 CFR 1910.114, Process Safety Management of Highly Hazardous Chemicals, also require ERPs. In addition, various state and local regulations must be considered, understood, and reflected in a complete emergency preparedness program.

These laws and regulations, which will be covered later in this chapter, should be viewed as describing only the minimum requirements. Beyond the legal and self-protective reasons for chemical facilities to improve their management of emergency response programs lie other important driving forces. One of these is the moral and ethical responsibility to employees and the public to work toward elimination of events destructive to the quality of life. In practice, full recognition of this responsibility requires actions in prevention, prediction, and preparation that go beyond the letter of the law.

The financial benefits of good emergency preparedness program management are important as well. Emergencies are always expensive. Uncontrolled emergencies can become financially devastating crises. Nervous communities, fearful of the dangers dimly perceived through industrial fences, can and have put some chemical operations out of business. Creating and maintaining a comprehensive emergency preparedness program does not come free. However, the investment is proving to be money well spent by most chemical facilities.

Preventing and Predicting Emergencies: Getting Started

The objective of a comprehensive emergency preparedness program is the protection of people, property, and the environment from unplanned hazardous events. Organizations handling hazardous materials recognize that the process of creating an effective program starts with the identification, evaluation, reduction, and control of hazards (commonly called risk management), and proceeds through the preparation, drilling, and maintenance of plans and procedures designed to contain an emergency situation should one occur.

To assure an effective emergency preparedness program, chemical facilities need to make periodic, formal assessments of their vulnerability to and preparedness for emergencies. Managers must be involved in developing and monitoring key indicators that will help in assessing an organization's ability to prevent or deal with an emergency. Awareness of any program weaknesses revealed by examining these indicators leads to corrective action to ensure that the potential for incidents decreases, and emergency

EMERGENCY PREVENTION	EMERGENCY RESPONSE PLANNING
1. Frequent management presence in operating areas.	1. Emergency manual that documents the areas or processes with emergency potential and describes the emergency response plan.
2. Proper storage and identification of hazardous materials.	2. Regular emergency response plan training, and drills each quarter.
3. Formal and systematic inspections of key equipment, safety devices, and safety interlocks.	3. Liaison with the community for response coordination.
4. Audits for compliance with safety rules and procedures.	4. Emergency response management organization with documented functional tasks and assigned personnel.
5. Periodic critical reviews of existing operating and maintenance procedures.	5. Systems to recognize and report an emergency in a timely manner.
6. Training programs updated to meet current needs.	6. Documentation of hazardous chemicals with potential to leave the site, and evacuation plan for affected areas.
7. Control systems for maintenance hot work and changes in process and equipment.	7. Adequate emergency response equipment.
8. Systematic process hazards reviews during design, start-up, and routine operations.	8. Proper method to account for personnel during and after an emergency.
9. Formal reviews all of events that could or did cause serious process incidents.	9. Procedures to review and modify the emergency plan following drills or actual emergencies.

Fig. 3.1 Emergency preparedness checklist

preparedness improves. Sample checklists including some of these key indicators are shown in Fig. 3.1.

Central to all emergency preparedness programs is a written ERP. An ERP obviously is a key element of emergency preparedness, yet it is only one procedural part of a systematic process that includes the following sequential steps:

Identify and evaluate hazards
Mitigate hazards wherever possible
Identify and evaluate remaining hazards
Identify and evaluate resources
Develop emergency procedures and ERP
Train facility personnel
Communicate plans to the public
Integrate with community ERPs
Conduct and critique drills
Review and revise ERPs
Do it again

The creation and maintenance of an ERP requires the allocation of valuable resources, as well as strong leadership from a manager. No one else in the organization has the authority to commit the resources required. Only the power of the manager's office can overcome the organizational inertia that is often encountered.

Many people do not like to "prepare for the worst." They may harbor sincere doubts about the value of planning for events that they consider unlikely to occur. Some people are complacent after years free of serious accidents, and honestly believe that "It can't happen here." Others may be concerned that an open discussion of potentially disastrous events will needlessly upset employees and neighbors. These and other "start-up" problems could make trouble for a manager initiating (or reviving) an emergency preparedness program.

Managers need to find ways to stimulate employee interest and enthusiasm in the planning process. This is best done through the involvement of employees. One strategy that minimizes problems is the early involvement of those employees who serve as emergency responders in their communities. Some of them have witnessed the tragic consequences of failing to heed early danger signs and being unprepared for events. Thus, they can be willing and able catalysts in the emergency response planning process.

Open communication with all employees early in the process is also important. Sharing objectives and encouraging contributions stimulates thoughtful discussion, leads to more complete hazard recognition and mitigation, and ensures a greater chance of effective action when the emergency alarms do sound. Armed with facts, employees acting as informal ambassadors to the community can reduce the potential for public alarm over a facility's preparations to deal with serious but low-probability events.

Hazard Identification and Mitigation

As shown in Fig. 3.2, emergency preparedness begins with the identification and mitigation of hazards. When properly done, the systematic analysis and evaluation of chemical process hazards stimulate actions that eliminate the potential for many emergency situations and pinpoint the situations that remain. The direction of these analytical and corrective efforts is an integral part of managing a comprehensive emergency preparedness program.

Process Safety Management Team

Hazard identification and mitigation are the responsibility of a facility's line organization—the same people who are responsible for all other organizational performance parameters. However, a line organization often needs support in identifying and mitigating complex chemical process hazards, support that can be effectively provided through the formation of a standing team functioning under the manager's direction. This team, called a Process Safety Management Team (PSMT), includes representatives of each unit of a facility from various levels of the organization. They meet on a regular basis (usually monthly) and report frequently to the manager on the status of their activities.

Identifying Hazards: PSR Teams

Supported by the PSMT, the line organization develops a structured approach for performing process hazard analyses on a repetitive basis. One effective way to do this is to organize knowledgeable facility personnel into ad hoc Process Safety Review (PSR) teams. These teams, supplemented by outside specialists as required, are responsible for studying all processes, identifying all potential hazards, and recommending appropriate corrective or control measures.

The makeup of a PSR team is critical to the success of this process. The effectiveness of the team depends on the skills, knowledge, and cooperative effort of its members, and the leadership ability of its chairperson. Each member must be familiar with the process being studied, and have at least a working knowledge of the basic engineering principles and chemistry involved. The team should include supervisors from operations and maintenance and a technical support person. Knowledgeable process operators and maintenance personnel would be valuable team members as well. Others who could contribute to the team's work include design engineers, specialists in electrical and instrument systems, safety engineers, reaction kinetics consultants, and equipment vendors.

The selection and training of PSR team members must be carefully monitored. Most managers make this the responsibility of the standing PSMT, which also may assist a PSR team in choosing the most appropriate hazards evaluation method for a specific process.

Review Methods

There are many structured methods for carrying out effective reviews of process hazards. The four most commonly used methods are:
What if/checklist
Failure mode and effect analysis
Hazard and operability study (HAZOP)
Fault tree analysis

What If/Checklist. The most frequently used method of process hazard review, the what if/checklist, is effective in reviews of relatively uncomplicated processes from raw materials to final product. The team formulates and answers "What if?" questions at each handling or processing step to evaluate the effects of component failures or procedural errors. They use a checklist to ensure that all important subjects are addressed. This method should be used as the first step in all process hazard reviews.

Failure Mode and Effect Analysis. When the team studies a specific item of equipment, such as a reaction vessel, they often use the failure mode and effect analysis method. Its semi-quantitative approach assists in prioritizing hazards.

HAZOP. The HAZOP procedure systematically questions every part of a process to discover how deviations from the intention of the design can occur, and to determine if the consequences of such deviations are hazardous.

Fault Tree Analysis. Fault tree analysis, the most complex of the commonly used methods, is employed to determine the possible causes of a preselected undesired event. Through the use of logic diagrams and failure rate data, the team can make a quantitative evaluation of the frequency of the undesired event.

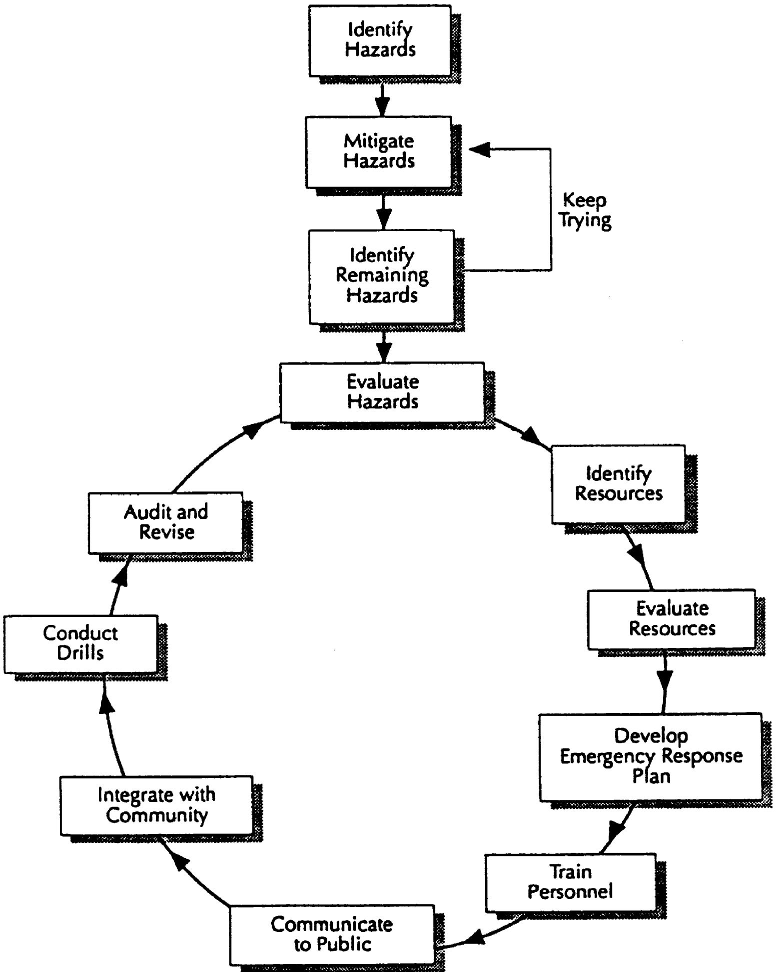


Fig. 3.2 Elements of a comprehensive emergency preparedness program

Recommendations and Reports

Regardless of the method used, the PSR team's most important responsibility is to alert management to serious hazards that may have been overlooked or given inadequate attention. To fulfill this responsibility, a PSR team must take the following steps:

1. Identify the hazards that could cause explosion, fire, release of toxic materials, serious injury, or inappropriate exposure to chemicals.
2. Evaluate the magnitude of the hazards for the areas of probable involvement; the consequences of an event in terms of injuries, environmental harm, and property damage; and, qualitatively or quantitatively, the probability of the hazards' occurrence.
3. Develop practical recommendations to eliminate or control the hazards identified.

The PSR team reviews in detail up-to-date reference material such as:

☐ Architectural drawings
☐ Equipment layout drawings
☐ Process schematics
☐ Instrument diagrams
☐ Chemical and physical characteristics of process materials
☐ Equipment design specifications
- Operating procedures
☐ Process conditions
☐ Emergency shutdown procedures

When a facility has more than one chemical process, the PSMT develops a priority order for reviews and recommends a review frequency to the facility manager. The frequency usually ranges from once every 2–3 years for high-hazard-class processes such as explosives manufacture or acetylene purification, to once every 5–7 years for low-hazard-class processes such as alcohol purification, steam generation, and operations involving combustible materials.

Changes not anticipated in the original design of equipment often pose serious problems. Some examples of such changes are:

☐ Introduction of different raw materials
☐ Changes in temperatures, pressures, speeds
☐ Deterioration of equipment

At the completion of each PSR, the team prepares a written report that defines needs, makes recommendations to remedy problems, and recommends priorities for the correction of deficiencies. Following review and acceptance of the report, the facility manager assigns responsibility for corrective action to the appropriate operating personnel. The PSMT then assists the manager in monitoring the status of the recommendations from all PSRs.

Mitigating Hazards: Release Detection and Mitigation

The release of flammable or toxic chemicals from uncontrolled pressure relief vents or as a result of equipment failure may present a serious threat to employees or neighbors who fail to guard against exposure. Every chemical facility must address this potential problem and prepare to protect people from these hazards by working toward reducing the potential hazard and its consequences. Typically, the line organization and PSR teams are responsible for hazard mitigation. They conduct a hazard study that includes the following activities:

1. Detailed appraisal of the potential for an accidental release of toxic gas or vapor.
2. Evaluation of instruments and other methods for detecting such leaks.
3. Provisions for rapid alerting of threatened personnel, and for communicating with emergency responders.
4. Identification of buildings in which people might be trapped by such a release.
5. Assessment of the capability of buildings or rooms to prevent the infiltration of gas or vapor.
6. Evaluation of plans for building evacuations, including the provision and maintenance of appropriate personal protective equipment.

Early *detection* of a chemical leak is necessary in order to limit its effect on people and the environment. The most fundamental method for detecting a chemical release is the systematic patrol of operating units by personnel trained to recognize potentially hazardous vapors using odor or visual observations. The frequency and scope of the patrols will vary with the nature of the process equipment and materials; however, every facility should have documented patrol procedures in place.

When particularly hazardous vapors are present, the patrol procedures should include special provisions to protect the patrollers from the fumes. For example, in facilities producing or consuming hydrogen sulfide, operators may carry emergency respiratory equipment to escape any unexpected fumes that they may encounter, and patrol in pairs or individually, under constant surveillance.

Many chemical facilities supplement operator patrols with an instrumented detection system. Such a system may be a necessary resource when a hazards study concludes that the system will substantially increase the available escape or emergency response time, or where:

1. The harmful substance is odorless or deadens the sense of smell at hazardous concentrations.
2. The harmful substance is toxic at concentrations undetectable by smell.
3. Large numbers of people may be exposed quickly.
4. Ventilation systems might draw toxic fumes into a building before other means of detection could trigger protective action.

At the core of an instrumented leak detection system is a gas detector. There are many kinds of detectors on the market with varying degrees of sensitivity and selectivity. All require careful regular testing and maintenance.

An engineering study of many site-specific factors is required before one makes a choice and designs an appropriately instrumented system. Some generally accepted guidelines are:

1. Early warning of a leak is enhanced if the sensors can be placed near the process equipment subject to leakage.
2. Air movement characteristics are critical in achieving reliable detection.
3. Most detectors respond to several gases or fumes, so the possible presence of all airborne substances affecting the detection system must be considered.
4. When it is necessary to monitor work areas or the intakes to ventilation systems, a highly sensitive system is desirable.

Most detection systems are designed not only to report the presence of hazardous fumes through instrument readouts but to sound an alarm and automatically initiate corrective or protective action. In an office or shop, for example, the system can be designed to shut down all ventilating fans and close exterior air inlet dampers.

When process safety reviews have identified chemical releases as potential sources of facility emergencies, the organization must provide the training and materials needed to ensure a prompt and appropriate reaction to *mitigate* the hazards. Some countermeasures that are effective in limiting the spread of a hazardous material spill or release should be included in the design of chemical process equipment, and described in emergency response procedures. The following discussion of relea e mitigation is largely excerpted from *Guidelines for Vapor Release Mitigation*, prepared by

R. W. Prugh and R. W. Johnson for the Center for Chemical Process Safety of the AIChE (copyright 1988 by the American Institute of Chemical Engineers, reproduced by permission of the Center for Chemical Process Safety of the AIChE).

"Water, steam, and air curtains and water sprays are primarily effective in dispersing and/or diluting vapors with air to reduce the severity of effects of a hazardous vapor release. In some cases, vapors can be partially 'knocked down' or absorbed after release."

"Ignition source control and deliberate ignition are also possible vapor release countermeasures." "For areas around processes handling flammable vapors, ignition source control is practiced to reduce the probability of vapor ignition if a leak occurs." "Administrative controls are exercised on plants where flammable materials are processed. Such controls may include hot work permits, restricted smoking areas, not allowing lighters or matches on the site, and electrical grounding and bonding procedures."

"Deliberate ignition is a countermeasure against spills of highly toxic materials which are also flammable, such as hydrogen sulfide, hydrogen cyanide, and methyl mercaptan. Igniting nontoxic flammable materials such as hydrocarbons may present hazards outweighing possible advantages." In any case, deliberate ignition must be carefully planned and executed so that the resulting fire is truly controllable.

"Practical methods for combating vapor from liquid leaks are dilution, neutralization, or covering. All three reduce the vaporization rate of the pool. Water dilution is effective for spills of water-miscible or water-soluble material. Spraying water into the spill reduces the vapor pressure by reducing the concentration of the liquid." "For acidic spills, limestone or soda ash is often used" to react with the spilled liquid to produce a less volatile salt or ester. "A foam cover can be effective in reducing vaporization from spills," and "dense liquids can be covered with lighter immiscible nonreactive liquids" to accomplish the same thing.

Preparing for Emergencies: Identifying and Evaluating Resources

The process of hazard identification, evaluation, and mitigation, when sustained over a period of time and coupled with other good safety management practices, can prevent most process-related emergencies. However, the potential for various low-probability process-related events will remain. Should one occur, there must be resources available to promptly bring the event under control. These resources include designated personnel, plans, systems, and facilities that are needed for effective action and communication. The facility line organization is responsible for providing and maintaining these resources.

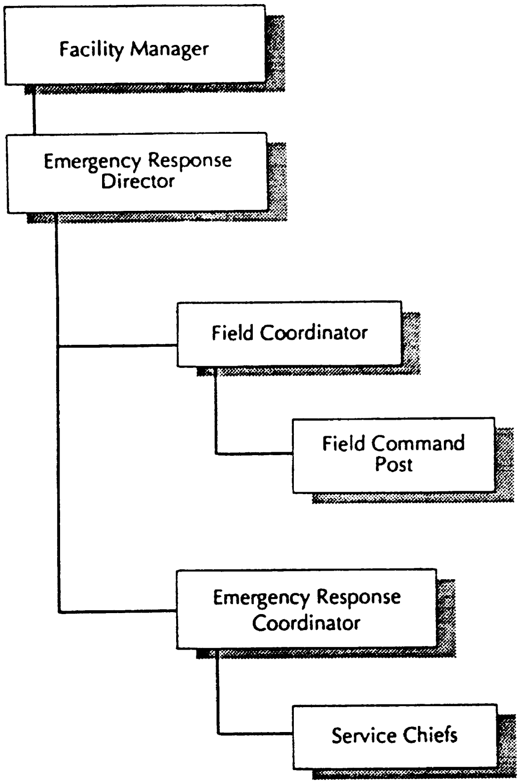


Fig. 3.3 Generic emergency management organization

Personnel

Emergency Preparedness Team. As with the PSMT and the PSR team, the efforts of the line will benefit from the support of a standing Emergency Preparedness Team functioning under the direction of the facility manager. This team, which includes representatives of each unit of the facility, leads the development and maintenance of an ERP for the facility, and monitors specific procedures and training for handling emergency situations of all kinds. It should meet on a regular basis (usually monthly), and report frequently to the manager on the status of its activities.

Emergency Management Organization (EMO). Emergencies demand rapid, well-coordinated decisions, communications, and action to bring them under control as swiftly as possible. There must be a formal EMO whose purpose is to achieve this objective. The structure of the organization is not critical as long as it is capable of rapid assembly, carries out its assigned responsibilities, and meets the needs of the facility. A generic EMO is shown in Fig. 3.3; the EMO is described below, in the section on "Developing an Emergency Response Plan."

Fire Brigades. Most chemical facilities of any size have established fire brigades as a key resource in their emergency preparedness programs. By virtue of its training and familiarity with the equipment and the physical layout of a site, the brigade can be a major factor in *preventing* incidents as well as in minimizing injuries and property damage due to fire, explosion, or other causes.

The size and the structure of a fire brigade vary, depending on the hazards present, the size of the facility, and the level of internal fire-fighting capability desired; and large, multi-process facilities usually have several unit or area brigades. Each brigade has a designated captain and an alternate. The training of brigade members must be commensurate with the duties and the functions that they are expected to perform, and thus depends on the fire prevention plan option selected by site management. (Fire prevention plans are discussed later in this section).

It is best to have a written statement establishing the brigade and its duties, specifying its size and organizational structure, and outlining the type, amount, and frequency of training provided. In the United States, OSHA has promulgated minimum standards for fire brigades, which may be found in 29 CFR 1910.156.

Plans

Work Unit Plans. In a large facility, each work unit has emergency procedures for its own area of operations, and these work unit plans form the foundation of the facility ERP. Thus, the facility's emergency response can be no more effective than the recognition and response capabilities of the employees in the facility work units.

The primary objectives of these unit plans are (1) to control and contain any emergency condition within the unit, and (2) to provide protection for unit personnel and equipment from events originating outside the unit. Unit plans specify who has authority to take emergency action, and how escalation to a full facility emergency occurs. Subjects that must be covered include:

☐ Unit emergency command and coordination
☐ Communications within the unit
☐ Communications with other units
☐ Emergency assessment
☐ Unit evacuation and personnel accounting
☐ Emergency shutdowns
☐ Communications with the facility management
☐ Criteria and procedures for securing resources from outside the unit

Unit plans clearly assign primary responsibility for initial emergency assessment and reaction to the lead unit employee on site at the time of the event. Among the factors to be considered by that individual in making the initial assessment are:

☐ Previous experience with similar situations
☐ How long the situation has existed
☐ What might occur "If . . ."
☐ Properties and hazards of materials involved
- Other complicating situations

Using the information gained from the assessment, unit personnel take whatever immediate action is required to protect people and property on and off the site.

Work unit plans should contain guidelines for assuring uniformity in the assessment of and reaction to unusual events. An effective approach is to develop an emergency classification system that includes criteria for classification and guidance for the appropriate response actions for each level of emergency that is defined.

For example, an emergency judged to be controllable within the unit, with no evacuation outside the unit and no impact in the community, is a Level One Emergency. Small releases of hazardous materials usually fall into this category. Appropriate response items to consider would include:

☐ Protection of unit personnel
☐ Shutdown or isolation of affected equipment
☐ Notification on- and off-site
☐ Containment and cleanup

Similarly, other events are categorized by their potential or actual severity. Preplanned response checklists are prepared for each one. An incident that requires response by facility personnel outside the affected unit but has no impact on the community is a Level Two Emergency. A Level Two Emergency requires activation of the facility ERP. If some community impact becomes probable, the emergency is a Level Three Emergency. That level of emergency triggers additional response requirements, including interaction with the community.

The unit plan must be formal and in writing. Because the emergency procedures of each individual operating area or work unit form the foundation of the facility ERP, they should be reviewed by the Emergency Preparedness Team for completeness and consistency across the site before they are included in the facility ERP.

Fire Prevention Plan. One of the most basic resources that all chemical facilities must have is a fire prevention plan. To be fully effective, this plan must be in writing, and it must be reviewed with all employees on an established schedule.

A comprehensive fire prevention plan includes the following subjects:

1. Major facility fire hazards
2. Storage and handling practices for combustible and flammable materials
3. Identification of potential ignition sources
4. Procedures to maintain systems and equipment installed to control ignition sources
5. Names and titles of those responsible for the system and equipment maintenance and the control of fuel source hazards

6. Procedures to minimize accumulations of flammable and combustible waste materials
7. Plans for communicating to all employees the fire hazards and their specific responsibilities in the event of fire

Fire response options open to facility managers range from a plan to evacuate all employees without attempting fire fighting on any scale, to full involvement of an established fire brigade in fighting advanced-stage structural fires. The fire prevention plan must specify which option applies to the facility. The option selected determines the type and extent of education and training required, as well as the type of equipment needed on the site.

Fire prevention and response information may be covered in other written facility documents (such as job descriptions, rules, or procedures) and communicated to employees on the same schedule as the fire prevention plan.

Evacuation Plans. Many emergency events require the evacuation of at least some facility personnel, if only for precautionary reasons. This presents other needs, which must be addressed in the evacuation plan. Personnel who are directed to evacuate their normal work stations need visible wind direction indicators to help them move away from a dangerous gas cloud drifting downwind of the release point. If the building is enveloped in the cloud, however, they will need personal protective equipment in order to evacuate the building safely. When there are processes that must be rapidly shut down and/or isolated as evacuation proceeds, the plan must include written procedures and appropriate means to ensure incident-free action.

Headcount Plan. The evacuation plan also must establish gathering points for all personnel in order to identify those who are missing.

Alarm Systems

The procedure for accounting for personnel, called the headcount, must work effectively and rapidly in the first minutes of the emergency to determine if anyone is missing or known to be injured. Communications to the fire brigades or other first responders must be quick and accurate so that search and rescue operations and medical aid can be successfully implemented.

Each work unit in a facility must have its own plan for headcounts as a part of its emergency preparedness program. This plan must be in writing, and it must be practiced frequently enough to ensure good execution in a time of real emergency. It should include:

☐ A designated assembly point for each employee
☐ An alternate assembly point, should the primary location be inaccessible
☐ A procedure for counting at each assembly point
☐ A designated unit headcount coordinator and backup
☐ Primary and alternate phone numbers where coordinator may be reached
☐ Phone numbers for reaching the personnel chief, who coordinates and summarizes the site-wide accounting

Headcount procedures must include guidelines for accounting for visitors to the facility at the time of the emergency. These visitors may include truck drivers, vendors, and contractors. Some guidelines are:

1. All visitors must sign in and out, preferably at one designated gate.
2. Visitors will have a designated host employee responsible for their safety whenever possible.
3. Everyone temporarily on the facility must report to the assembly point of the unit he or she is in at the time of the emergency.
4. Unit procedures must include reporting these “extra” people by name to head-count headquarters.
5. The personnel chief of the EMO must have access to the log at the visitors’ gate.

Another essential resource that every chemical facility must have is an effective alarm system, which initiates action by endangered personnel and emergency responders.

A satisfactory facility emergency alarm system must meet three additional requirements. It must, at a minimum, immediately alert *all* the people on the site. (A single signaling device should be adequate for small sites; however, large facilities may require many devices placed to reach all occupied areas of the site.) The system must function even when the facility has a general power failure. Alarm activation controls should be located so that an emergency condition is unlikely to prevent access to them.

On most older chemical facilities, the basic component of the emergency warning system is the fire alarm. It is essential that there be different signals to distinguish between an actual fire and other threatening events, as the action taken in response to a toxic gas release, for example, may be quite different from that taken in response to a fire. On multi-process sites, the signals also should identify the work unit involved.

Depending upon the number of people at a facility and their familiarity with response and evacuation procedures, it may be necessary to supplement signals with verbal instructions delivered by public address equipment, radios, or automated telephone systems. For reliability at any time of day, a system using verbal instructions must be located in a regularly manned job station such as a control room or gate house.

Any alarm system used must communicate clearly the nature of the emergency event and its severity. In a Level One Emergency some people may have to be evacuated in order to ensure their safety, but by and large the emergency can be contained within the unit. In that case, information usually may be given to unaffected employees through normal telephone communications.

In the event of a Level Two or a Level Three Emergency, each facility must have an easily recognizable signal that clearly communicates that a major incident has occurred. Some facilities have alarm boxes that will automatically sound this signal when pulled. Others have boxes that require manual operation. Still others rely on a telephone message to trigger the signal.

In some facilities, the “major” emergency alarm is a steady blast on the facility steam whistle. Others use sirens or special bells. Whatever is chosen, the alarm must reach everyone on the site, triggering activation of the facility ERP and the initiation of a head-count procedure.

It is essential that every employee on a chemical facility knows how, where, and when to turn in an appropriate and effective emergency alarm. The specifics will vary considerably, but the procedure for turning in an alarm is similar in all cases.

Employees first need to know how to report a fire. On most facilities, this is done by using a fire alarm box or by telephoning a central station such as a guard house from which the fire alarm can be sounded. Chemical facilities need to have well-developed training and drill programs in place to be sure that response procedures and equipment are maintained.

When fire alarms are supplemented by verbal instructions, the facility personnel who are contacted need to know the location and the nature of the fire, the action under way to control it, and the actions required (if any) by those not involved. There are many commercially available communications systems that deliver this information efficiently, even on large sites.

Facilities for Protection and Communication

Safe Havens. When accidental releases of a toxic gas occur or threaten to occur, the immediate protection of on-site personnel is of paramount importance. One resource available for this purpose is a building or room that by the nature of its construction and its heating and ventilating characteristics can prevent the infiltration of intolerable concentrations of the toxic substance. The best location for a safe haven is determined by an engineering study. That location then is documented in the work unit and facility ERPs.

Safe havens may be rare on some sites; however, many buildings can provide personnel with temporary protection until the incident has been analyzed and a decision made on the need for evacuation. To qualify as a designated temporary safe haven, a building must be reasonably well sealed against air infiltration, with adjustable ventilation systems that can reduce or close off exhaust vents and outside makeup air. Emergency procedures should state how long a building can be considered a safe haven under specified exposure conditions.

Main Emergency Control Center (MECC). Using a list of safe havens developed by the work units, the Emergency Preparedness Team helps management select one building or room to be used as a control center in case of a major emergency. This MECC should be as remote from potential hazards as possible in order to serve as a reasonably safe haven to those involved in directing, coordinating, and communicating activities for the duration of an emergency.

MECCs need to have basic information readily available and maintained in an up-to-date, standby condition. Such a collection would include:

- ☐ Copies of the facility ERP
- ☐ Facility maps and diagrams
- ☐ Process material isolation points
- ☐ Fire control maps and diagrams
- ☐ Maps of the surrounding area
- ☐ Aerial photos of the facility and the surrounding area
- Names, addresses, phone numbers for: —all facility employees—off-site company people to be notified—groups and organizations who may be notified—community officials who may be notified
- ☐ Material Safety Data Sheets (MSDSs) for facility materials
- ☐ MSDSs for the materials of neighboring chemical facilities
- ☐ Copies of mutual aid agreements
- ☐ Highlights of the facility’s history, products, and performance

The MECC should have the following basic equipment:

- ☐ Adequate telephones and lines
- ☐ Unlisted telephone and/or a hotline
- ☐ Two-way radios
- ☐ Fume path projector
- ☐ FAX machine
- ☐ Regular radio with tape recorder
- ☐ Regular television with recorder and playback video cassette recorders
- ☐ Chart pads and stands
- ☐ Battery-powered lighting
- ☐ Personal protective equipment
- ☐ Hand-held tape recorders

Alternate Control Centers. Recognizing the unpredictable nature of a developing emergency, it is advisable to select at least one alternate on-site main control center. Both the MECC and the alternate on-site center should have backup power supplies.

The presence of substantial quantities of explosive or toxic materials on-site may justify establishing yet a third location off-site, in case of a complete evacuation of the facility. This could be a fixed location in the nearby community, or a mobile unit such as a truck or van properly equipped for managing an emergency. Some of the supplies and equipment for the alternate off-site main control center may have to be packed in a readily transported kit rather than being on standby at the alternate off-site location.

Media Headquarters. In an emergency, it is essential to maintain regular contact with the public and the media. A separate location for communications will allow that contact to continue without interfering with the operations of the MECC or overloading its communications system. There, the public affairs chief and the facility manager can receive media representatives and provide periodic updates on the emergency situation. The media headquarters should be equipped with several direct outside telephones and the facilities to make the reporters comfortable for the duration of the emergency. Permanently mounted facility plot plans or aerial photographs are helpful resources, as are handouts describing in general terms the facility and its products.

Developing an ERP

Following the identification and evaluation of the resources available to the facility, the Emergency Preparedness Team, under the direction of the manager, develops the written ERP for the entire facility. The Emergency Preparedness Team is responsible for ensuring that the written facility ERP informs all employees of their roles in an emergency and the hazards to which they may be exposed.

An ERP must be tailored to a specific facility, reflecting its unique conditions and individual needs. Among the variables that affect the details of a plan are:

☐ Materials used, produced, or stored
☐ Nature of the operations
☐ Available employee skills
☐ Geographic location
☐ Proximity to other facilities
☐ Available emergency resources
☐ Mutual aid agreements

No two ERPs will be exactly alike, nor is there one best outline. The plan of a production facility will differ from that of a warehouse. However, every plan must include:

☐ Initial alarm procedures
☐ Emergency escape procedures and routes
☐ Emergency operating and shutdown procedures for critical operations
☐ Accounting procedures for all on-site personnel
☐ Rescue and medical duties for specified personnel
☐ Procedures for communicating the emergency inside and outside the organization
☐ Structure, duties, and resources of the EMO, including the names or titles of people with detailed knowledge of the plan and its assigned duties
☐ Reference material

Plan Design

The ERP is a working plan, applicable to any event with emergency potential occurring at any time and at any location on the facility. It should be written concisely, with diagrams and checklists used wherever possible, so that it may be effectively used for guidance during an actual emergency. To accomplish this, it is helpful to divide the ERP into several major sections, which are in turn divided by subject.

For example, an ERP may be broken down into three sections: the EMO, Action Plans, and Reference Material.

The first section documents the structure, duties, resources, and communications systems for the facility EMO, and the conditions that will trigger the EMO's response. It includes plans for site-wide notification and response.

The second section includes detailed action plans for each particular type of emergency. This section includes summaries or outlines of the emergency procedures developed by each work unit of a facility, and may contain considerable process-oriented information.

The third section consists of reference material on plan philosophy, training, plan maintenance, drills, and similar supportive data that usually are not needed at the time of an emergency. Members of the EMO and others with assigned response duties should be able to quickly locate and refer to the appropriate action guidelines, to determine that key functions are being performed.

Using this approach, the outline of an ERP for a multiprocess chemical facility would look like this:

Section I
A. Table of Contents
B. EMO Structure and Tasks
C. Notification: On-Site
D. Accounting for Personnel
E. Headcount Center
F. EMO Center
G. Emergency Scene
H. Notification: Off-site
Section II
A. Table of Contents
B. Work Unit 1
  1. Emergency types
  2. Action checklists
  3. Responsibilities
  4. Communications
  5. Shutdown procedures
C. Work Unit 2 (same as above)
Section III
A. Table of Contents
B. Plan Philosophy and Objectives
C. Training Programs
D. Drills and Tests

E. Return to Normal
F. Appendices
  1. Special situations (e.g., bomb threats)
  2. Hazardous material data
  3. (Others as required)

EMO Structure

Perhaps the most essential element of an ERP is the establishment of the organization that will manage an emergency response once the event has developed beyond a Level One Emergency. The EMO must be capable of rapid assembly, and must be able to carry out all of its assigned responsibilities.

The EMO is composed of the facility manager, the emergency response coordinator, the field coordinator, and various service chiefs. The number of service chiefs and their duties will vary. There need not be a separate individual in charge of each service, and small facilities may combine two or more services under one individual, as long as each function is performed.

The facility manager, in cooperation with the Emergency Preparedness Team, assigns people, by their title or function at the facility, to the EMO, with designated alternates for each EMO function. The duties and responsibilities of each EMO function are defined and documented in the ERP. Because there must be a functional EMO in place regardless of when a facility emergency occurs, designated shift personnel take specific EMO responsibilities until the primary members can reach the site.

All the positions established for the EMO should have written position descriptions, which may be placed in the third section of the ERP as an appendix to the plan. In the first section of the ERP, the EMO summary need only include a concise checklist for each position, with diagrams showing relationships and major communication flow paths.

Facility Manager. In the EMO, the facility manager has the overall responsibility for protecting people on-site, facility property, the environment, and the public during and after an emergency. With the assistance of the public affairs and communications chiefs, the manager usually serves as the spokesperson for the facility and the company, communicating with representatives of the media and other concerned audiences.

Emergency Response Coordinator. The designated emergency response coordinator, who may also serve as the chairperson of the Emergency Preparedness Team, directs all emergency control activities from the MECC. All other service chiefs on the EMO report to the coordinator. Using information from the emergency scene and from the service chiefs, the coordinator makes the key decisions on what should be done, and coordinates activities on and off the site. The coordinator reports to the manager, who should be available nearby for overall guidance and counsel.

Field Coordinator. The job of the field coordinator is to correct the emergency situation as rapidly as possible with minimum risk to those in or near the affected area. He or she establishes a field command post as close to the scene of the emergency as can be safely done. The post often is in or near a radio-equipped building or vehicle, thus permitting rapid establishment of communications between the command post and the emergency response coordinator.

Ideally, the field coordinator is familiar with the operations and materials involved in the emergency; so she or he often is the highest-ranking supervisor of the affected area who is available at the time of the event. That supervisor may remain as field coordinator for the duration of the emergency, or may be relieved by another designated member of management.

In the latter case, the area supervisor may become a service chief with greater hands-on involvement, with the field coordinator concentrating on marshaling required resources and maintaining effective communications. The choice depends to a great extent on the size and complexity of the facility, and the resources available for the EMO.

Public Affairs Chief. Working in close coordination with the facility manager, the public affairs chief releases appropriate information to the news media, regulatory personnel, government officials, and other public groups and individuals. No information is to be made public by anyone other than the facility manager without specific direction from the public affairs chief. The objective is to provide full and accurate statements in a timely fashion, so that public attention is focused on facts and useful information rather than on rumors and speculation.

The public affairs chief also establishes and monitors the media headquarters, which is isolated from the EMCC to avoid interference with operations there. The public affairs chief's function includes assisting the manager in the preparation of formal statements and background information to be distributed to reporters. It also may include arranging with local radio and television stations to make periodic announcements during an emergency so that the public and the employees not on the site are properly informed. The public affairs chief also arranges to monitor and perhaps record the broadcasts of local radio and television stations in order to determine what additional statements or clarifications to the public may be required.

Communications Chief. The communications chief establishes and maintains communications capability with appropriate people on and off the site. This chief must be familiar with the various communicating systems available, including telephones, public address systems, two-way radios, and messengers. The job includes recommending and coordinating revisions or additions to the communicating systems during an emergency, and assisting the facility manager and the public affairs chief with any communications as needed.

Engineering Chief. The engineering chief's primary responsibility is to maintain electrical power for vital services. These services include on-the-scene portable lighting, continuous fire pump operation, and a steady supply of nitrogen, steam, and process cooling water. Another key duty is to assemble repair groups capable of isolating damaged sections of pipelines, electrical lines, and other necessary equipment in order to contain problems and maintain or restore operations outside the affected area. Because these repair groups must be drawn largely from site personnel such as electricians, welders, riggers, and pipefitters, the position of engineering chief should be filled by someone from the facility maintenance organization.

Emergency/Fire Chief. The emergency/fire chief is responsible for fire-fighting and fume control activities. The best person for this job has good knowledge of and access to the fire-fighting, rescue, and fume control resources available on and off the site. On a small facility this may be the captain of the fire brigade. The job includes direction of designated facility personnel, and coordination with any outside forces brought in to bring the fire or fume condition under control. The environmental chief may assist in determining optimum courses of action based on actual or threatened adverse effects on air and water emissions from the facility.

Medical Chief. The medical chief ensures that the proper medical care is provided to people on the facility who have been injured or exposed to toxic materials. When facilities have medical professionals on the site, one of them carries out this function. This chief's responsibilities include establishing field stations to treat affected personnel, and, in cooperation with the transportation chief, providing transportation for injured people to other medical facilities.

The medical chief also participates in discussions with community officials and appropriate facility personnel regarding actual or potential medical problems for people outside the site boundaries. Representatives from nearby local hospitals, ambulance services, fire departments, police forces, and emergency management groups also may be involved in these discussions. The medical chief communicates any action or contingency plans developed in the discussions to the emergency response coordinator and the public affairs chief.

Environmental Chief. The environmental chief oversees all activities designed to minimize adverse effects on the quality of air and water as a result of an emergency. The function includes coordinating air and water quality monitoring on and off the site during and after the emergency, and assisting the emergency/fire chief in selecting the optimum approach to abating a fire or fume condition. The environmental chief also provides assistance in projecting the path and concentration of a fume release, using computer modeling (if available) or maps with plastic overlays, and developing an effective plan of action.

Personnel Chief. The personnel chief is responsible for accounting for all personnel on the facility at the time of the emergency. Unit emergency programs must include training on how this is accomplished. Unit supervisors initiate the process with headcounts at designated rally points. The success of rescue and medical activities depends on how quickly and thoroughly this information is obtained and reported to the personnel chief.

The personnel chief also coordinates communication with relatives of injured or deceased employees, and makes certain that this is completed before any names are released publicly. This requires close coordination with the public affairs chief and the manager.

The personnel chief need not be located at the MECC, but it is imperative that the personnel chief be in close touch with other members of the EMO. Telephones in the headcount center should have answering and message recording equipment to capture any unit reports arriving before the headquarters is staffed.

Security Chief. The security chief makes sure that entry to and egress from the facility are properly controlled. This involves securing gates; limiting entry to authorized personnel; registering all who pass through gates; meeting visitors, including representatives of the media, and escorting them to proper locations; and controlling all traffic on the site. The function also includes communicating with local police so that access to the facility is maintained, and crowd control procedures are enforced. Usually, the individual serving as security/chief at the time of the emergency also is responsible for initiating procedures to summon key facility personnel and urgently needed outside agencies.

Transportation Chief. The transportation chief coordinates and controls all transportation on the facility. This includes directing the assembly of available vehicles and crews, and identifying needs beyond site capabilities such as cranes, trackmobiles, and bulldozers that must be obtained from outside organizations. The function also includes providing suitable transportation for facility employees who monitor the effects of emergencies beyond site boundaries, or interact with community officials at an off-site location.

Training Personnel

Having a written facility ERP that is supported by established work unit emergency plans and procedures it is an important part of the manager's job to ensure that unplanned events will be promptly controlled with minimum risk to people, property, and the environment. There is much important work to be done, however, before the ERP is anything more than a paper resource. Facility personnel must be trained to use the ERP effectively. There must be frequent drills to test the plan and the people against the standards established by management as well as those established by law.

Within the facility, there are three groups of people who require training. First, there are members of the fire brigade, who must be trained to fight fires at the level specified by facility management. Next are the employees who have been assigned active roles in controlling emergencies of all types. This group includes members of the EMO and their alternates, plus designated support people such as headcount coordinators, guards, and emergency repair personnel. Finally, everyone else on the facility must have a basic understanding of the ERP, and must know how to respond when specific alarms sound.

The facility manager is responsible for ensuring that the appropriate training and retraining are done in a timely and effective manner. An employee should be trained when he or she is hired, at least annually thereafter, and when the employee's work area changes or the plan is revised. Most managers make this primarily a line organization function. They expect facility supervisors to use all available resources and means of education to accomplish the tasks. Some specialized training assistance, however, must be provided, particularly for the members of a fire brigade.

Fire Brigade Training

Training programs for fire brigades have two major objectives. One is to inform the brigade members of new hazards at their facility and innovations in fire-fighting techniques and equipment. The other objective is to provide hands-on training for developing skills in emergency operations and using equipment, including:

☐ Portable fire extinguishers
☐ Hoses and accessories
☐ Portable lighting
☐ Forcible entry tools
☐ Ladders
☐ Salvage equipment
☐ First-aid supplies
☐ Replacement parts
☐ Personal protective equipment
☐ Transportation equipment

A comprehensive training program for fire brigades must include classroom and hands-on training.

Outside resources can provide valuable assistance in the education and training of brigade members. Local fire departments and state fire schools are usually enthusiastic partners in such efforts. Often they are the key to securing adequate resources at a reasonable cost. In industrialized communities, mutual aid agreements may include cooperative training provisions with other chemical facilities, which provide opportunities for even more effective use of available resources. All mutual aid agreements should require cross-training in special hazards at the other facilities.

EMO Training

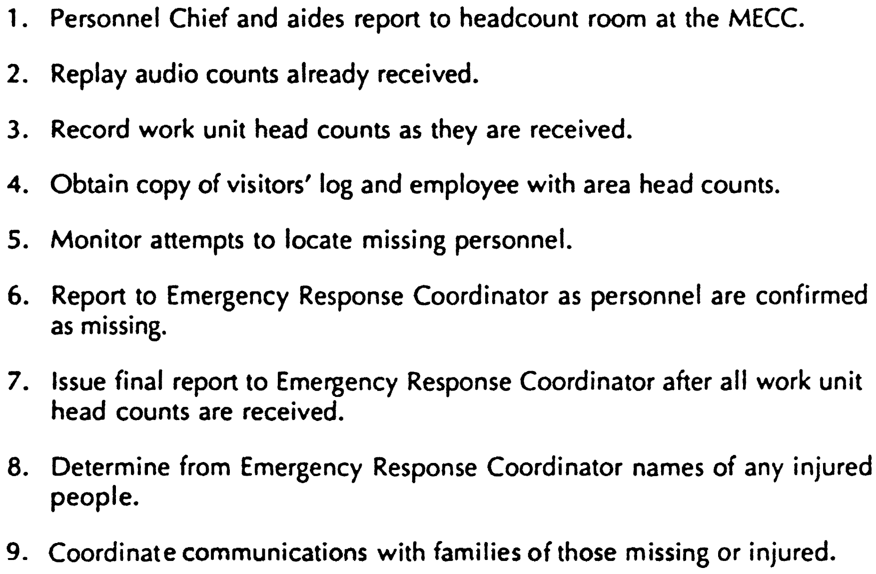
Training for personnel assigned to the facility EMO can be led by the emergency response coordinator, who meets periodically with each member of the EMO to review and refine position descriptions and the associated functional checklists. An example of such a checklist is shown in Fig. 3.4. The meetings may be followed by limited drills involving only the people and responsibilities included in the individual EMO function. Some facility managers assign specific emergency response training duties to each member of the EMO.

For example, the engineering chief organizes, equips, and trains the repair groups who will be called on to physically stabilize a situation at the time of an emergency. The emergency/fire chief could be given the responsibility for maintaining a trained force of fire and fume fighters, which includes competent leadership on all shifts.

Employee Training

There are many ways to be sure that all employees understand the ERP. The Emergency Preparedness Team can assist in the assessment of existing training programs, alerting the organization to training weaknesses and suggesting or providing creative ways to overcome those deficiencies. Unit supervisors must periodically review the ERP in scheduled group safety meetings or with individuals. Key plan elements can be reproduced on wallet-size cards, desktop displays, or telephone stick-ons. Individual or group discussions of how to react in given situations can detect weaknesses in procedures, training, or understanding.

Training programs for the three groups of employees discussed here should be documented, reviewed regularly, and included in a reference section of the facility ERP. But



1. Personnel Chief and aides report to headcount room at the MECC.
2. Replay audio counts already received.
3. Record work unit head counts as they are received.
4. Obtain copy of visitors' log and employee with area head counts.
5. Monitor attempts to locate missing personnel.
6. Report to Emergency Response Coordinator as personnel are confirmed as missing.
7. Issue final report to Emergency Response Coordinator after all work unit head counts are received.
8. Determine from Emergency Response Coordinator names of any injured people.
9. Coordinate communications with families of those missing or injured.

Fig. 3.4 Personnel chief functional checklist

no matter how comprehensive the training programs may be, their effectiveness is unknown until a drill of the ERP is conducted and its systems and procedures are tested.

Facility Drills

To evaluate the effectiveness and completeness of an ERP, a facility must conduct periodic announced and unannounced Level Two and Level Three emergency drills. These are in addition to the more frequent Level One unit drills that are held to ensure that the more limited response procedures of a work unit are complete and well understood. Major internal emergency drills should be held at least four times each year, and scheduled to involve each working shift at least once a year.

To gain the maximum benefit from the drills, assigned observers should witness all aspects of response activity, and gather soon after to participate in a verbal evaluation of actions taken. These observers should include members of the Emergency Preparedness Team. This process is enhanced by capturing on-scene action on videotape or in still photographs. Critiquing can be extremely valuable in identifying necessary plan changes, training needs, and resource deficiencies of all kinds. The process of critiquing must be controlled and managed. The leader must be sure to identify the strengths as well as the weaknesses that were revealed in the drill. The objective is to stimulate actions for positive change, not to assign blame.

Drills should be held on weekends and at night occasionally to test segments of the organization that work at times when all of the specialized resources of the facility are not immediately available. Initially, limited scale drills can test segments of the ERP, such as manning the MECC with the shift personnel that are available and carrying out a headcount without involving daytime employees.

As the proficiency of the organization increases, drill scenarios can become more complex. Complicating factors approaching worst real-life conditions should be introduced periodically, including:

☐ Telephone switchboard overload
☐ Absence of key EMO members
☐ Arrival of major TV network anchorperson
☐ Simulated mass casualties
☐ Two-way radio failure
☐ Evacuation of primary and backup MECCs
☐ Major community impact

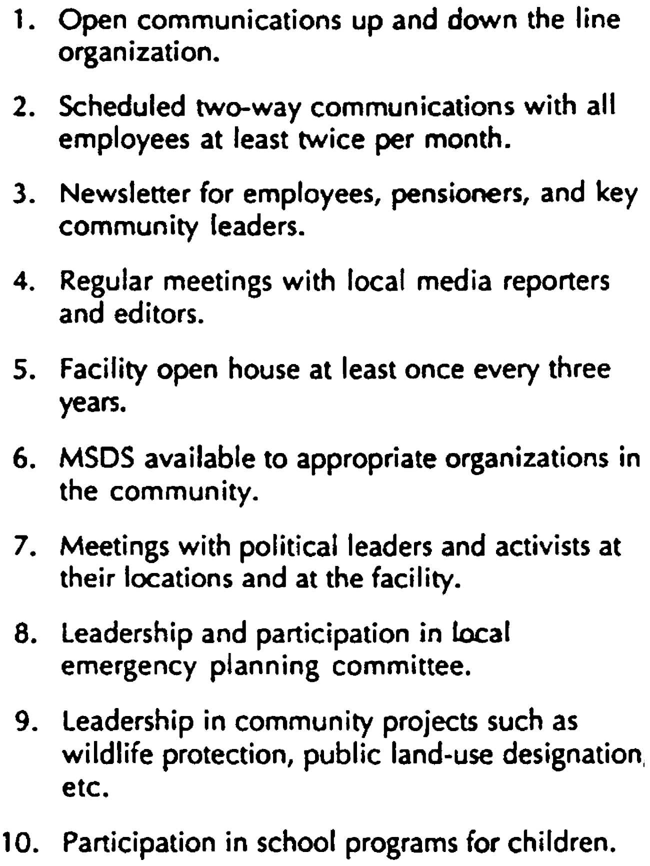
The Emergency Preparedness Team, with approval of the manager, designs the drill scenarios, monitors the organization's performance, provides leadership for critiquing, and recommends corrective actions. The team also develops and monitors a plan for involving the community in the important task of integrating a facility's ERP into the public emergency preparedness programs of the region.

Involving the Community

Of major concern to the chemical industry is the public perception that facility managers have little concern for the welfare of their neighbors. Managers themselves have contributed to such false impressions by failing to interact with their communities in a consistent and meaningful way.

The public clearly wants to know more about the risks presented by a chemical facility. Increasingly, the public wants to help decide which risks are acceptable and which are not.

Recognizing that industrial facilities exist only with the consent of their host communities, most industrial organizations are assigning a high priority to building



1. Open communications up and down the line organization.
2. Scheduled two-way communications with all employees at least twice per month.
3. Newsletter for employees, pensioners, and key community leaders.
4. Regular meetings with local media reporters and editors.
5. Facility open house at least once every three years.
6. MSDS available to appropriate organizations in the community.
7. Meetings with political leaders and activists at their locations and at the facility.
8. Leadership and participation in local emergency planning committee.
9. Leadership in community projects such as wildlife protection, public land-use designation, etc.
10. Participation in school programs for children.

Fig. 3.5 Risk communication checklist

stronger bridges with their neighbors. The chemical industry in particular is finding that the involvement of the community in the process of emergency preparedness planning presents an excellent opportunity for constructive two-way communications. This has proved particularly productive if these communications are part of an aggressive and continuous risk communications process.

Communications

Each chemical facility needs to create opportunities for sharing information with its surrounding community; it should not wait for these opportunities to occur. Facility managers in particular should be active in the process, and should monitor the performance of their organizations. Some items that should be on a manager's checklist for risk communications are shown in Fig. 3.5.

The process should start with employee communications. In its employees, a facility has an important, informal communication link with its community that is often ignored. Employees deserve to know at least as much about their facility as their neighbors, and they deserve to know it first. They should be aware through communications (if not through actual involvement) of the facility's entire emergency preparedness program, from prevention to preparation. The employees also should be familiar with the products made and their end uses. They should understand the potential hazards of the processes and materials with which they work, and how to protect themselves and the public from those hazards.

Other communication channels should be developed and regularly used. Scheduled meetings with representatives of community emergency service groups are useful for exchanging information and objectives concerning emergency preparedness, for promoting the sharing of resources, for gaining familiarity with one another's physical facilities and people, for identifying problems, and for recommending action for their solution.

In the 1980s, some chemical facilities and their communities formalized this approach and broadened participation in their meetings to include public officials and representatives of regulatory agencies and the media. Operating as Hazardous Material Advisory Councils (HMACs) and meeting regularly with established leadership and agendas, these groups quickly became key resources in organizing a community's efforts to better understand the potential for hazardous material incidents and to protect against them. Typically the responsibilities of an HMAC include coordinating a regional risk assessment, assisting the development of a community response plan specific to hazardous materials, and assisting with educational programs for various segments of the public.

The Responsible Care program of the American Chemistry Council embraces HMACs as one good way to communicate relevant and useful information that responds to public concerns for safety, health, and the environment. However, managers are finding many other ways to interact with the public to achieve a fuller measure of community awareness and involvement in affairs of mutual interest. Among the many options from which a manager may choose are the following:

- Hosting facility tours featuring emergency prevention and mitigation procedures
- Speaking at community meetings (service clubs, schools, governing bodies, etc.)
- Sending newsletters to selected neighbors
- Preparing informative brochures or newspaper inserts
- Appearing on local TV or radio

An open and sincere comprehensive risk communication process led by the facility manager creates a better-informed public that is able to understand real risks (vs. perceptions) and is likely to respond effectively in case of an actual emergency.

Integrating Plans

Most communities have long had written ERPs designed for natural events such as floods and windstorms; some communities have had written plans dealing with emergencies created by people, such as bomb threats and civil disturbances; but, until recently, few had specific plans for responding to emergencies involving hazardous materials. As a result, the consequences of accidental chemical releases have been in many cases tragically magnified by the undisciplined reactions of people near the release source. It has been reported, for example, that when the alarms sounded at Bhopal, residents of the nearby homes ran toward the plant rather than taking action to protect themselves from the enveloping fumes.

It is not enough to train the personnel of a chemical facility to implement an ERP effectively. Appropriate people in the community, especially near neighbors, need to understand the elements of an ERP that are designed to protect them and the role they play in making the plan work. There must be a continuous effort to integrate the facility ERP into community emergency planning at local and district levels. Drills involving external resources that test all the plan elements against the standards mutually established with the community and those imposed by laws and regulations are necessary to ensure successful implementation of the plan.

Off-Site Warning

Designing an effective off-site warning system presents some major challenges. Despite excellent ongoing communications between a chemical facility and its neighbors, there is no positive way to ensure that the general public will respond quickly and appropriately to a warning alarm of any kind. Thus, it is essential that the selected warning system be developed with the close cooperation of the community. Even then, it is difficult to predict such factors as the inclination of people to be warned and the degree of public confidence in the validity of an alarm.

No off-site warning system will assure complete coverage of the intended audience. Best results are achieved by combining two or more systems for sequential alerting—the first to trigger preplanned immediate action by the public at greatest risk, followed by other communications that provide further information and guidance to a larger audience. Some of the systems most commonly used are:

- ☐ Facility fixed-sound sources, such as sirens and whistles
- ☐ Mobile alerting by police or fire personnel, either from vehicles with loudspeakers or door-to-door
- ☐ Fixed public address systems in the community or in the facility
- ☐ Automated telephone calling
- ☐ Alert radios energized by a special signal to produce a warning tone followed by broadcast messages
- ☐ Strobe lights in situations where the noise level is a problem
- ☐ Local radio stations and the emergency broadcast system
- ☐ Local TV stations

More sophisticated and less commonly used warning systems include helicopters equipped with loudspeakers, modified cable TV installations, and computer networks between a chemical facility and community emergency response groups.

Local Emergency Plans

Existing plans for a coordinated response to emergencies in a community vary greatly in content and organization, but the plans have two common objectives. They are to:

- ☐ Define authority and responsibilities of various emergency service participants
- ☐ Describe the interaction between those participants, government, and industry

In creating their plans, most communities draw on the Integrated Emergency Management System (IEMS) developed by the Federal Emergency Management Agency (FEMA).

A local plan has many of the same elements as a chemical facility ERP. It includes:

- ☐ An EMO, with designated functional responsibilities
- ☐ The location of the emergency operating center and its resources
- ☐ Guidelines for classifying emergencies
- ☐ Activation and declaration checklists
- ☐ Communications requirements and available systems
- ☐ Evacuation and sheltering plans
- ☐ Methods for securing added resources
- ☐ Descriptions of local hazards

Most local plans are written to be nonspecific as to the cause of the emergency, with various appendices describing the details of response to specific events. These appendices are based on the results of risk assessments made by the community with the cooperation of industry.

One such appendix should relate to emergencies caused by fixed facility or transportation incidents in which hazardous materials are involved. Chemical facilities must provide substantial support to the community in preparing this portion of the local emergency plan, and provide resources and training leadership that are not available elsewhere in the community. Where a HMAC exists, there is an effective forum for doing this. In any case, a chemical facility manager should seek ways to help the community prepare for and recover from incidents of this nature.

Local Emergency Planning Committees

An important contribution to community and industry cooperation in emergency preparedness was the passage in 1986 of the SARA, which contained an emergency planning and community right-to-know provision. Title III, or EPCRA, as this portion of SARA is commonly called, is intended to encourage and support hazardous materials emergency planning efforts at the state and local level, and to provide citizens and local agencies with appropriate information concerning potential hazards in their communities.

The major portions of Title III require

☐ A statewide organization for planning emergency action and receiving hazardous chemical information
☐ Notification to the community of emergency releases of chemicals
☐ Reports of hazardous chemical inventories and copies of MSDSs to be furnished to the community
☐ An annual inventory of hazardous chemical releases to the environment

Drills and Critiques

The optimum frequency of major drills involving personnel outside a chemical facility is dependent upon a number of variables:

☐ Location of the facility
☐ Dependence upon community emergency agencies
☐ Size and complexity of the facility
☐ Site and off-site risk assessments
☐ Population patterns

An important element of emergency preparedness is the establishment of an appropriate major drill frequency in cooperation with off-site agencies. A reasonable goal is to hold one such drill each year. The scenario might include an on-site, internally generated hazardous material emergency 1 year and a transportation emergency somewhere in the adjacent community the following year.

For facilities and communities just beginning to test their plans, desktop or simulated drills are effective for identifying procedural problems that need to be corrected before they proceed to full-scale drills. In these simulations, staffing of the appropriate emergency center would occur, but the emergency service groups would not actually mobilize at the scene of the incident.

As people gain confidence in the completeness and the effectiveness of the ERP, it becomes important to measure the performance of all who are involved. Monitors record and later report on all aspects of response actions, including:

Elapsed times before critical actions occur
Actions and coordination of responding groups
Actions of uninvolved personnel
Alarm and communication effectiveness
Emergency control center management
Control at the emergency scene
Accounting for personnel
Medical aid for simulated casualties
- Off-site notifications

Handling media representatives—real or simulated

Following each drill there must be an organized critique that provides the information needed to strengthen the plan and/or its implementation. All the people actively involved should be represented at the critique, and a written report of conclusions and recommendations should be widely distributed. It then is the responsibility of the facility Emergency Preparedness Team and the local emergency planning committee (LEPC) to coordinate and assist in solving any problems identified—a process that begins emergency preparedness activities again: identifying hazards; evaluating and strengthening resources; modifying the emergency plan; training people; communicating and integrating plans; and testing them once again.

Laws, Regulations, and Support

Laws

A number of legal requirements must be incorporated in a facility's ERP. Emergency prevention, preparedness, and response planning are regulated at the federal, state, and, occasionally, local levels. At the federal level, these laws include

☐ Clean Air Act (CAA)
☐ Clean Water Act (CWA)
☐ Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA)
☐ Emergency Planning and Community Right-to-Know Act (EPCRA, or SARA Title III)
☐ Energy Reorganization Act (Nuclear Regulatory Commission)
☐ Hazardous Materials Transportation Act (HMTA)
☐ Occupational Safety and Health Act (OSHA)
☐ Resource Conservation and Recovery Act (RCRA)
☐ Toxic Substances Control Act (TSCA)
☐ U.S. Coast Guard requirements
☐ Environmental Protection Agency (EPA) regulations

A list of these laws and their *Code of Federal Regulations* (CFR) citations appears in the bibliography at the end of the chapter. States may have their own laws and regulations that also govern emergency response planning. References to these laws may be found in the Bureau of National Affairs (BNA) *Environment Reporter*.

These laws and their regulations are enforced by all levels of governmental agencies. A knowing or willful violation has serious implications for both companies and individuals, who may be held civilly or criminally liable for noncompliance. The penalties can be severe, ranging from daily-assessed fines to imprisonment. Thus, it is prudent to know the regulations that apply to the facility, and to ensure that the regulatory requirements are met.

The regulations regarding emergency planning and response are comprehensive, covering every aspect from prevention to reporting. RCRA's regulations cover the entire process, from planning to training to formal reports on the facility's response to an emergency involving hazardous waste. Other regulations focus on a specific aspect of emergency response, or part of the facility's operations. For example, the TSCA requires that spills or releases that contaminate the environment be reported orally and in writing within a certain time frame. All but one of the laws put the burden of planning on the facility. EPCRA, the exception, requires state and local agencies to prepare an ERP for the community. Facilities that meet criteria specified in EPCRA regulations have to assist in the development of the plan; however, they are not responsible for creating it.

The OSHA Process Safety Management regulation (29CFR 1910.119) and the EPA's Risk Management Plan regulation require significant attention to emergency planning and response. Inevitably, some of the regulations and requirements overlap. Most chemical facilities are subject to more than one law, and could be expected to prepare separate plans for specific parts of their facility. RCRA's Contingency Plan, for example, must be developed and maintained apart from other ERPs. The key to managing all the requirements and satisfying the regulations in an efficient, coordinated manner is first to understand the requirements and how they apply to the facility, and then to look for the common denominators among the requirements. The finished product, or master plan, will satisfy all the common denominators that apply, and will avoid duplication of effort. It also can be used as the basis for plans that must be maintained separately or that have requirements in addition to the common denominators.

Meeting the Requirements

Regulations governing emergency response planning can be broken down into four general categories:

☐ Preparation
☐ Plans
☐ Reports and other communications
☐ Drills and evaluations

Figure 3.6 shows a matrix-type summary of the major federal laws and their requirements for emergency planning and response. Such a matrix is very helpful in determining what the requirements are and how they apply to a facility. It could be further tailored to cover only the requirements that apply to a specific site or operation.

Prevention and Preparation

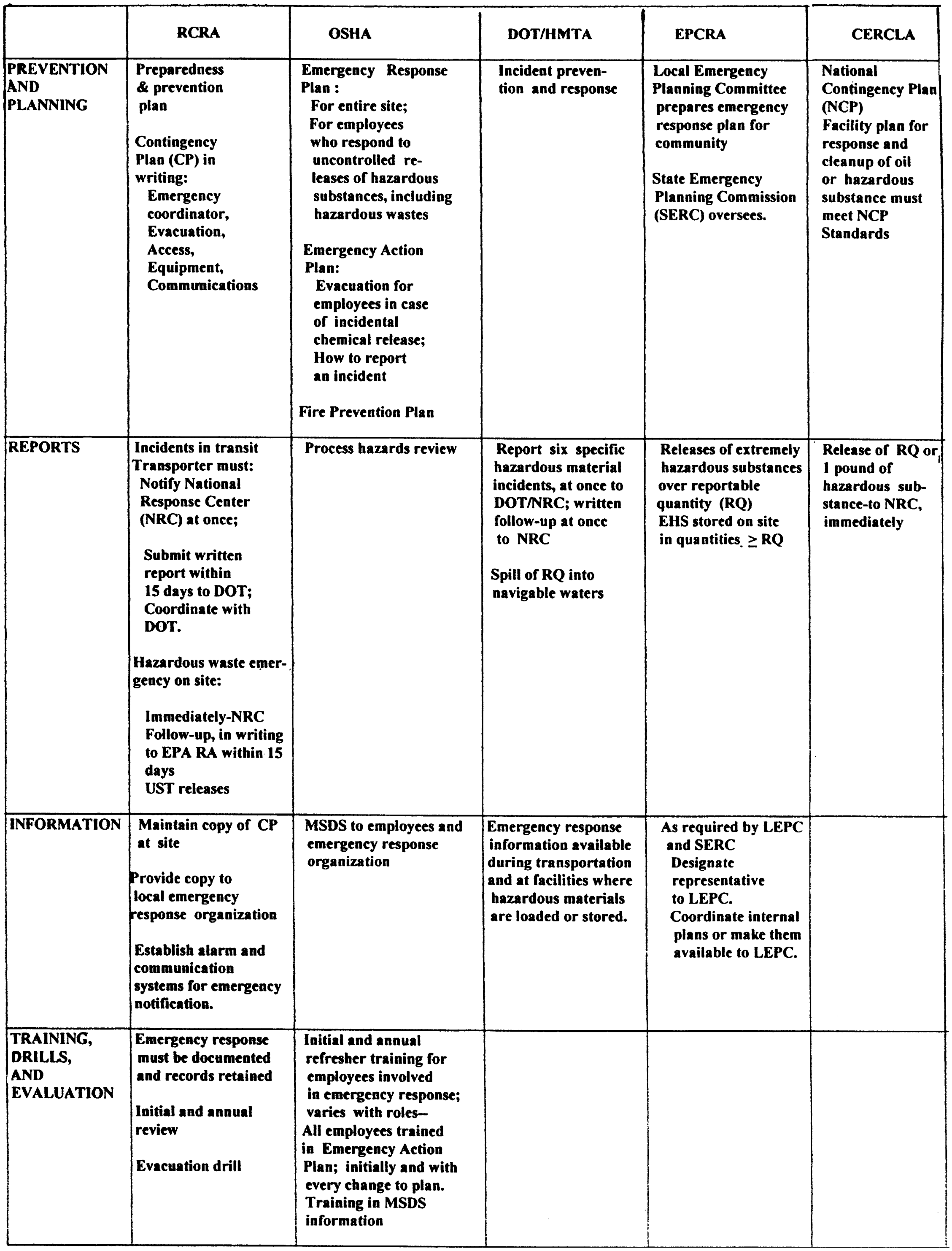
Some regulations require that a facility conduct a risk assessment and/or other preparatory activities. The RCRA calls it a preparedness and prevention plan. A facility subject to the RCRA must determine how structures, processes, and operations can be changed in order to minimize the possibility of an emergency involving hazardous waste. The facility also has to determine the communications and alarm systems that will be used in the event of such an emergency. The CWA includes prevention in its requirements for the Spill Prevention, Containment, and Countermeasure (SPCC) Plan. The 1990 amendments to the CAA added an accident prevention plan for extremely hazardous substances (EHSs).

Plans

At the very least, a facility is required to develop a plan describing how it will respond to an incident that threatens human health and/or the environment. Generally, the plan includes notification, evacuation, protection of employees, and control of the incident. This ERP usually must be in writing. For example, the OSHA requires a minimum of three plans: emergency response, emergency action, and fire prevention. The CAA requires that the state implementation plan have an emergency air pollution episode plan.

Communications

There are two aspects to emergency communications: the actual equipment used to communicate information about the incident and the types of communications or information-sharing required. The RCRA has specific requirements for the types of emergency communication equipment (alarm systems, phone or radio communications) that must be present. Under the EPCRA, facilities must provide information about their operations and substances used or stored on site when the Local Emergency Planning Committee (LEPC) or State Emergency Response Commission (SERC) requests it. If the facility uses or stores extremely hazardous substances (EHSs) in reportable quantities, it must appoint a representative to the LEPC.



	RCRA	OSHA	DOT/HMTA	EPCRA	CERCLA
PREVENTION AND PLANNING	Preparedness & prevention plan Contingency Plan (CP) in writing: Emergency coordinator, Evacuation, Access, Equipment, Communications	Emergency Response Plan : For entire site; For employees who respond to uncontrolled releases of hazardous substances, including hazardous wastes Emergency Action Plan: Evacuation for employees in case of incidental chemical release; How to report an incident Fire Prevention Plan	Incident prevention and response	Local Emergency Planning Committee prepares emergency response plan for community State Emergency Planning Commission (SERC) oversees.	National Contingency Plan (NCP) Facility plan for response and cleanup of oil or hazardous substance must meet NCP Standards
REPORTS	Incidents in transit Transporter must: Notify National Response Center (NRC) at once; Submit written report within 15 days to DOT; Coordinate with DOT. Hazardous waste emergency on site: Immediately-NRC Follow-up, in writing to EPA RA within 15 days UST releases	Process hazards review	Report six specific hazardous material incidents, at once to DOT/NRC; written follow-up at once to NRC Spill of RQ into navigable waters	Releases of extremely hazardous substances over reportable quantity (RQ) EHS stored on site in quantities ≥ RQ	Release of RQ or 1 pound of hazardous substance-to NRC, immediately
INFORMATION	Maintain copy of CP at site Provide copy to local emergency response organization Establish alarm and communication systems for emergency notification.	MSDS to employees and emergency response organization	Emergency response information available during transportation and at facilities where hazardous materials are loaded or stored.	As required by LEPC and SERC Designate representative to LEPC. Coordinate internal plans or make them available to LEPC.	
TRAINING, DRILLS, AND EVALUATION	Emergency response must be documented and records retained Initial and annual review Evacuation drill	Initial and annual refresher training for employees involved in emergency response; varies with roles– All employees trained in Emergency Action Plan; initially and with every change to plan. Training in MSDS information			

Fig. 3.6 Emergency prevention, planning, and response

CWA	CAA	TSCA	NUREG 054	USCG	EPA
Spill Prevention Control and Countermeasure (SPCC) Plan Shows how facility will: prevent, respond, follow-up to oil spills in harmful quantities. Must be in writing	State Iplementation Plan (SIP) must include an Emergency Air Pollution Episode plan Accident Prevention Program for EHS States may require prevention and emergency response plan		Emergency response plan, including: Emergency planning zones; Prevention, mitigation, and limitation of core damage and consequences of release Subject to annual review		Risk management plan
Release of RQ of oil or hazardous substance to NRC, immediately.	Announce uncontrolled releases of pollutants over certain set levels Releases ≥ RQ to NRC	Emergency incidents of environmental contamination At once to EPA, Written follow-up within 10 days	Notify state and local officials: Change in condition Protective action recommendations Notify community within 15 minutes. 100% notification within 15 miles.	All spills into navigable waters of oil, hazardous substances, ≥ RQ, at once, to US Coast Guard or to NRC	Worst case scenarios
Keep copy of SPCC at site available to EPA RA during normal working hours. Submit SPCC amendments to EPA RA and state water pollution control agency					Emergency response plan
Train: SPCC responders Employees who operate, maintain equipment			Annual training On-site Off-site Annual graded emergency response exercises		Training required for all employees in emergency response plan

Several laws require that a copy of the ERP be made available to employees and representatives of government agencies during working hours. The OSHA requires that facilities provide material safety data sheets (MSDS) for all hazardous substances present on the site.

Reports

Reports are another important communications aspect of the ERP. Most laws insist on prompt notification of the proper agencies immediately after an incident occurs. These reporting requirements can be complex, particularly in view of the fact that many laws have their own lists of hazardous substances and reportable quantities. The CERCLA requires that releases of a reportable quantity of what it defines as a hazardous substance must be reported immediately to the National Response Center (NRC). The EPCRA requires facilities that store and/or release reportable quantities of substances on its EHS list to report that information to the LEPC and the SERC. EPA's Risk Management Plan requires prior disclosure of possible "worst case" incident scenarios.

Written follow-up reports are often a requirement. The HMTA has identified six specific hazardous material incidents that must be reported immediately and again in writing. The RCRA gives managers 15 days after a hazardous waste emergency to submit a written report to the EPA.

Training, Drills, Audits, and Evaluations

An emergency plan is relatively useless unless the employees affected by it are trained in its use. The RCRA, OSHA, HMTA, CWA, and the Energy Reorganization Act require annual and refresher training. In addition, the facilities must keep records of the training, and must make them available to the appropriate agency when they are requested. Some laws go so far as to require practice drills. Nuclear power plants must conduct on- and off-site training, and go through annual graded emergency response exercises. The plan and the response executed according to the plan then are evaluated so that the plan can be improved.

Sources of Assistance

Seeing all the requirements together can be overwhelming. Fortunately, there are agencies, associations, and programs that can assist in the preparation of a comprehensive emergency prevention and response plan.

The federal government and the agencies responsible for the laws that govern emergency response planning provide 800-number hotlines and manuals that describe various aspects of emergency prevention, planning, and response.

The volunteer or professional emergency responders in the community have valuable practical experience that can be put to work in developing the facility ERP. Working with them also establishes a forum for communications and understanding with the community.

The ACC Community Awareness and Emergency Response (CAER) program provides comprehensive guidelines for the development and implementation of an ERP. The CAER program has been expanded to include all aspects of the chemical industry in an initiative called Responsible Care.

Other services of the ACC include CHEMTREC, a 24-h emergency response service for people who respond to emergencies involving chemicals; CHEMNET, a mutual aid agreement between chemical producers and emergency response contractors; and workshops and videotape training programs for first responders and other emergency response personnel.

Additional sources of assistance and information include other professional associations, such as the American Institute of Chemical Engineers, and publications, seminars, workshops, and videotapes offered by educational organizations. Considerable information is available on the Internet through web sites such as www.fema.gov which is maintained by the FEMA.

A bibliography; a list of laws, regulations, and standards; and a compilation of suggested reading material follow.

Select Bibliography

American Chemistry Council (Formerly Chemical Manufacturers Association) (1989) CAER: the next phase program handbook. Chemical Manufacturers Association (now American Chemistry Council), Washington, DC, with assistance from HMM Associates, Concord, MA

American Chemistry Council (Formerly Chemical Manufacturers Association) (1986) Community emergency response exercise handbook. Chemical Manufacturers Association (now American Chemistry Council), Washington, DC

American Chemistry Council (Formerly Chemical Manufacturers Association) (1987) Emergency warning systems guidebook. Chemical Manufacturers Association (American Chemistry Council), Washington, DC

American Chemistry Council (Formerly Chemical Manufacturers Association) (1986) Site emergency response planning handbook. Chemical Manufacturers Association (now American Chemistry Council), Washington, DC

The Conservation Foundation (1986) Risk Communication: Proceedings of the National Conference on Risk Communication. The Conservation Foundation, Washington, DC

Covello VT, Sandman PM, Slovic P (1988) Risk communication, risk statistics, and risk comparisons. Chemical Manufacturers Association (American Chemistry Council), Washington, DC

Federal Emergency Management Agency (1985) FEMA handbook. Federal Emergency Management Agency, Washington, DC

National Response Team (1987) Hazardous materials emergency planning guide (NRT 1). National Response Team of the National Oil and Hazardous Substances Contingency Plan. Washington, DC

National Response Team (1988) Criteria for review of hazardous materials emergency plans. National Response Team of the National Oil and Hazardous Substances Contingency Plan. Washington, DC

Occupational Safety and Health Administration (1988) How to prepare for workplace emergencies, OSHA 3088 (revised). U.S. Department of Labor, Washington, DC

Prugh RW, Johnson RW (1988) Guidelines for vapor release mitigation. Center for Chemical Process Safety, The American Institute of Chemical Engineers, New York

SARA Title III (Superfund Amendment and Reauthorization Act, Emergency Planning and Community Right-to-Know) 4OCFR 355.3, Section 302(c), Emergency planning notification Section 303(d), Appointment of emergency coordinator, provision of information; Section 304, Emergency release notification requirements

EPA 40 CFR Part 68, Risk Management Plan

U.S. Coast Guard 33CFR 126.9 Reporting requirements for discharge of petroleum products or dangerous liquid commodities into navigable waters of USA 33CFR 153.023 Reporting requirements for discharge of reportable quantity (RQ) of oil or hazardous substance into navigable waters

Standards

ANSI/National Fire Protection Association Standards # 72 National Fire Alarm Code # 110 Emergency Power and Standby Systems # 600 Private Fire Brigades # 1561 Emergency Services Incident Management System

Suggested Reading: Regulations

Bureau of National Affairs. Environment Reporter. Bureau of National Affairs, Inc., Washington, DC

ENSR Consulting and Engineering (1988) Air quality handbook: a guide to permitting and compliance under the clean air act and air toxics programs, 10th edn. ENSR Consulting and Engineering, Acton, MA

ERT, Inc., and Sidley & Austin (1987) Superfund handbook, 2nd edn. ERT, Inc. and Chicago, IL: Sidley & Austin, Concord, MA

Lowry GG, Robert C (1988) Lowry's handbook of right-to-know and emergency planning. Lewis Publishers, Inc., Chelsea, MI

Office of Solid Waste, U.S. Environmental Protection Agency (1986) Solving the Hazardous Waste Problem: EPA's RCRA Program (EPA/530-SW-86-037). U.S. Environmental Protection Agency, Washington, DC

Prevention and Planning

Abrams MJ, Lewis J (1987) Preplanning, the key emergency response. Papers Presented at the Spring 1987 National Meeting of the American Institute of Chemical Engineers. American Institute of Chemical Engineers, New York

Adler V, Sorenson JH, Rogers GO (1989) Chemical and nuclear emergencies: interchanging lessons learned from planning and accident experience. Proceedings of a Joint NEA/CEC Workshop on Technical Aspects of Emergency Planning, Brussels, Belgium (CONF-89906137-1). U.S. Department of Energy, Washington, DC

U.S. Environmental Protection Agency (1988) Guide to exercises in chemical emergency preparedness programs. U.S. Environmental Protection Agency, Washington, DC

Laws, Regulations, and Standards: Laws and Regulations

CAA (Clean Air Act) 4OCFR 51 SIP Emergency air pollution episode plan CAA Reauthorization Section 304, Prevention of sudden, accidental releases CERCLA (Comprehensive Environmental Response, Compensation, and Liability Act)

4OCFR 300–306 Section 103(a) Spill reporting requirements 4OCFR 355 Emergency planning CWA (Clean Water Act) 4OCFR 112.3-7 Spill prevention, control, and countermeasure plan (SPCC) requirements, modifications, preparation, and implementation Energy Reorganization Act (was Atomic Energy Act) 1OCFR 50.47, also Appendix E NRC (Nuclear Regulatory Commission) Standards for on-site and off-site emergency plans HMTA (Hazardous Materials Transportation Act) 49CFR Parts 171–177 DOT requirements for providing information and advice on meeting emergencies; FEMA requirements for evaluation of training programs for incident prevention and response

OSHA (Occupational Safety and Health Act) 29CFR 1910.1200 Hazard communication standard 29CFR 1910–210 and Appendices, Hazardous waste operations and emergency response 29CFR 1910.35-38 Requirements for evacuation route and exit posting, emergency lighting, accessibility of exits, and emergency action plan: 29CFR 1910.156 Fire brigades: 29CFR 1910.157-164 Fire extinguishing and detection systems: 29CFR 1910.165 Employee alarm systems: 29CFR 1910.119 Process Safety Management of Highly Hazardous Chemicals

RCRA (Resource Conservation and Recovery Act) 4OCFR 263.30,31 Transporter responsibilities in hazardous waste transportation incidents 4OCFR 264.30-37 and 4OCFR 265.30-37 Preparedness and prevention 4OCFR 264.50-56 and 4OCFR 265.50-56 Contingency Plan

American Nuclear Society (1988) Emergency Response—Planning, Technologies, and Implementation; Proceedings of the ANS Topical Meeting (CONF-880913. UC-610). Savannah River Laboratory, E. I. du Pont de Nemours and Company, Aiken, SC

Bell DW, Burns CC (1987) Offsite emergency plan development and maintenance considerations. Papers Presented at the Spring 1987 National Meeting of the American Institute of Chemical Engineers. American Institute of Chemical Engineers, New York

Davis DS et al (1988) Prevention reference manual: overviews on preventing and controlling accidental releases of selected toxic chemicals (EPA/600/8-88-074). U.S. Environmental Protection Agency, Air and Energy Engineering Research Lab, Research Triangle Park, NC

Dickerson MH (1986) Emergency planning, response, and assessment: a concept for a center of excellence. Proceedings from an International Seminar on Nuclear War International Cooperation: The Alternatives (CONF-8608149-1). U.S. Department of Energy, Washington, DC

Federal Emergency Management Agency (1989) Hazardous materials exercise evaluation methodology (HM-EEM) and manual. Federal Emergency Management Agency, Washington, DC

Fingleton DJ, Tanzman EZ, Bertram KM (1986) Development of a model emergency response plan for catastrophic releases of toxic gases. Proceedings of the Air Pollution Control Association Annual Meeting (CONF-860606-19). U.S. Department of Energy, Washington, DC

Gudiksen P et al (1986) Emergency response planning for potential accidental liquid chlorine releases. U.S. Department of Energy, Washington, DC

International Association of Fire Chiefs (1985) Fire service emergency management handbook. Federal Emergency Management Agency, Washington, DC

Jones E (1987) Contingency planning and emergency response in construction activities: training the construction worker. Proceedings of an Oak Ridge Model Conference on Waste Problems (CONF-871075-5). U.S. Department of Energy, Washington, DC

Kalnins RV (1986) Emergency preparedness and response. Symposium on the Characterization of Thermodynamic and Transport Properties of Polymer Systems. American Institute of Chemical Engineers, New York

Krimsky S, Plough A (1988) Environmental hazards—communicating risks as a social process. Auburn House Publishing Company, Medford, MA

Michael EJ et al (1986) Emergency planning considerations for specialty chemical plants. American Institute of Chemical Engineers Summer National Meeting. American Institute of Chemical Engineers, New York

Morentz JW, Griffith D (1987) Using computers for chemical emergency planning and response. Papers Presented at the Spring 1987 National Meeting of the American Institute of Chemical Engineers. American Institute of Chemical Engineers, New York

National Response Team (1988) Criteria for review of hazardous materials emergency plans (NRT-1A). National Response Team of the National Oil and Hazardous Substances Contingency Plan, Washington, DC

National Response Team (1987) HAZMAT planning guide (WH-562A). National Response Team, Washington, DC

National Response Team (1987) Technical guidance for hazards analysis: emergency planning for extremely hazardous substances (Supplement to Hazardous Materials Emergency Planning Guide NRT 1). National Response Team, Washington, DC

Philley JO (1987) Emergency preparedness training tips. American Institute of Chemical Engineers National Meeting—Summer'87. American Institute of Chemical Engineers, New York

U.S. Environmental Protection Agency (1986) Bibliography on chemical emergency preparedness and prevention. U.S. Environmental Protection Agency, Washington, DC

U.S. Environmental Protection Agency (1988) It's not over in October: a guide for local emergency planning committees. U.S. Environmental Protection Agency, Washington, DC

U.S. Environmental Protection Agency (1988) Seven cardinal rules of risk communication (EPA 87–020). U.S. Environmental Protection Agency, Washington, DC

U.S. Environmental Protection Agency (1989) Why accidents occur: insights from the accidental release information program. Chemical Accident Prevention Bulletin (OSWER-89-008.1), Series 8, No. 1 (July 1989). U.S. Environmental Protection Agency, Washington, DC

Waldo AB (ed) (1986) The community right-to-know handbook. Thompson Publishing Group, Washington, DC

Transportation

Office of Hazardous Materials Transportation, U.S. Department of Transportation. Emergency Response Guidebook. Available through the U.S. Government Printing Office (GPO) U.S. Government Bookstores. Updated every three years

Applied Statistical Methods and the Chemical Industry

4

Stephen Vardeman and Robert Kasprzyk

Introduction

The discipline of statistics is the study of effective methods of data collection, data summarization, and (data based, quantitative) inference making in a framework that explicitly recognizes the reality of nonnegligible variation in real-world processes and measurements.

The ultimate goal of the field is to provide tools for extracting the maximum amount of useful information about a noisy physical process from a given investment of data collection and analysis resources. It is clear that such a goal is relevant to the practice of Industrial chemistry. The primary purposes of this chapter are to indicate in concrete terms the nature of some existing methods of applied statistics that are particularly appropriate to industrial chemistry, and to provide an entry into the statistical literature for those readers who find, in the discussion here, reasons to believe that statistical tools can help them be effective in their work.

This chapter will begin with some simple ideas of modern descriptive statistics, including numerical and graphical data summarization tools, and the notions of fitting equations to data and using theoretical distributions. Next, some tools for routine industrial process monitoring and capability assessment, concentrating primarily on the notion of control charting, will be presented. This will be followed by a more extensive discussion of common statistical data collection strategies and data analysis methods for multifactor experimental situations met in both laboratory and production environments. This section will touch on ideas of partitioning observed variation in a system response to various sources thought to influence the response, factorial and fractional factorial experimental designs, sequential experimental strategy, screening experiments, and response surface fitting and representation. Next come brief discussions of two types of special statistical tools associated specifically with chemical applications, namely, mixture techniques and nonlinear mechanistic model building. A short exposition of chemical industry implications of relationships between modern business process improvement programs and the discipline of statistics follows. The chapter concludes with a reference section listing sources for further reading.

Simple Tools of Descriptive Statistics

There are a variety of data summarization or description methods whose purpose is to make evident the main features of a data set. (Their use, of course, may be independent of whether or not the data collection process actually employed was in any sense a "good" one.) To illustrate some of the simplest of these methods, consider the data listed in Table 4.1. These numbers represent aluminum impurity contents (in ppm) of 26 bihourly samples of recycled PET plastic recovered at a Rutgers University recycling pilot plant.

A simple plot of aluminum content against time order, often called a *run chart*, is a natural place to begin looking for any story carried by a data set. Figure 4.1 shows such a plot for the data of Table 4.1, and in this case reveals only one potentially interesting feature of the data. That is, there is perhaps a weak hint of a downward trend in the aluminum contents that might well have been of interest to the original researchers. (If indeed the possible slight decline in aluminum contents is more than "random scatter," knowledge of its physical origin, whether in actual composition of recycled material or in the measurement process, presumably would have been helpful to the effective running of the recycling facility. We will save a discussion of tools for rationally deciding whether there is more than random scatter in a plot like Fig. 4.1 until the next section).

S. Vardeman (✉)
Departments of Statistics and of Industrial Engineering and Manufacturing Systems Engineering, Iowa State University, Ames, IA, USA
e-mail: vardeman@iastate.edu

R. Kasprzyk
Dow Chemical Company

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_4, © Springer Science+Business Media New York 2012

Fig. 4.1 A run chart for 26 consecutive aluminum contents

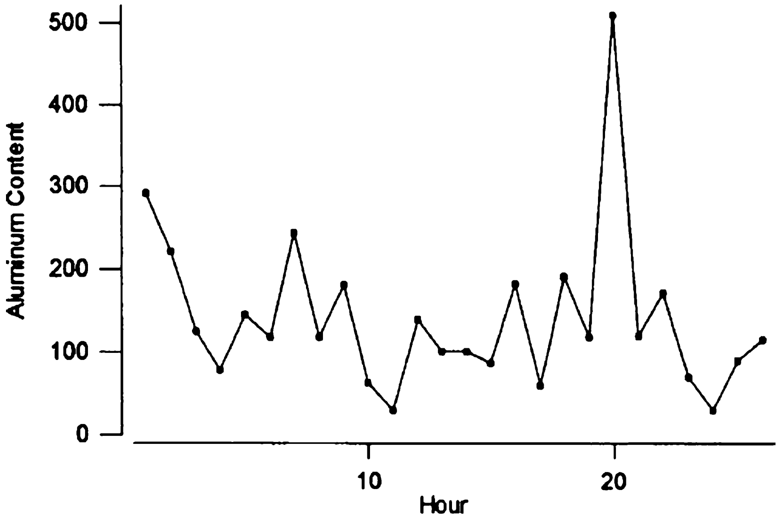
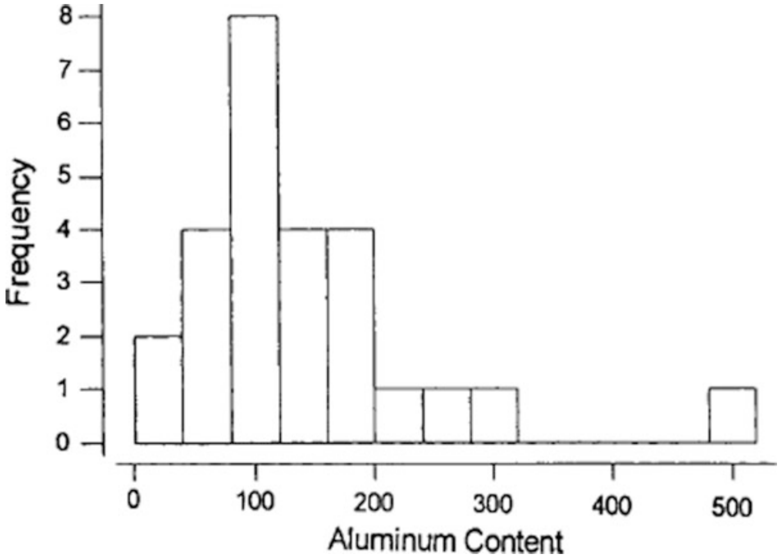


Fig. 4.2 A histogram for 26 aluminum contents



The run chart is a simple, explicitly dynamic tool of descriptive statistics. In those cases where one decides that there is in fact little information in the time order corresponding to a data set, there are a variety of simple, essentially static, statistical tools that can be used in describing the pattern of variation in a data set. Figures 4.2, 4.3, 4.4 and 4.5 show graphical representations of the data of Table 4.1 in, respectively, *histogram*, *stem and leaf plot*, *dot plot*, and *box plot* forms.

The histogram/bar chart idea of Fig. 4.2 is likely familiar to most readers, being readily available, for example, through the use of commercial spreadsheet software. It shows how data are spread out or distributed across the range of values represented, tall bars indicating high frequency or density of data in the interval covered by the base of the bar. Figure 4.2 shows the measured aluminum contents to be somewhat asymmetrically distributed (statistical jargon is that the distribution is "skewed right"), with a "central" value perhaps somewhere in the vicinity of 120 ppm.

Histograms are commonly and effectively used for final data presentation, but as working data analysis tools they suffer from several limitations. In one direction, their appearance is fairly sensitive to the data grouping done to make them, and it is usually not possible to recover from a histogram the exact data values used to produce it, should

0	30,30
0	60,63,70,79,87,90
1	01,02,15,18,19,19,20,25,40,45
1	72,82,83,91
2	22,44
2	91
3	
3	
4	
4	
5	11

Fig. 4.3 A stem and leaf plot for 26 aluminum contents

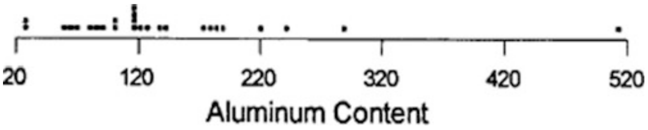


Fig. 4.4 A dot plot for 26 aluminum contents

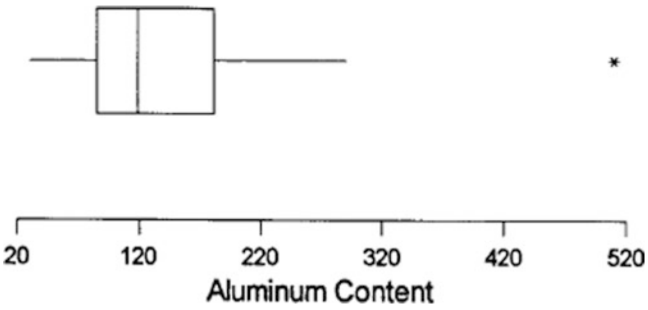


Fig. 4.5 A box plot for aluminum contents

Table 4.1 Twenty-six consecutive aluminum contents (ppm)[a]

291	222	125	79	145	119	224	118	182	63	30	140	101
102	87	183	60	191	119	511	120	172	70	30	90	115

[a] Based on data in Albin [32]

one wish to try other groupings. In another direction, histograms are somewhat unwieldy, for example, not being particularly suitable to the comparison of, say, 10 or 12 data sets on a single page. The graphical devices of Figs. 4.3, 4.4, and 4.5 are less common than the histogram, but address some of these shortcomings.

The stem and leaf diagram of Fig. 4.3 and the dot plot of Fig. 4.4 carry shape information about the distribution of aluminum contents in a manner very similar to the histogram of Fig. 4.2. But the stem and leaf and dot diagrams do so without losing the exact identities of the individual data points. The box plot of Fig. 4.5 represents the "middle half" of the data with a box divided at the 50th percentile (or in statistical jargon, the median) of the data, and then uses so-called whiskers to indicate how far the most extreme data points are from the middle half of the data.

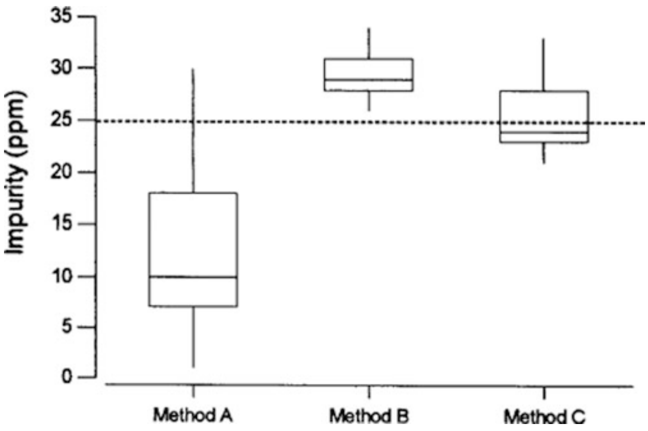


Fig. 4.6 Side-by-side box plots for three laboratory test methods

Box plots preserve much of the shape information available from the other displays (e.g., portraying lack of symmetry through differing sizes of box "halves" and/or whisker lengths), but do so in a way that is conducive to simultaneous representation and comparison of many data sets on a single graphic, through the placement of box plots side by side. Figure 4.6 illustrates this point with a graphical comparison of three laboratory test methods to a standard.

A total of 90 samples of a stock solution known to contain 25 ppm of an impurity were analyzed by a single lab team using three different test methods (30 of the samples being allocated to each of the three methods), and the box plots in Fig. 4.6 portray the measured impurity levels for the different methods. The figure shows quite effectively that Method A is neither precise nor accurate, Method B is quite precise but not accurate, and Method C is somewhat less precise than B but is accurate. This kind of knowledge can form the basis of an informed choice of method.

Figures 4.2, 4.3, 4.4, 4.5, and 4.6 give only a hint of the spectrum of tools of statistical graphics that are potentially helpful in data analysis for Industrial chemistry. For more details and much additional reading on the subject of modern statistical graphics, the reader is referred to the book by Chambers et al. [1] listed in the references section.

Complementary to graphical data summaries are *numerical summarizations*. For the simple case of data collected under a single set of conditions, the most commonly used measures deal with the location/center of the data set and the variability/spread of the data. The (*arithmetic*) *mean* and the *median* are the most popular measures of location, and the *variance* and its square root, the *standard deviation*, are the most widely used measures of internal variability in a data set.

For n data values $y_1, y_2, \ldots, y_n$ the median is

$$\tilde{y} = \text{the ``middle'' of } \frac{n+1}{2} \text{th ordered data value} \tag{4.1}$$

and the mean is

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{4.2}$$

The reader is invited to check that upon ordering the $n = 26$ values in Table 4.1, the 13th smallest value is 119 and the 14th smallest value is also 119, so that the only sensible interpretation of (4.1) for the aluminum content data is that

$$\tilde{y} = \text{the } 13.5\text{th ordered data value} = \frac{119 + 119}{2} = 119 \text{ ppm}$$

On the other hand, from (4.2) the mean of the aluminum contents is

$$\begin{aligned}\bar{y} &= \frac{1}{26}(291 + 222 + 125 + \cdots + 30 + 90 + 115)\\ &\approx 142.7 \text{ ppm}\end{aligned}$$

The median and mean are clearly different measures of location/center. The former is in the middle of the data in the sense that about half of the data are larger and about half are smaller. The latter is a kind of "center of mass," and for asymmetrical data sets like that of Table 4.1 is usually pulled from the median in the direction of any "skew" present, that is, is pulled in the direction of "extreme" values.

The variance of n data values $y_1, y_2, \ldots, y_n$ is essentially a mean squared deviation of the data points from their mean. In precise terms, the variance is

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 \tag{4.3}$$

and the so-called standard deviation is

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{4.4}$$

For the example of the aluminum contents, it is elementary to verify that

$$\begin{aligned}s^2 &\approx \frac{1}{26-1}[(291 - 142.7)^2 + (222 - 142.7)^2 + \cdots\\ &\quad + (115 - 142.7)^2]\\ &\approx 9{,}644(\text{ppm})^2\end{aligned}$$

so that

$$s = \sqrt{s^2} \approx 98.2 \text{ ppm}$$

An appropriate interpretation of s is not completely obvious at this point, but it does turn out to measure the spread of a data set, and to be extremely useful in drawing quantitative inferences from data. (In many, but not all, circumstances met in practice, the *range* or largest value in a data set minus the smallest value is on the order of four to six times s.) The variance and standard deviation are time-honored and fundamental quantifications of the variation present in a single group of measurements and, by implication, the data-generating process that produced them.

When data are collected under several different sets of conditions, and those conditions can be expressed in quantitative terms, effective data summarization often takes the form *of fitting an approximate equation* to the data. As the basis of a simple example of this, consider the data in Table 4.2. The variable x, hydrocarbon liquid hourly space velocity, specifies the conditions under which information on the response variable y, a measure of isobutylene conversion, was obtained in a study involving the direct hydration of olefins.

Table 4.2 Seven liquid hourly space velocity/mol% conversion data Pairs[a]

Liquid hourly space velocity, x	Mole % isobutylene conversion, y
1	23.0, 24.5
2	28.0
4	30.9, 32.0, 33.6
6	20.0

[a] Based on a graph in Odioso et al. [33]

For purposes of economy of expression, and perhaps some cautious interpolation between values of x not included in the original dataset, one might well like to fit a simple equation involving some parameters $\underline{b}$, say,

$$y \approx f(x|\underline{b}.) \tag{4.5}$$

to the data of Table 4.2. The simplest possible form for the function $f(x|\underline{b})$ that accords with the "up then back down again" nature of the conversion values y in Table 4.2 is the quadratic form

$$f(x|\underline{b}) = b_0 + b_1 x + b_2 x^2 \tag{4.6}$$

and a convenient method of fitting such an equation (that is linear in the parameters b) is *the method of least squares*. That is, to fit a parabola through a plot of the seven (x, y)

pairs specified in Table 4.2, it is convenient to choose b_0, b_1, and b_2 to minimize the sum of squared differences between the observed conversion values y and the corresponding fitted values y on the parabola. In symbols, the least squares fitting of the approximate relationship specified by (4.5) and (4.6) to the data of Table 4.2 proceeds by minimization of

$$\sum_{i=1}^{7} \left[y_i - \left(b_0 + b_1 x_i + b_2 {x_i}^2\right)\right]^2$$

over choices of the coefficients b. As it turns out, use of standard statistical "regression analysis" software shows that the fitting process for this example produces the approximate relationship

$$y \approx 13.64 + 11.41x - 1.72x^2$$

and Fig. 4.7 shows the fitted (summarizing) parabola sketched on the same set of axes used to plot the seven data points of Table 4.2.

The least squares fitting of approximate functional relationships to data with even multidimensional explanatory variable x typically goes under the (unfortunately obscure) name of *multiple regression* analysis, and is given an introductory treatment in most engineering statistics textbooks, including, for example, the ones by Devore [2], Vardeman and Jobe [3], and Vardeman [4] listed in the references. A lucid and rather complete treatment of the subject can also be found in the book by Neter et al. [5].

A final notion that we wish to treat in this section on descriptive statistics is that of representing a distribution of responses and/or the mechanism that produced them (under a single set of physical conditions) by a *theoretical distribution*. That is, there are a number of convenient theoretical distributional shapes, and it is often possible to achieve great economy of expression and thought by seeing in a graphical representation such as Figs. 4.2, 4.3, 4.4, and 4.5 the possibility of henceforth describing the phenomenon portrayed via some one of those theoretical distributions. Here we will concentrate on only the most commonly used theoretical distribution, the so-called *Gaussian* or *normal* distribution.

Figure 4.8 is a graph of the function of x

$$g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad (4.7)$$

where $g(x)$ specifies the archetypical "bell-shaped curve" centered at the number μ, with spread controlled by the number σ (and is in fact usually called the Gaussian probability density with mean μ and standard deviation σ).

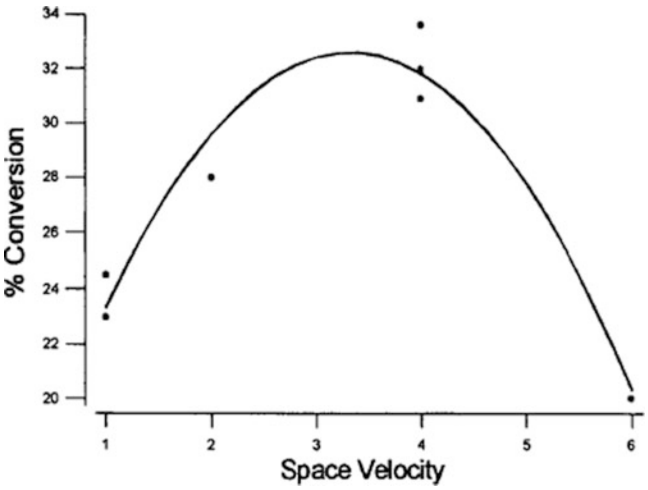


Fig. 4.7 A scatter plot of seven space velocity/mol% conversion data pairs and a fitted parabola

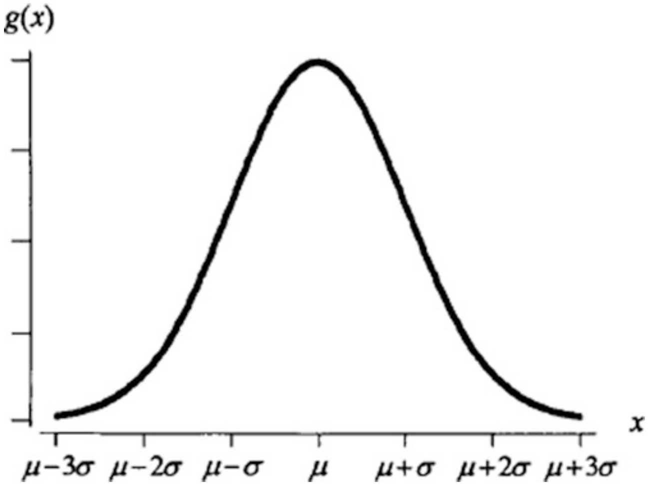


Fig. 4.8 The Gaussian probability density with mean μ and standard deviation σ

Figure 4.8 can be thought of as a kind of idealized histogram. Just as fractional areas enclosed by particular bars of a histogram correspond to fractions of a data set with values in the intervals represented by those bars, areas under the curve specified in (4.7) above particular intervals might be thought of as corresponding to fractions of potential data points having values in those intervals. (It is possible to show that the total area under the curve represented in Fig. 4.8, namely, $\int_{-\infty}^{\infty} g(x)\,\mathrm{d}x$, is 1.) Simple tabular methods presented in every elementary statistics book avoid the need to regularly use numerical integration in evaluating such areas. These methods can, for example, be used to show that roughly 68% of a Gaussian distribution lies between $\mu-\sigma$ and $\mu+\sigma$, roughly 95% lies between $\mu-2\sigma$ and $\mu+2\sigma$, and roughly 99.7% lies between $\mu-3\sigma$

Table 4.3 Twenty-six logarithms of aluminum contents

5.67	5.40	4.83	4.37	4.98	4.78	5.50	4.77	5.20	4.14	3.40	4.94	4.62
4.62	4.47	5.21	4.09	5.25	4.78	6.24	4.79	5.15	4.25	3.40	4.50	4.74

```
3 | .40,.40
3 |
4 | .09,.14,.25,.37,.47
4 | .50,.62,.62,.74,.77,.78,.79,.79,.83,.94,.98
5 | .15,.20,.21,.25,.40
5 | .50,.67
6 | .24
```

Fig. 4.9 A stem and leaf plot for the logarithms of 26 aluminum contents

and $\mu + 3\sigma$. Part of the convenience provided when one can treat a data-generating process as approximately Gaussian is that, given only a theoretical mean μ and theoretical standard deviation σ, predictions of fractions of future data values likely to fall in intervals of interest are thus easy to obtain.

At this point let us return to the aluminum content data of Table 4.1. The skewed shape that is evident in all of Figs. 4.2, 4.3, 4.4, and 4.5 makes a Gaussian distribution inappropriate as a theoretical model for (raw) aluminum content of such PET samples. But as is often the case with right skewed data, considering the *logarithms* of the original measurement creates a scale where a normal distribution is more plausible as a representation of the phenomenon under study. Thus, Table 4.3 contains the natural logs of the values in Table 4.1, and the corresponding stem and leaf plot in Fig. 4.9 shows the transformed data to be much more symmetrically distributed than the original data. The possibility opened up by this kind of transformation idea is one using statistical methods based on the normal distribution to reach conclusions about $\ln y$ and then simply exponentiating to derive conclusions about the original response y itself. The applicability of statistical methods developed for normal distributions is thereby significantly broadened.

In addition to providing convenient conceptual summarizations of the nature of response distributions, theoretical distributions such as the normal distribution form the mathematical underpinnings of methods of formal quantitative *statistical inference*. It is outside our purposes in this chapter to provide a complete introduction to such methods, but thorough and readable accounts are available in engineering statistics books such as those of Devore [2] and Vardeman and Jobe [3]. Here, we will simply say that, working with a Gaussian description of a response, it is possible to quantify in various ways how much information is carried by data sets of various sizes. For instance, if a normal distribution describes a response variable y, then in a certain well-defined sense, based on $n = 26$ observations producing a mean $\bar{y}$ and a standard deviation s, the interval with end points

$$\bar{y} - 2.060s\sqrt{1 + \frac{1}{26}} \tag{4.8}$$

and

$$\bar{y} + 2.060s\sqrt{1 + \frac{1}{26}} \tag{4.8a}$$

has a 95% chance of predicting the value of an additional observation. For instance, applying formula (4.8) to the log values in Table 4.3, the conclusion is that the interval from 3.45 to 6.10 ln(ppm) has a 95% chance of bracketing an additional log aluminum content produced (under the physical conditions of the original study) at the recycling plant. Exponentiating, the corresponding statement about raw aluminum content is that the interval from 31 to 446 ppm has a 95% chance of bracketing an additional aluminum content. Methods of statistical inference like that represented in (4.8) are called *prediction interval* methods. The book by Hahn and Meeker [6] provides a thorough discussion of such methods, based not only on the Gaussian distribution but also on other theoretical distributional shapes as well.

Tools of Routine Industrial Process Monitoring and Capability Assessment

Probably the two most basic generic industrial problems commonly approached using statistical methods are those of (1) monitoring and maintaining the stability/consistency of a process and (2) assessing the capability of a stable process. This section provides a brief introduction to the use of tools of "control" charting in these enterprises.

Working at Bell Labs during the 1920s and 1930s, Walter Shewhart developed the notion of routinely plotting data from an industrial process in a form that allows one to separate observed variability in a response into two kinds of variation. The first is that variation which appears to be inherent, unavoidable, short-term, baseline, and characteristic of the process (at least as currently configured). This variation Shewhart called *random* or *common cause variation*. The second kind of variability is that variation which appears to be avoidable, long-term, and/or due to sources outside of those seen as legitimately impacting process behavior. He called this variation *assignable* or *special cause variation*.

Shewhart reasoned that by plotting summary statistics from periodically collected data sets against time order of collection, one would be able to see interpretable trends or other evidence of assignable variation on the plots, and could intervene to eliminate the physical causes of that variation. The intention was to thereby make process output stable or consistent to within the inherent limits of process precision. As a means of differentiating plotted values that should signal the need for intervention from those that carry no special message of process distress, he suggested drawing so-called control limits on the plots. (The word "control" is something of a misnomer, at least as compared to common modern engineering usage of the word in referring to the active, moment-by-moment steering or regulation of processes. The nonstandard and more passive terminology "monitoring limits" would actually be far more descriptive of the purpose of Shewhart's limits.) These limits were to separate plausible values of the plotted statistic from implausible values when in fact the process was operating optimally, subject only to causes of variation that were part of standard conditions.

By far the most famous implementations of Shewhart's basic logic come where the plotted statistic is either the mean, the range, or less frequently, the standard deviation. Such charts are commonly known by the names *x-bar charts*, *R charts*, and *s charts*, respectively. As a basis of discussion of Shewhart charts, consider the data given in Table 4.4. These values represent melt index measurements of specimens of extrusion grade polyethylene, taken four per shift in a plastics plant.

Figure 4.10 shows plots of the individual melt indices, means, ranges, and standard deviations from Table 4.4

Table 4.4 Measured melt indices for ten groups of four specimens[a]

Shift	Melt index	$\bar{y}$	R	S
1	218, 224, 220, 231	223.25	13	5.74
2	228, 236, 247, 234	236.25	19	7.93
3	280, 228, 228, 221	239.25	59	27.37
4	210, 249, 241, 246	236.50	39	17.97
5	243, 240, 230, 230	235.75	13	6.75
6	225, 250, 258, 244	244.25	33	14.06
7	240, 238, 240, 243	240.25	5	2.06
8	244, 248, 265, 234	247.75	31	12.92
9	238, 233, 252, 243	241.50	19	8.10
10	228, 238, 220, 230	229.00	18	7.39

[a] Based on data from page 207 of Wadsworth et al. [7]

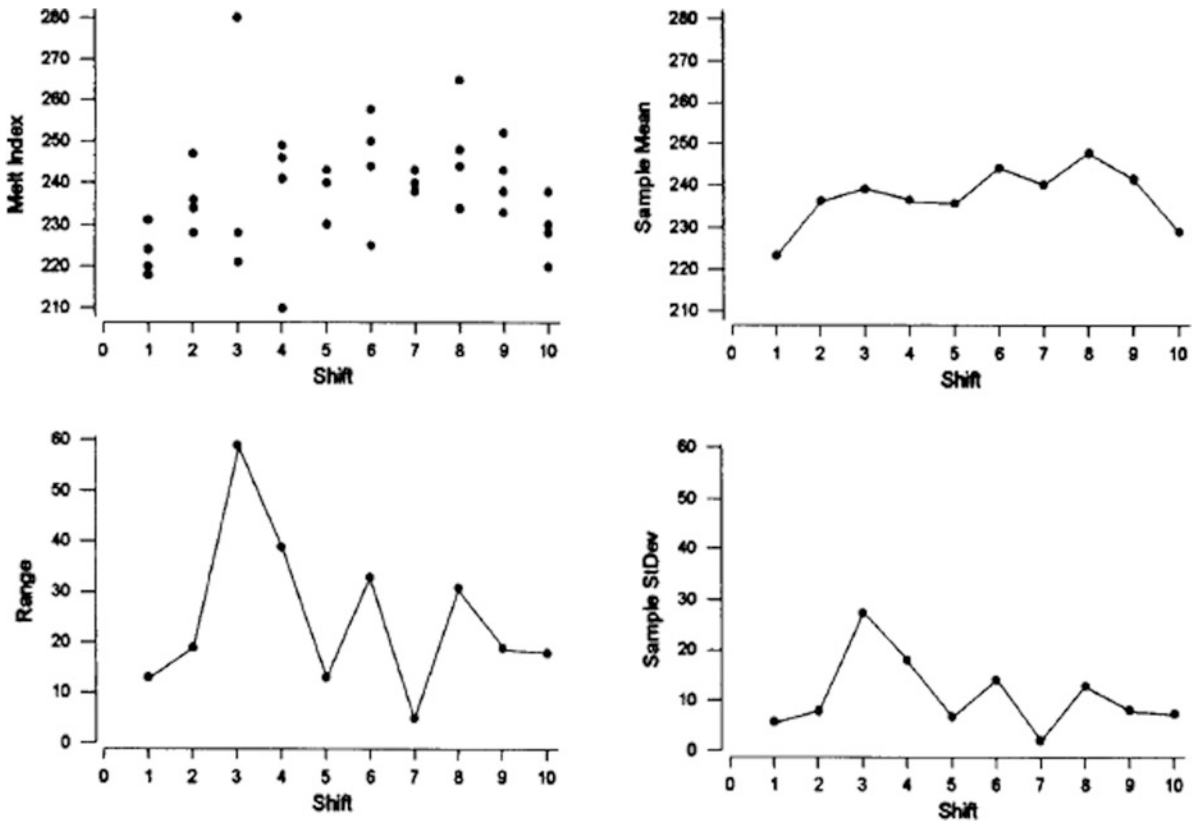


Fig. 4.10 Plots of melt index, $\bar{Y}$, R, and s against shift number

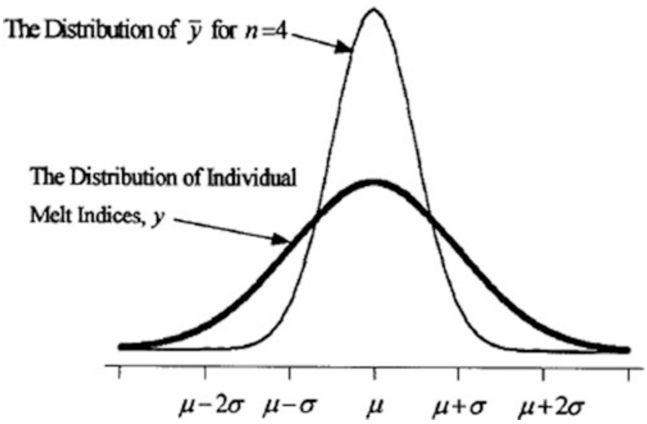


Fig. 4.11 The distribution of individuals, Y, and sample means, $\bar{Y}$

against shift number. The last three of these are the beginnings of so-called Shewhart $\bar{x}$, R, and s control charts.

What remain to be added to the plots in Fig. 4.10 are appropriate control limits. In order to indicate the kind of thinking that stands behind control limits for Shewhart charts, let us concentrate on the issue of limits for the plot of means. The fact is that mathematical theory suggests how the behavior of *means* $\bar{y}$ ought to be related to the distribution of *individual* melt indices y, provided the data-generating process is stable, that is, subject only to random causes. If individual responses y can be described as normal with some mean μ and standard deviation σ, mathematical theory suggests that averages of n such values will behave as if a different normal distribution were generating them, one with a mean $\mu_{\bar{y}}$ that is numerically equal to μ and with a standard deviation $\sigma_{\bar{y}}$ that is numerically equal to $\sigma/\sqrt{n}$. Figure 4.11 illustrates this theoretical relationship between the behavior of individuals and the behavior of means.

The relevance of Fig. 4.11 to the problem of setting control chart limits on means is that if one is furnished with a description of the typical pattern of variation in y, sensible expectations for variation in $\bar{y}$ follow from simple normal distribution calculations. So Shewhart reasoned that since about 99.7% (most) of a Gaussian distribution is within three standard deviations of the center of the distribution, means found to be farther than three theoretical standard deviations (of $\bar{y}$) from the theoretical mean (of $\bar{y}$) could be safely attributed to other than chance causes. Hence, furnished with standard values for μ and σ (describing individual observations), sensible control limits for $\bar{y}$ become

$$\begin{aligned}\text{Upper control limit (UCL) for } \bar{y} &= \mu_{\bar{y}} + 3\sigma_{\bar{y}}\\ &= \mu + 3\frac{\sigma}{\sqrt{n}}\end{aligned}$$

and

$$\begin{aligned}\text{Lower control limit (LCL) for } \bar{y} &= \mu_{\bar{y}} - 3\sigma_{\bar{y}}\\ &= \mu - 3\frac{\sigma}{\sqrt{n}}\end{aligned} \tag{4.9}$$

Returning to the context of our example represented by the data of Table 4.4, Wadsworth et al. [7] state that the target value for melt index in the original application was in fact 235. So if standard process behavior is "on target" behavior, the value $\mu = 235$ seems appropriate for use in (4.9). No parallel value for σ was provided by the authors. Common practice in such situations is to use the data in hand (the data of Table 4.4) to produce a plausible value for σ to use in (4.9). There are many possible ways to produce such a value, but to understand the general logic behind the standard ones, it is important to understand what σ is supposed to measure. The variable σ is intended as a theoretical measure of baseline, short-term, common cause variation. As such, the safest way to try to approximate it is to somehow use only measures of variation *within* the groups of four values in Table 4.4 *not* influenced by variation *between* groups. (Measures of variation derived from considering all the data simultaneously, e.g., would reflect variation between shifts as well as the shorter-term variation within shifts.) In fact, the most commonly used ways of obtaining from the data in hand a value of σ for use in (4.9) are based on the averages of the (within-group) ranges or standard deviations. For example, the ten values of R given in Table 4.4 have a mean

$$\bar{R} = \frac{1}{10}(13 + 19 + 59 + \ldots + 19 + 18) = 24.9$$

and some standard mathematical theory suggests that because the basic group size here is $n = 4$, an appropriate multiple of $\bar{R}$ for use in estimating σ is

$$\frac{\bar{R}}{2.059} \approx 12.1 \tag{4.10}$$

(The divisor above is a tabled factor commonly called d_2, which increases with n.)

Finally, substituting 235 for μ and 12.1 for σ in (4.9) produces numerical control limits for $\bar{y}$:

$$\text{LCL} = 235 - 3\frac{(12.1)}{\sqrt{4}} = 216.9$$

and

$$\text{UCL} = 235 + 3\frac{(12.1)}{\sqrt{4}} = 253.1$$

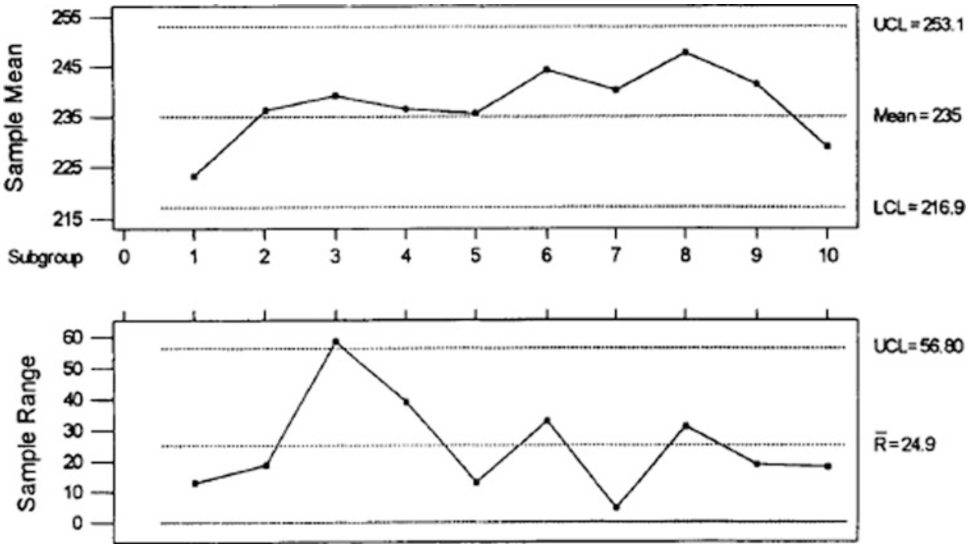


Fig. 4.12 Control charts for $\bar{Y}$and R based on melt indices

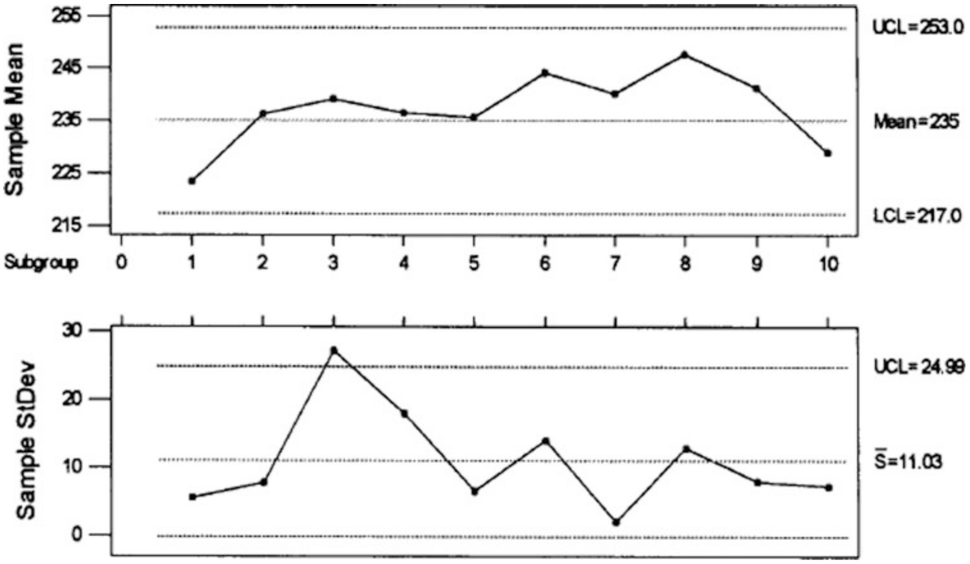


Fig. 4.13 Control charts for $\bar{Y}$and s based on melt indices

Comparison of the $\bar{y}$ values in Table 4.4 to these limits reveals no "out of control" means, that is, no evidence in the means of assignable process variation. Figures 4.12 and 4.13 show control charts for all of $\bar{y}$, R, and s, where control limits for the last two quantities have been derived using standard calculations not shown here.

The *R and s* charts in Figs. 4.12 and 4.13 are related representations (only one is typically made in practice) of the shift-to-shift behavior of melt index *consistency*. It is seen that on both charts, the shift #3 point plots above the upper control limit. The strong suggestion thus is that melt index consistency was detectably worse on that shift than on the others, so that from this point of view the process was in fact *not stable* over the time period represented in Table 4.4. In practice, physical investigation and hopefully correction of the origin of the instability typically would follow, as well as some reconsideration of our earlier assessment of 12.1 as a plausible figure to represent the inherent short-term variability of melt index. (If shift #3 could be treated as a special case, explainable as an unfortunate but correctable situation that was not expected to reoccur, there might be reason to revise $\bar{R}$ downward by deletion of shift #3 from the calculation, and thereby to reduce one's view of the size of baseline process variability. Notice that, in general, such a

Table 4.5 Moisture contents for 20 polymer samples[a]

Sample	Moisture, y	Moving range, MR
1	36	–
2	20	16
3	16	4
4	21	5
5	32	11
6	34	2
7	32	2
8	34	2
9	23	11
10	25	2
11	12	13
12	31	19
13	25	6
14	31	6
15	34	3
16	38	4
17	26	12
18	29	3
19	45	16
20	27	18

[a] Based on data from page 190 of Burr [8]

downward revision of $\bar{R}$ might well also have the effect of causing one to need to rethink his or her assessment of the constancy of the melt index mean.)

There is a variation on the basic "$\bar{x}$ and R chart" idea that we wish to illustrate here next, because of its frequent use in chemical industry applications. That is the making of a so-called x and *MR* chart pair. The motivation for this modification of the ideas outlined thus far in this section is that in many chemical process monitoring contexts the natural "group size" is $n = 1$. A mean of $n = 1$ observation(s) is simply that observation itself, and the limits of (4.9) make perfectly good sense for the case of $n = 1$. That is, the analog of an $\bar{x}$ chart for $n = 1$ cases is clear, *at least if one has an externally provided value for* σ. But what, if anything, to do for an $n = 1$ counterpart of the R chart and how to develop an analog of (4.10) in cases where σ is not a priori known are perhaps not so obvious. Table 4.5 contains data representing moisture contents in 0.01% of bihourly samples of a polymer, and the question at hand is what besides simply the bihourly y values might be plotted in the style of a Shewhart control chart for such data.

The final column of Table 4.5 gives 19 so-called moving ranges of pairs of successive moisture contents. It is often argued that although these *MR* values are actually affected not only by variation within a 2-h production period but by some variation between these periods as well, they come as close to representing purely short-term variation as any measure available from $n = 1$ data. Accordingly, as a kind of $n = 1$ analog of an R chart, moving ranges are often charted in addition to individual values y. Further, the average moving range is used to estimate σ in cases where information on the inherent variability of individuals is a priori lacking, according to the formula

$$\text{Estimated } \sigma = \frac{\overline{\text{MR}}}{1.128}$$

where $\overline{\text{MR}}$ is the mean of the moving ranges (and plays the role of $\bar{R}$ in (4.10)), and 1.128 is then $n = 2$ version of the factor d_2 alluded to immediately below (4.10).

In the case of the data of Table 4.5,

$$\overline{\text{MR}} = \frac{1}{19}(16 + 4 + 5 + \ldots + 16 + 18) \approx 8.2$$

so that a (possibly somewhat inflated due to between period variation) data-based estimate of within-period variability σ for use, for example in limits (4.9), is

$$\frac{8.2}{1.128} \approx 7.2$$

Figure 4.14 shows both an x (individuals) chart and an *MR* (moving range) chart based on these calculations. As no standard value of moisture content was provided in Burr's text [8], the value $\bar{y} = 28.55$ was used as a substitute for μ in (4.9). The MR chart limits are based on standard $n = 2$ (because ranges of "groups" of two observations are being plotted) R chart control limit formulas. Figure 4.14 shows no evidence of assignable variation in the moisture contents.

Statistical research in the last decade has cast serious doubt on the wisdom of adding the *MR* chart to the x chart in $n = 1$ situations. The price paid for the addition in terms of "false alarm rate" is not really repaid with an important increase in the ability to detect process change. For a more complete discussion of this issue see Sect. 4.4 of Vardeman and Jobe [9].

The use of Shewart control charts is admirably documented in a number of statistical quality control books including those by Vardeman and Jobe [9], Wadsworth et al. [7], Duncan [10], Burr [8], Grant and Leavenworth [11], and Ott et al. [12]. Our purpose here is not to provide all details necessary for their use, but only to give the reader an introduction to the overall function that they serve. It should be said, however, that in recent years other statistical process monitoring tools such as the so-called CUmulative SUM (CUSUM) schemes and Exponentially Weighted Moving Average (EWMA) schemes have been developed as competing methodologies, and can in some circumstances be practically more effective than the original Shewhart charts. Indeed, many computerized controllers for real-time chemical process monitoring and adjustment now employ some

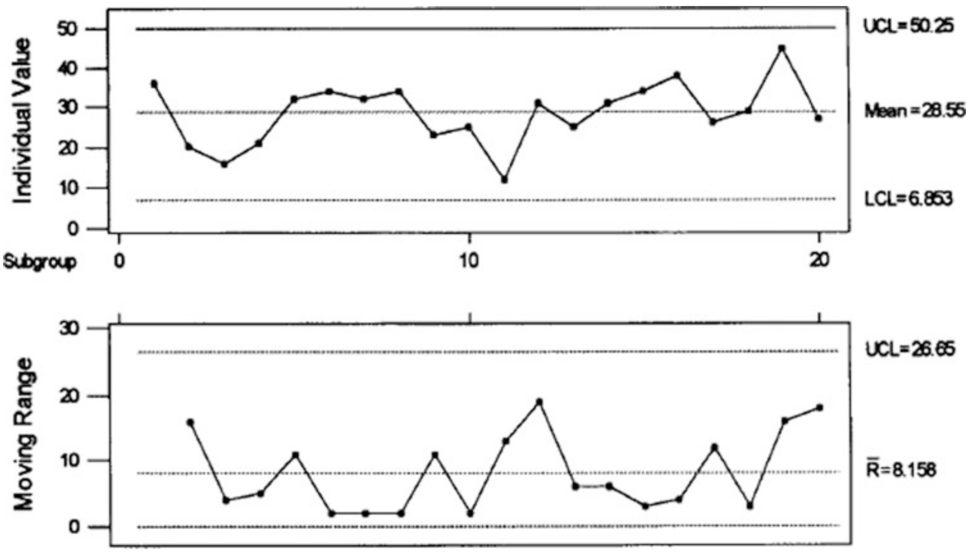


Fig. 4.14 Control charts for *Y* and *MR* based on moisture contents of 20 polymer samples

form of CUSUM or EWMA logic. For more on these topics, including their integration with model-based process controllers, the reader is referred to Sects. 4.1 and 4.2 of Vardeman and Jobe [9] and Vander Wiel et al. [13].

Shewhart's basic conceptualization of common and special cause variation not only leads to control charts as quantitative, rational tools to guide one in knowing when (and when not!) to intervene in an industrial process to correct potential ills, but also provides a framework for considering the question of what is the best/most consistent performance one can hope for from a particular version of a process. That is, it provides a framework for discussing process capability assessment.

If $\hat{\sigma}$ is some (standard deviation type) estimate of the baseline variation inherent in an industrial process (obtained, e.g., from a calculation such as (4.10) or from data taken from the process after eliminating all physical sources of assignable variation), it essentially specifies what is possible in terms of consistency of process output. There are, however, several common ways of using such an estimate to produce related measures of process capability.

For one thing, remembering again the fact that an interval from $\mu-3\sigma$ to $\mu+3\sigma$ (i.e., of length 6σ) will bracket about 99.7% of a normal distribution, the figure 6σ is sometimes stated as "the process capability." This usage would say that in the context of the polyethylene melt index example of Table 4.4 the $\hat{\sigma} = 12.1$ figure from (4.10) implies a melt index process capability of approximately $6(12.1) \approx 72.6$. If properly monitored, the process appears capable of producing almost all individual melt indices in a 73-point range.

Where there are stated specifications for individual measurements y, σ is sometimes turned into a kind of index comparing it to the difference in upper and lower engineering specifications. For example, one such *process capability index* is

$$C_{\mathrm{p}} = \frac{\mathrm{USL} - \mathrm{LSL}}{6\sigma}$$

where *USL–LSL* is the difference in specifications. Fairly obviously, the larger the value of C_{p}, the more comfortably (properly targeted) process output values will fit in an interval from *LSL* to *USL*.

Another process capability measure that is frequently used in the industrial chemistry sector is

$$C_{\mathrm{pk}} = \mathrm{minimum}\left\{C_{\mathrm{pu}} = \frac{\mathrm{USL} - \mu}{3\sigma}, \qquad C_{\mathrm{pl}} = \frac{\mu - \mathrm{LSL}}{3\sigma}\right\}$$

where μ is an overall process average for an in-control/stable/predictable process, and σ is as before. This measure is clearly similar to C_{p}, but it takes into account the placement of the process mean in a way that is ignored by C_{p}. A large value of C_{pk} indicates that not only is the process short-term variation small enough for the process output values to potentially fit comfortably between *LSL* and *USL*, but that the process is currently so targeted that the potential is being realized.

Statistical Methods and Industrial Experimentation

One of the most important areas of opportunity for the new application of statistical methods in the chemical industry in the twenty-first century is that of increasing the effectiveness of industrial experimentation. That is, it is one thing to bring an existing industrial process to stability (a state of "statistical" control), but it is quite another to determine how to make fundamental changes in that process that will improve its basic behavior. This second activity almost always involves some form of experimentation, whether it be in the laboratory or in a plant. As we indicated in the introduction, efficient methods and strategies of such data collection (and corresponding analysis) are a central concern of applied statistics. In this section, we hope to give the reader some insight into the kinds of statistical tools that are available for use in chemical industry experimentation.

We will here take as our meaning of the term "experimentation" the observation of a (typically noisy) physical process under more than one condition, with the broad goal of understanding and then using knowledge of how the process reacts to the changes in conditions. In most industrial contexts, the "conditions" under which the process is observed can be specified in terms of the settings or so-called levels chosen for several potentially important process or environmental variables, the so-called factors in the experiment. In some cases, the hope is to identify those (often largely unregulated) factors and combinations of factors that seem to most influence an observed response variable, as a means of targeting them for attention intended to keep them constant or otherwise to eliminate their influence, and thereby to improve the consistency of the response. In other situations the hope is to discover patterns in how one or more critical responses depend on the levels of (often tightly controlled) factors, in order to provide a road map for the advantageous guiding of process behavior (e.g., to an increased mean reaction yield) through enlightened changing of those levels.

This section is organized into two subsections. In the first, we will illustrate the notion of variance component estimation through an example of a nested or hierarchical data collection scheme. In the second, we will discuss some general considerations in the planning of experiments to detail the pattern of influence of factors on responses, consider so-called factorial and fractional factorial experimental designs, illustrate response surface fitting and interpretation tools and the data requirements they imply, and, in the process, discuss the integration of a number of statistical tools in a sequential learning strategy.

Table 4.6 Forty-four copper content measurements from 11 bronze castings[a]

Casting	Specimen	Determination	Copper content, y, (%)
1	1	1	85.54
1	1	2	85.56
1	2	1	85.51
1	2	2	85.54
2	1	1	85.54
2	1	2	85.60
2	2	1	85.25
2	2	2	85.25
3	1	1	85.72
3	1	2	85.77
3	2	1	84.94
3	2	2	84.95
4	1	1	85.48
4	1	2	85.50
4	2	1	84.98
4	2	2	85.02
5	1	1	85.54
5	1	2	85.57
5	2	1	85.84
5	2	2	85.84
6	1	1	85.72
6	1	2	85.86
6	2	1	85.81
6	2	2	85.91
7	1	1	85.72
7	1	2	85.76
7	2	1	85.81
7	2	2	85.84
8	1	1	86.12
8	1	2	86.12
8	2	1	86.12
8	2	2	86.20
9	1	1	85.47
9	1	2	85.49
9	2	1	85.75
9	2	2	85.77
10	1	1	84.98
10	1	2	85.10
10	2	1	85.90
10	2	2	85.90
11	1	1	85.12
11	1	2	85.17
11	2	1	85.18
11	2	2	85.24

[a] Based on data taken from Wernimont [34]

Identifying Major Contributors to Process Variation

A statistical methodology that is particularly relevant where experimentation is meant to identify important unregulated sources of variation in a response is that of variance component estimation, based on the so-called ANalysis Of Variance (ANOVA) calculations and random effects models. As an example of what is possible, consider the data of Table 4.6. Shown here are copper content measurements for some bronze castings. Two copper content determinations were made on each of two physical specimens cut from each of 11 different castings.

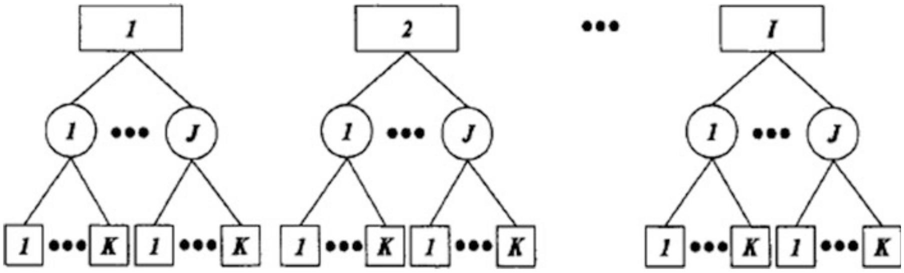


Fig. 4.15 A balanced hierarchical data structure

The data of Table 4.6 were *by design* collected to have a "tree type" or so-called hierarchical/nested structure. Figure 4.15 shows a diagram of a generic hierarchical structure for *balanced* cases like the present one, where there are equal numbers of branches leaving all nodes at a given level (there are equal numbers of determinations for each specimen and equal numbers of specimens for each casting).

An important goal in most hierarchical studies is determining the size of the contributions to response variation provided by the different factors, that is, the different levels of the tree structure. (In the present context, the issue is how variation between castings compares to variation between specimens within a casting, and how they both compare to variation between determinations for a given specimen. If the overall variability observed were considered excessive, such analysis could then help guide efforts at variation reduction by identifying the largest contributors to observed variability.) The structure portrayed in Fig. 4.15 turns out to enable an appealing statistical analysis, providing help in that quantification.

If one lets

y_{ijk} = the copper content from the kth determination of the jth specimen from casting i

$\bar{y}_{ij.} = \frac{1}{2}\sum_k y_{ijk} =$ the mean copper content determination from the ith casting

$\bar{y}_i\cdot\cdot = \frac{1}{2}\sum_j \bar{y}_{ij}\cdot =$ the mean copper content determination from the ith casting

and

$\bar{y}_{\ldots} = \frac{1}{11}\sum_i y_{i\cdot\cdot} =$ the overall mean copperdetermination

it is possible to essentially break down the variance of all 44 copper contents (treated as a single group) into interpretable pieces, identifiable as variation between $\bar{y}_{i\cdot\cdot}$s(casting means), variation between $\bar{y}_{ij\cdot}$s(specimen means) within castings, and variation between $\bar{y}_{ijk}$s (individual measurements) within a specimen. That is, it is an algebraic identity that for 44 numbers y_{ijk} with the same structure as those in Table 4.6

$$\begin{aligned}(44-1)s^2 &= \sum_{i,j,k}(y_{ijk}-\bar{y}_{\ldots})^2\\ &= \sum_{i,j,k}(\bar{y}_{i\cdot\cdot}-\bar{y}_{\ldots})^2+\sum_{i,j,k}(\bar{y}_{ij\cdot}-\bar{y}_{i\ldots})^2\\ &\quad+\sum_{i,j,k}(y_{ijk}-\bar{y}_{ij\cdot})^2 \qquad (4.11)\end{aligned}$$

The sums indicated in (4.11) are over all data points; so, for example, the first summand on the right is obtained for the copper content data by summing each $(\bar{y}_{i\cdot\cdot}-\bar{y}_{\ldots})^2$ a total of 2 × 2 = 4 times, one for each determination on a given casting. With the obvious meaning for the s and the substitution of the total number of data values for 44, the identity in (4.11) applies to any balanced hierarchical data structure. It is a so-called ANOVA identity, providing an intuitively appealing partitioning of the overall observed variability in the data, an *analyzing of the* (*observed*) *variation*.

Some tedious arithmetic "by hand," or use of nearly any commercially available statistical package that includes an ANOVA program, shows that for the copper content data of Table 4.6 the numerical version of (4.11) is approximately

$$5.1385 = 3.2031 + 1.9003 + 0.0351 \qquad (4.12)$$

Although we will not provide any details here, the reader is alerted to the fact that it is common practice to present the elements of an identity such as (4.12) in a tabular form called an "ANOVA table." The use for the elements of (4.12) that we wish to illustrate here is their role in estimating casting, specimen, and determination "variance components."

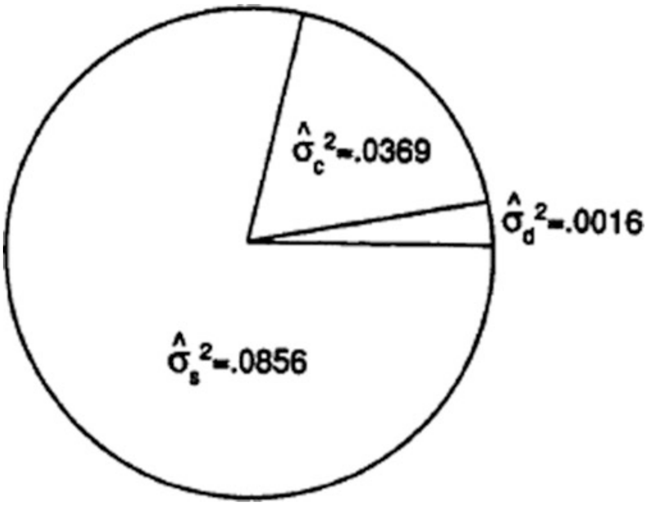


Fig. 4.16 Three estimated variance components for copper contents

That is, if one models an observed copper determination as the sum of a *random* casting-dependent *effect* whose distribution is described by a variance σ_c^s, a *random* specimen-dependent *effect* whose distribution is described by a variance σ_s^2, and a *random* determination-dependent *effect* whose distribution is described by a variance σ_d^2, the elements of (4.12) lead to estimates of the *variance components* σ_c^2, σ_s^2, and σ_d^2 in the model. Note that in such a random effects model of the data-generating process, copper measurements from the same casting share the same casting effect, and copper measurements from the same specimen share the both same casting and the same specimen effects. The individual σ^2 values are conceptually the variances that would be seen in copper contents if only the corresponding sources of variation were present. The sum of the σ^2 values is conceptually the variance that would be seen in copper contents if single determinations were made on a number of different castings.

Standard statistical methodology for estimation of the variance components (which we will not detail here but can, e.g., be found in Sect. 4.5 of Vardeman and Jobe [9] or Chap. 11 of Hicks and Turner [14]) produces

$$\sigma_d^2 = \frac{0.0351}{11 \cdot 2 \cdot (2-1)} \approx 0.0016(\%)^2$$

as an estimate of σ_d^2,

$$\hat{\sigma}_s^2 = \frac{1}{2}\left(\frac{1.9003}{11 \cdot (2-1)} - 0.0016\right) \approx 0.0856(\%)^2$$

as an estimate of σ_s^2, and

$$\begin{aligned}\hat{\sigma}_c^2 &= \frac{1}{2 \cdot 2}\left(\frac{3.2031}{(11-1)} - \frac{1.9003}{11 \cdot (2-1)}\right)\\ &\approx 0.0369(\%)^2\end{aligned}$$

as an estimate of σ_c^2. Figure 4.16 is a pie chart representation of these three estimated variance components as fractions of their sum (the variance predicted if single determinations were made on single specimens from each casting), and graphically identifies inhomogeneity between specimens cut from a single casting as the biggest contributor to observed variation.

On the standard deviation scale the estimates translate to $\hat{\sigma}_d \approx 0.04\%$, $\hat{\sigma}_s \approx 0.29\%$, $\hat{\sigma}_c \approx 0.19\%$. So, for example, the data of Table 4.6 indicate that even if castings and specimens were all exactly alike, it would still be reasonable to expect measured copper contents to vary according to a standard deviation of about 0.04%, presumably due to unavoidable measurement error.

Variance component estimation methodology is not limited to balanced hierarchical experiments, but they do provide an important and straightforward context in which to introduce the technology. More detailed information on the case discussed here and extensions to other kinds of data structures can be found in books by Vardeman [4], Neter et al. [5], Mason et al. [15], and Hicks and Turner [14].

Discovering and Exploiting Patterns of Factor Influence on Responses

Having discussed statistical methodology particularly appropriate to studies whose primary purpose is simply to identify factors with the largest influence on a response, we will now consider methods aimed more directly at detailed experimental quantification of the pattern of factor influence on one or more responses. As an example, we will use a "sanitized" account of some statistical aspects of a highly successful and economically important process improvement project. (Data presented here are not the original data, but resemble them in structure. Naturally, details of the project not central to our expository purposes and those of a proprietary nature will be suppressed.) A more complete version of this case study appears as Chap. 11 of Vardeman [4].

The process monitoring, capability assessment, and variance source identification ideas discussed thus far are almost logical prerequisites for industrial experimentation to detail the nature of dependence of response variables on factors of interest. When an industrial process has been made to operate in a stable manner, its intrinsic variability reduced to the extent practically possible, and that baseline performance quantified and understood, the prospects of success are greatly enhanced for subsequent efforts to understand the effects of potential fundamental process changes.

Preliminary work by various groups left a project team with a batch production process behaving in a stable but unsatisfactory fashion. Obvious sources of variation (both in the process itself and "upstream") had been identified and, to the degree practically possible, eliminated. The result was a process with an average output *purity* of 88% and an associated purity standard deviation of around 5%, and an average *yield* of 43% and an associated yield standard deviation of around 5% as well. The project team was charged with finding ways to increase the purity and yield means to, respectively, 95% and 59% while it is hoped, also further reducing the standard deviations. To accomplish this, the team recognized the need for an improved understanding of how various process variables under their control influenced purity (which we will call y_1) and yield (which we will call y_2). Experimentation to provide this was authorized, and, in particular, attention was focused on four factors consisting of three reactant concentrations and the process run time. We will call the Reactant A mole ratio x_1, the Reactant B mole ratio x_2, the Reactant C mole ratio x_3, and the run time (in hours) x_4.

The choice of experimental factors (what to vary in data collection) is a nontrivial matter of fundamental importance that is best handled by people with firsthand process knowledge. There are a number of popular techniques and tools (such as so-called cause and effect diagrams, discussed for instance in Sect. 2.1 of Vardeman and Jobe [9]) for helping groups brainstorm and reach a consensus on such matters. Further, in cases where a priori knowledge of a process is scarce, relatively small preliminary screening experiments can help reduce a large list of potential factors to a smaller list apparently worthy of more detailed study. (The fractional factorial plans that will be illustrated shortly often are recommended for this purpose).

Once a particular set of experimental factors has been identified, questions about exactly how they should be varied must be answered. To begin with, there is the choice of levels for the factors, the matter of how much the experimental factors should be varied. Particular experimental circumstances usually dictate how this is addressed. Widely spaced (substantially different) levels will in general lead to bigger changes in responses, and therefore clearer indications of how the responses depend upon the experimental factors, than will closely spaced (marginally different) levels. But they may do so at the expense of potentially creating unacceptable or even disastrous process conditions or output. Thus, what may be an acceptable strategy in a laboratory study might be completely unacceptable in a production environment and vice versa.

Given a set or range of levels for each of the individual experimental factors, there is still the question of exactly what combinations of levels actually will be used to produce experimental data. For example, in the process improvement study, standard process operating conditions were $x_1 = 1.5$, $x_2 = 1.15$, $x_3 = 1.75$, and $x_4 = 3.5$, and the project team decided on the ranges

$$1.0 \leq x_1 \leq 2.5, 1.0 \leq x_2 \leq 1.8, 1.0 \leq x_3 \leq 2.5, \tag{4.13}$$

and

$$2.0 \leq x_4 \leq 5.0$$

as defining the initial limits of experimentation. But the question remained as to exactly what sets of mole ratios and corresponding run times were appropriate for data collection.

A natural (but largely discredited) strategy of data collection is the one-variable-at-a-time experimental strategy of picking some base of experimental operations (such as standard operating conditions) and varying the level of only one of the factors away from that base at a time. The problem with such a strategy is that sometimes two or more factors act on responses jointly, doing things in concert that neither will do alone. For example, in the process improvement study, it might well have been that an increase in either x_1, or x_2 alone would have affected yield very little, whereas a simultaneous increase in both would have caused an important increase. Modern strategies of industrial experimentation are conceived with such possibilities in mind, and attempt to spread out observations in a way that gives one some ability to identify the nature of the response structure no matter how simple or complicated it turns out to be.

There are several issues to consider when planning the combinations of levels to include in an experiment. We have already said that it is important to "vary several factors simultaneously." It is also important to provide for some replication of at least a combination or two in the experiment, as a means of getting a handle on the size of the experimental error or baseline variation that one is facing. The replication both verifies the reproducibility of values obtained in the study and identifies the limits of that reproducibility. Also, one must balance the urge to "cover the waterfront" with a wide variety of combinations of factor levels against resource constraints and a very real law of diminishing practical returns as one goes beyond what is really needed in the way of data to characterize response behavior. In addition, the fact that real-world learning is almost always of a sequential rather than a "one shot" nature suggests that it is in general wise to spend only part of an experimental budget on early study phases, leaving resources adequate to follow up directions suggested by what is learned in those stages.

It is obvious that a minimum of two different levels of an experimental factor must appear in a set of experimental combinations if any information is to be gained on the effects of that factor. So one logical place to begin thinking about a candidate design for an industrial experiment is with the set of all possible combinations of two levels of each of the experimental factors. If there are p experimental factors, statistical jargon for such an arrangement is to call it a (complete) $2 \times 2 \times 2 \times \cdots \times 2$ or 2^p factorial plan. For example, in the process improvement situation, an experiment consisting of the running of all 16 possible combinations of

$$\begin{aligned} x_1 &= 1.0 \quad \text{or} \quad x_1 = 2.5 \\ x_2 &= 1.0 \quad \text{or} \quad x_2 = 1.8 \\ x_3 &= 1.0 \quad \text{or} \quad x_3 = 2.5 \end{aligned}$$

and

$$x_4 = 2.0 \quad \text{or} \quad x_4 = 5.0$$

would be called a complete $2 \times 2 \times 2 \times 2$ or 2^4 factorial experiment. Notice that in geometric terms, the (x_1, x_2, x_3, x_4) points making up this 2^4 structure amount to the 16 "corners" in four-dimensional space of the initial experimental region defined in (4.13).

A complete factorial experimental plan is just that, in some sense "complete." It provides enough information to allow one to assess (for the particular levels used) not only individual but also joint or interaction effects of the factors on the response or responses. But when in fact (unbeknownst to the investigator) a system under study is a relatively simple one, principally driven by only a few individual or low-order joint effects of the factors, fewer data actually are needed to characterize those effects adequately. So what is often done in modern practice is initially to run only a carefully chosen part of a full 2^p factorial, a so-called fractional factorial plan, and to decide based on the initial data whether data from the rest of the full factorial appear to be needed in order to adequately characterize and understand response behavior. We will not discuss here the details of how so-called 2^{p-q} fractional factorials are intelligently chosen, but there is accessible reading material on the subject in books by Box et al. [16], and by Vardeman and Jobe [3].

In the process improvement study, what was actually done in the first stage of data collection was to gather information from one-half of a full 2^4 factorial (a 2^{4-1} fractional factorial) augmented by four observations at the "center" of the experimental region (thereby providing both some coverage of the interior of the region, in addition to a view of some of its corners, and important replication as well).

The data in Table 4.7 are representative of what the group obtained.

Table 4.7 Data from an initial phase of a process improvement study

x_1	x_2	x_3	x_4	Purity, y_1(%)	Yield, y_2(%)
1.00	1.0	1.00	2.0	62.1	35.1
2.50	1.0	1.00	5.0	92.2	45.9
1.00	1.8	1.00	5.0	7.0	4.0
2.50	1.8	1.00	2.0	84.0	46.0
1.00	1.0	2.50	5.0	61.1	41.4
2.50	1.0	2.50	2.0	91.6	51.2
1.00	1.8	2.50	2.0	9.0	10.0
2.50	1.8	2.50	5.0	83.7	52.8
1.75	1.4	1.75	3.5	87.7	54.7
1.75	1.4	1.75	3.5	89.8	52.8
1.75	1.4	1.75	3.5	86.5	53.3
1.75	1.4	1.75	3.5	87.3	52.0

The order in which the data are listed is simply a convenient systematic one, not to be confused with the order in which experimental runs were actually made. The table order is far too regular for it to constitute a wise choice itself. For example, the fact that all $x_3 = 1.0$ combinations precede the $x_3 = 2.5$ ones might have the unfortunate effect of allowing the impact of unnoticed environmental changes over the study period to end up being confused with the impact of x_3 changes. The order in which the 12 experimental runs were actually made was chosen in a "completely randomized" fashion. For a readable short discussion of the role of randomization in industrial experimentation, the reader is referred to Box [17].

For purposes of this discussion, attention is focused on the yield response variable, y_2. Notice first that the four y_2 values from the center point of the experimental region have $\bar{y} = 53.2$ and $s = 1.13$ (which incidentally already appear to be an improvement over typical process behavior). As a partial indication of the logic that can be used to investigate whether the dependence of yield on the experimental factors is simple enough to be described adequately by the data of Table 4.7, one can compute some estimated "main effects" from the first eight data points. That is, considering first the impact of the variable x_1 (alone) on yield, the quantity

$$\begin{aligned} \bar{y}_{\text{high}_{x_1}} - \bar{y}_{\text{low}_{x_1}} &= \frac{1}{2}(45.9 + 46.0 + 51.2 + 52.8) \\ &\quad - \frac{1}{4}(35.1 + 4.0 + 41.4 + 10.0) \\ &= 26.35 \end{aligned}$$

is perhaps a sensible measure of how a change in x_1 from 1.00 to 2.50 is reflected in yield. Similar measures for the other variables turn out to be

$$\begin{aligned} \bar{y}_{\text{high}_{x_2}} - \bar{y}_{\text{low}_{x_2}} &= -15.20 \\ \bar{y}_{\text{high}_{x_3}} - \bar{y}_{\text{low}_{x_3}} &= 6.10 \end{aligned}$$

and

$$\bar{y}_{\text{high}_{x_4}} - \bar{y}_{\text{low}_{x_4}} = 0.45$$

These measures provide some crude insight into the directions and magnitudes of influence of the experimental variables on y_2. (Clearly, by these measures $x_1 = 2.50$ seems preferable to $x_1 = 1.00$, and the run time variable x_4 seems to have little impact on yield.) But they also provide strong evidence that the nature of the dependence of yield on the experimental factors is too complicated to be described by the action of the factors individually. For example, if it *were* the case that the separate actions of the experimental factors were adequate to describe system behavior, then standard statistical theory and the data indicate that the mean response for the $x_1 = 1.00$, $x_2 = 1.0$, $x_3 = 1.00$, and $x_4 = 2.0$ set of conditions would be around

$$\hat{y} = \bar{y}_{\text{corners}} - \frac{1}{2}(-26.35) - \frac{1}{2}(-15.20) - \frac{1}{2}(6.10) - \frac{1}{2}(0.45) = 27.45$$

(where $\bar{y}_{\text{corners}}$ is standing for the mean of the first eight yields in Table 4.7). But the observed yield of 35.1 is clearly incompatible with such a mean and the standard deviation value (of $s = 1.13$) derived from the repeated center point. Also, other simple evidence that (at least linear and) separate action of the four factors is not enough to describe yield adequately is given by the large difference between $\bar{y}_{\text{cornere}} = 35.8$ and the observed mean from the center point $\bar{y} = 53.2$. (As it turns out, calculations that we will not show here indicate the *possibility* that individual action of the factors *plus joint action* of the Reactant A and Reactant B mole ratios is sufficient to describe yield. But in any case, the point is that the data of Table 4.7 provide evidence that the pattern of dependence of yield on the experimental variables is not simple, and thus that completion of the 2^4 factorial is in order).

After a complete analysis of the first round of experimental data, the project team "ran the second half fraction" of the 2^4 factorial, and data similar to those in Table 4.8 were obtained. (Again, no significance should be attached to the order in which the observations in Table 4.8 are listed. It is not the order in which the experimental runs were made).

Table 4.8 Data from a second phase of a process improvement study

x_1	x_2	x_3	x_4	Purity, y_1(%)	Yield, y_2(%)
1.00	1.0	1.00	5.0	64.0	35.3
2.50	1.0	1.00	2.0	91.9	47.2
1.00	1.8	1.00	2.0	6.5	3.9
2.50	1.8	1.00	5.0	86.4	45.9
1.00	1.0	2.50	2.0	63.9	39.5
2.50	1.0	2.50	5.0	93.1	51.6
1.00	1.8	2.50	5.0	6.8	9.2
2.50	1.8	2.50	2.0	84.6	54.3

The data from the second phase of experimentation served to complete the project team's 2^4 factorial picture of yield and confirm the tentative understanding drawn first from the initial half fraction. It is seen that the combinations listed in Table 4.8 are in the same order as the first eight in Table 4.7 as regards levels of experimental variables x_1, x_2, and x_3, and that the corresponding responses are very similar. (This, by the way, has the happy practical implication that run time seems to have little effect on final purity or yield, opening the possibility of reducing or at least not increasing the standard run time.) Thorough data analysis of a type not shown here left the project team with a clear (and quantified version of the) understanding that Reactant A and B mole ratios have important individual and joint effects on the responses, and that, acting independently of the other two reactants, Reactant C also has an important effect on the responses. However, it did *not* yet provide a solution to the team's basic problem, which was to reach a 59% mean yield goal.

The data of Tables 4.7 and 4.8 do hold out hope that conditions producing the desired purity and yield can be found. That is, though none of the 16 corners of the experimental region nor the center point appeared to meet the team's yield goal, the data do show that there is substantial *curvature* in the yield response. (The joint effect of x_1 and x_2 amounts to a kind of curvature, and the nonlinearity of response indicated by a large difference between and $\bar{y}_{\text{corners}} \approx 35.8$ and $\bar{y} = 53.2$ at the center of the experimental region also is a kind of curvature.) If one could "map" the nature of the curvature, there is at least the possibility of finding favorable future operating conditions in the interior of the initial experimental region defined in (4.13).

It ought to be at least plausible to the reader that 2^4 factorial data (even supplemented with center points) are not really sufficient to interpolate the nature of a curved response over the experimental region. More data are needed, and a standard way of augmenting a 2^p design with center points to one sufficient to do the job is through the addition of so-called star points to produce a central composite design. Star points are points outside the original experimental region whose levels of all but one of the p experimental factors match those of the center point. Figure 4.17 shows graphical representations of central composite designs in $p = 2$ and $p = 3$ factors.

The project team conducted a third phase of experimentation by adding eight star points to their study and obtained data similar to those in Table 4.9.

The data in Tables 4.7, 4.8, and 4.9 taken together turn out to provide enough information to enable one to rather

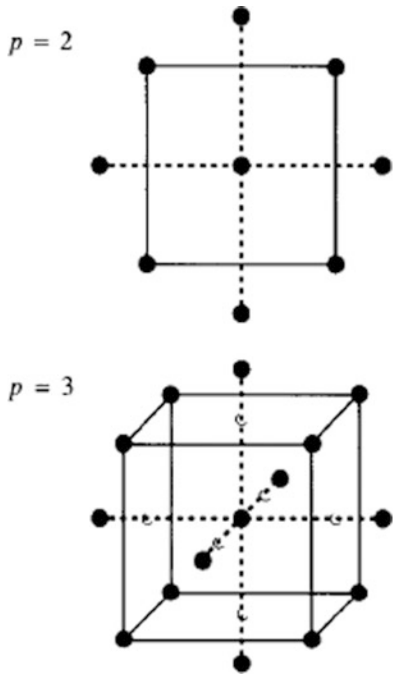


Fig. 4.17 $p = 2$ and $p = 3$ central composite designs

Table 4.9 Data from a third phase of a process improvement study

x_1	x_2	x_3	x_4	Purity, $y_1(\%)$	Yield, $y_2(\%)$
0.6895	1.4	1.75	3.5	20.8	13.0
2.8105	1.4	1.75	3.5	95.9	54.3
1.75	0.8344	1.75	3.5	99.9	62.4
1.75	1.9656	1.75	3.5	65.9	41.2
1.75	1.4	0.6895	3.5	64.4	32.7
1.75	1.4	2.8105	3.5	64.8	40.3
1.75	1.4	1.75	1.379	88.1	52.7
1.75	1.4	1.75	5.621	88.9	50.5

thoroughly quantify the "curved" nature of the dependence of y_2 on x_1, x_2, x_3, and x_4. A convenient and often successful method of accomplishing this quantification is through the least squares fitting of a general *quadratic response surface*. That is, central composite data are sufficient to allow one to fit an equation to a response that involves a constant term, linear terms in all the experimental variables, quadratic terms in all of the experimental variables, and cross-product terms in all pairs of the experimental variables. Appropriate use of a multiple regression program with the project data represented here produces the fitted equation

$$\begin{aligned} y_2 \approx\ & 15.4 + 37.9x_1 - 66.2x_2 + 48.8x_3 \\ & + 0.97x_4 - 16.1x_1^2 - 0.03x_2^2 \\ & - 13.6x_3^2 - 0.046x_4^2 + 26.5x_1x_2 \\ & + 0.344x_1x_3 - 0.217x_1x_4 \\ & + 1.31x_2x_3 - 0.365x_2x_4 + 0.061x_3x_4 \end{aligned}$$

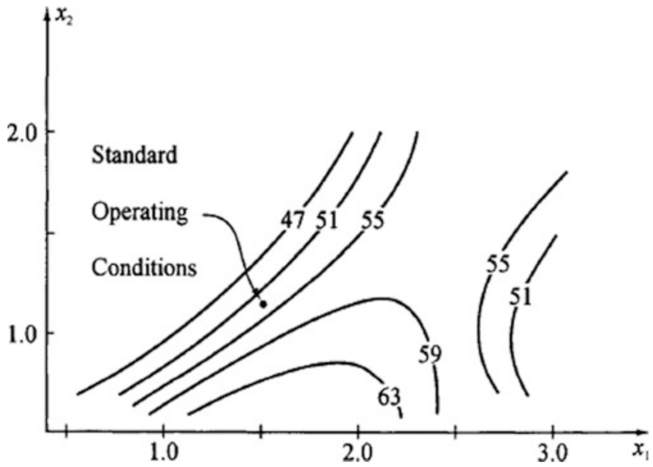


Fig. 4.18 A contour plot of fitted yield when $x_3 = 1.75$ (from *Statistics for Engineering Problem Solving* (*1st Ed.*) by S. B. Vardeman © 1994. Reprinted with permission of Brooks/Cole, a Division of Thomson Learning; http://www.thomsonlearning.com. FAX 800-730-2215)

This may not seem to the reader to be a particularly helpful data summary, but standard multiple regression tools can be used to deduce that an essentially equivalent, far less cluttered, and more clearly interpretable representation of the relationship is:

$$\begin{aligned} y_2 \approx\ & 13.8 + 37.8x_1 - 65.3x_2 + 51.6x_3 \\ & - 16.2x_1^2 - 13.6x_3^2 + 26.5x_1x_2 \end{aligned} \qquad (4.14)$$

Equation (4.14) provides an admirable fit to the data in Tables 4.7, 4.8, and 4.9, is in perfect agreement with all that has been said thus far about the pattern of dependence of yield on the experimental factors, *and* allows one to do some intelligent interpolation in the initial experimental region. Use of an equation like (4.14) ultimately allowed the project team to determine that an increase of x_1 only would, with minimal change in the existing process, allow them to meet their yield goal. (In fact, the single change in x_1 proved to be adequate to allow them to meet all of their yield *and* purity goals!)

Graphical representations similar to those in Figs. 4.18 and 4.19 for (4.14) with $x_3 = 1.75$ (the standard operating value for x_3) were instrumental in helping the team understand the message carried by their data and how yield could be improved. Figure 4.18 is a so-called contour plot (essentially a topographic map) of the fitted equation, and Fig. 4.19 is a more three-dimensional-looking representation of the same surface. Both types of display are commonly used tools of modern statistical experiment design and analysis. The contour plot idea is particularly helpful where several responses are involved, and by overlaying several such plots one can simultaneously picture the various implications of a contemplated choice of process conditions.

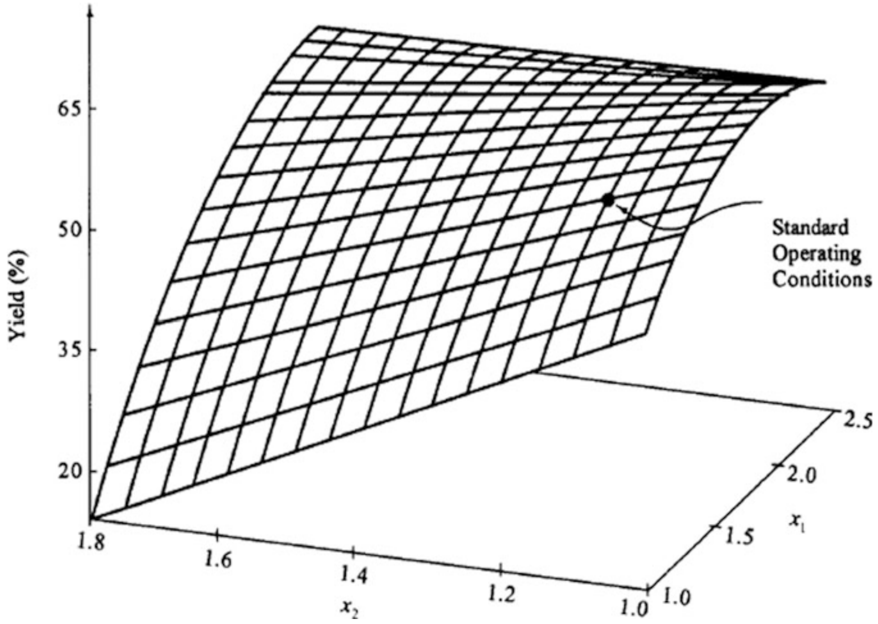


Fig. 4.19 A perspective graph of fitted yield when $x_3 = 1.75$ (from *Statistics for Engineering Problem Solving* by S. B. Vardeman © 1994. Reprinted with permission of Brooks/Cole, a Division of Thomson Learning; http://www.thomsonlearning.com. Fax 800-730-2215)

Special Statistical Tools for Chemical Applications

The statistical methods discussed thus far are of a quite general nature, routinely finding application beyond the bounds of the chemical industry. In this section, we will briefly highlight two statistical methodologies whose most important applications are to chemical problems. That is, we will touch on some of the ideas of mixture experiments and the role of statistics in mechanistic modeling.

Mixture Experiments

In many situations in industrial chemistry, some important measured property of a product is a function of the proportions in which a set of p ingredients or components are represented in a mixture leading to the product. For example, case studies in the literature have covered subjects ranging from octanes of gasoline blends, discussed by Snee [18] to strengths of different formulations of ABS pipe compound, treated in Koons and Wilt [19] to aftertaste intensities of different blends of artificial sweeteners used in an athletic sport drink, discussed by Cornell [20] to moduli of elasticity of different rocket propellant formulations, considered by Kurotori [21]. For experimenting in such contexts, special statistical techniques are needed. These tools have been discussed at length by Cornell [22, 23], and our purpose here is not to attempt a complete exposition, but only to whet the reader's appetite for further reading in this area.

The goal of mixture experimentation is to quantify how proportions $x_1, x_2, x_3, \ldots, x_p$ of ingredients 1 through p affect a response y. Usually, the hope is to fit some kind of approximate equation involving some parameters $\underline{b}$, say

$$y \approx f(x_1, x_2, \ldots, x_p | \underline{b})$$

to a set of n data points ($x_1, x_2, \ldots, x_p, y$), for the purpose of using the fitted equation to guide optimization of y, that is, to find the "best" blend. The logic of data collection and equation fitting is complicated in the mixture scenario by the fact that

$$x_1 + x_2 + \ldots + x_p = 1 \tag{4.15}$$

The linear constraint (4.15) means that (p way) factorial experimentation is impossible, and that special measures must be employed in order to use standard regression analysis software to do least squares equation fitting. We will briefly describe in turn some approaches to experimental design, equation fitting, and presentation of results for the mixture problem under its fundamental constraint (4.15).

In the case of $p = 3$ (a three-component mixture problem), the set of all possible combinations of values for x_1, x_2, and x_3 satisfying (4.15) can be conveniently represented as an equilateral triangular region. Figure 4.20 shows such a region and the so-called simplex coordinate system on the region. The corners on the plot stand for cases where the "mixture" involved is actually a single pure component. Points on the line segments bounding the figure represent two-component mixtures, and interior points represent genuine three-component mixtures. For example, the center of

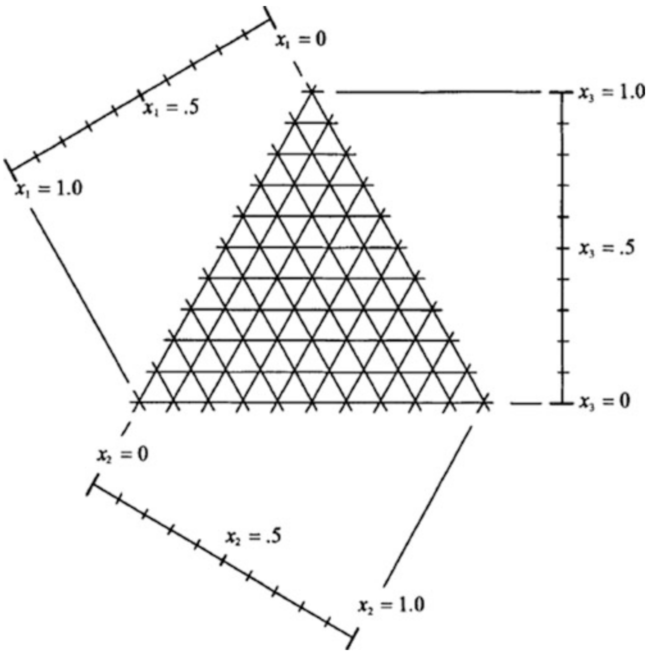


Fig. 4.20 The set of points with $x_1 + x_2 + x_3 = 1$ and a simplex coordinate system (from *Statistics for Engineering Problem Solving* by S. B. Vardeman © 1994. Reprinted with permission of Brooks/Cole, a Division of Thomson Learning; http://www.thomsonlearning.com. Fax 800-730-2215)

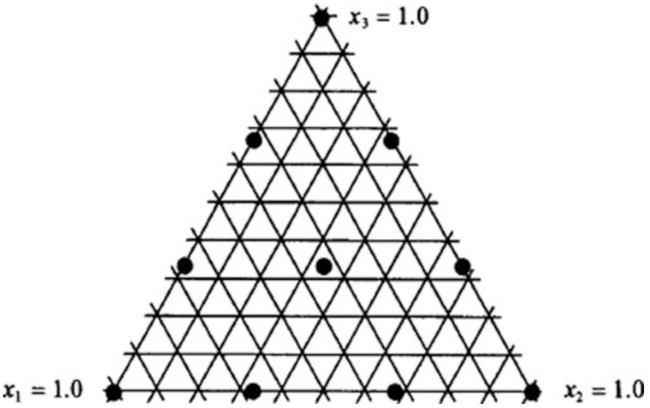


Fig. 4.21 A $p = 3$ simple lattice design

the simplex corresponds to a set of conditions where each component makes up exactly one-third of the mixture.

One standard mixture (experimental) design strategy is to collect data at the extremes (corners) of the experimental region along with collecting data on a regular grid in the experimental region. Figure 4.21 shows a $p = 3$ example of such a so-called simplex lattice design, and Table 4.10 lists the (x_1, x_2, x_3) points involved. (As in the cases of the data in Tables 4.7, 4.8, and 4.9, the order used in the listing in Table 4.10 is not one that would be used in sequencing data collection runs. Instead, a randomly chosen order often is employed).

Table 4.10 (x_1, x_2, x_3) points in a particular $p = 3$ simplex lattice design

x_1	x_2	x_3
1	0	0
0	1	0
0	0	1
$\frac{1}{3}$	$\frac{2}{3}$	0
$\frac{2}{3}$	$\frac{1}{3}$	0
$\frac{1}{3}$	0	$\frac{2}{3}$
$\frac{2}{3}$	0	$\frac{1}{3}$
0	$\frac{1}{3}$	$\frac{2}{3}$
0	$\frac{2}{3}$	$\frac{1}{3}$
$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

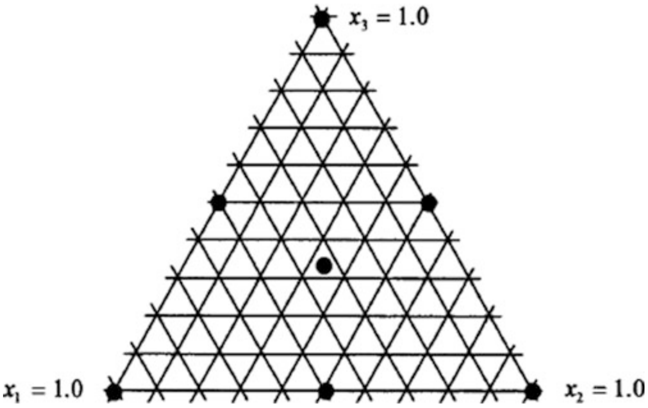


Fig. 4.22 A $p = 3$ simplex centroid design

Another standard mixture experiment strategy is the so-called simplex centroid design, where data are collected at the extremes of the experimental region and for every equal-parts two-component mixture, every equal-parts three-component mixture, and so on. Figure 4.22 identifies the blends included in a $p = 3$ simplex centroid design.

Often, the space of practically feasible mixtures is smaller than the entire set of $x_1, x_2, \ldots, x_p$ satisfying (4.15). For example, in many contexts, "pure" mixtures do not produce viable product. Concrete made using only water and no sand or cement obviously is a useless building product. One common type of constraint on the proportions $x_1, x_2, \ldots, x_p$ that produces quite simple experimental regions is that of lower bounds on one or more of the individual proportions. Cornell [22], for example, discusses a situation where the effectiveness in grease stain removal of a $p = 3$ bleach mixture was studied. Past experience with the product indicated that the proportions by weight of bromine, x_1, of powder, x_2, and of HCl, x_3, needed to satisfy the constraints:

$$x_1 \geq 0.30, x_2 \geq 0.25, \text{ and } x_3 \geq 0.02 \tag{4.16}$$

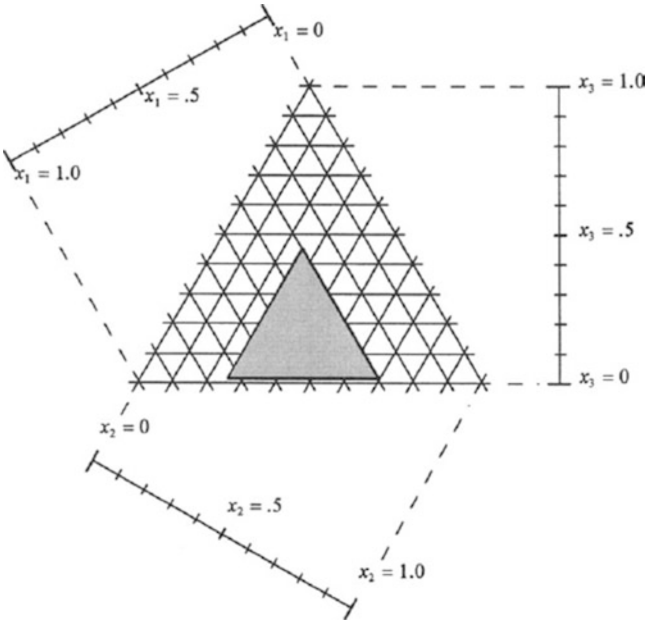


Fig. 4.23 The $p = 3$ simplex and a set of feasible bleach mixtures

for effective action of the product (i.e., the mixture needed to be at least 30% bromine, at least 25% powder, and at least 2% HCl by weight).‘

The effect of adding the lower bound constraints (4.16) to the basic mixture constraint (4.15) can be pictured as in Fig. 4.23. There, a triangular subregion of the basic $p = 3$ simplex depicts the feasible (x_1,x_2,x_3) points. The choice of experimental mixtures for such an experimental region can be made by direct analogy to or rescaling of designs such as the simplex lattice and simplex centroid designs illustrated above to cover the entire simplex. (It is common to refer to the rescaling process as the use of pseudo-components).

Constraint systems more complicated than simple lower bounds produce irregularly shaped experimental regions and less obvious methods of choosing $(x_1, x_2, \ldots, x_p)$ points to cover the experimental region. When $p = 3$, it is possible to sketch the region of feasible points on a simplex plot and use it to help guide the choice of mixture experiment strategy. Figure 4.24 illustrates the kind of region that can arise with other than exclusively lower bound constraints.

When more than three components are involved in a mixture study, such plots are, of course, no longer possible, and other more analytic methods of identifying candidate experimental mixtures have been developed. For example, McLean and Anderson [24] presented an algorithm for locating the vertices of an experimental region defined by the basic constraint (4.15) and any combination of upper and or lower bound constraints

$$0 \leq a_i \leq x_i \leq b_i \leq 1$$

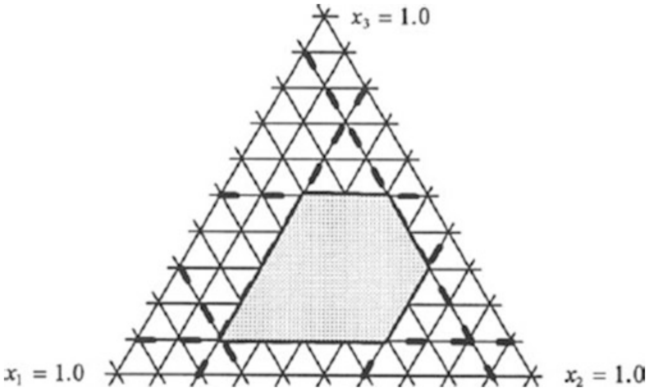


Fig. 4.24 An irregularly shaped experimental region in a $p = 3$ mixture study

on the proportions x_i. Cornell [22, 23] discusses a variety of algorithms for choosing good mixture experiment designs under constraints, and many of the existing algorithms for the problem have been implemented in the MIXSOFT software package developed by Piepel [25].

Empirical polynomial descriptions of (approximately) how a response y depends upon proportions $x_1, x_2, \ldots, x_p$ are popular mixture analysis tools. The process of fitting polynomials to mixture experiment data in principle uses the same least squares notion illustrated in the fitting of a parabola to the data of Table 4.2. However, the mechanics of using standard multiple regression analysis software in the mixture context is complicated somewhat by the basic constraint (4.15). For example, in view of (4.15) the basic ($p + 1$ parameter) linear relationship

$$y \approx b_0 + b_1x_1 + b_2x_2 + \ldots + b_px_p \tag{4.17}$$

is in some sense “overparameterized” in the mixture context, in that it is equivalent to the (p parameter) relationship

$$y \approx b_1x_1 + b_2x_2 + \ldots + b_px_p \tag{4.18}$$

if one identifies the coefficients in (4.18) with the sums of the corresponding coefficients in (4.17) and the coefficient b_0. As a result, it is the “no intercept” relationship (4.18) that is typically fit to mixture data when a linear relationship is used. In a similar way, when a second-order or (multivariable) quadratic relationship between the individual proportions and the response variable is used, it has no intercept term and no pure quadratic terms. For example, in the $p = 3$ component mixture case, the general quadratic relationship typically fit to mixture data is

$$\begin{aligned} y \approx\ & b_1x_1 + b_2x_2 + b_3x_3 + b_4x_1x_2 \\ & + b_5x_1x_3 + b_6x_2x_3 \end{aligned} \tag{4.19}$$

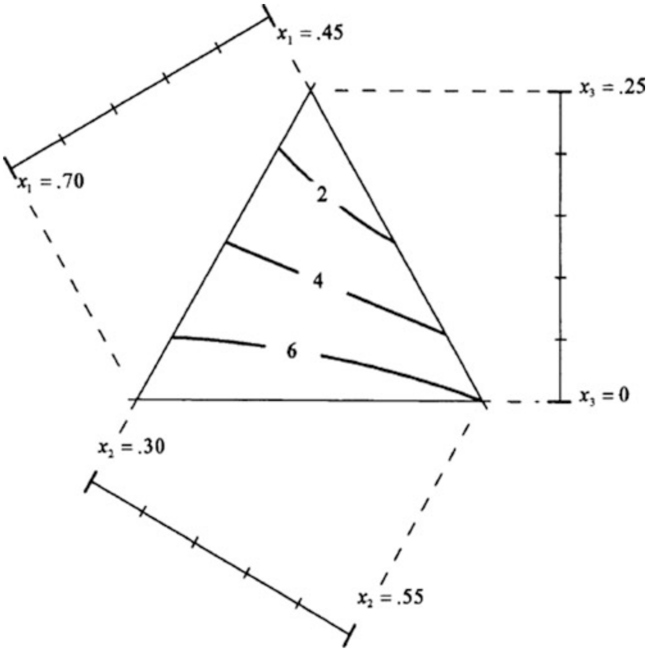


Fig. 4.25 A contour plot made on the $p = 3$ simplex (from *Statistics for Engineering Problem Solving* by S. B. Vardeman © 1994. Reprinted with permission of Brooks/Cole, a Division of Thomson Learning: http://www.thomsonlearning.com. Fax 800-730-2215)

(Any apparently more general relationship involving an intercept term and pure quadratic terms can by use of (4.15) be shown to be equivalent to (4.19) in the mixture context.) Relationships of the type of (4.19) are often called Scheffé models, after the first author to treat them in the statistical literature. Other more complicated equation forms are also useful in some applications, but we will not present them in this chapter. The interested reader is again referred to Cornell [22, 23] for more information on forms that have been found to be tractable and effective.

We should point out that the ability to fit equations of the form (4.18) or like (4.19), or of an even more complicated form, is predicated on having data from enough different mixtures to allow unambiguous identification of the parameters $\underline{b}$. This requires proper data collection strategy. Much of the existing statistical research on the topic of mixture experiment design has to do with the question of wise allocation of experimental resources under the assumption that a particular type of equation is to be fit.

One's understanding of fitted polynomial (and other) relationships often is enhanced through the use of contour plots made on coordinate systems such as that in Fig. 4.25. (This is even true for $p \geq 3$ component mixture scenarios, but the use of the idea is most transparent in the three-component case.) A plot like Fig. 4.25 can be a powerful tool to aid one in understanding the nature of a fitted equation and finding regions of optimum fitted response.

The mixture experiment counterpart to conventional screening/fractional factorial experimentation also is possible. So-called axial designs have been developed for the purpose of providing screening-type mixture data for use in rough evaluation of the relative effects of a large number of mixture components on a response variable. The same kind of sequential experimental strategy illustrated in the process improvement example is applicable in mixture contexts as well as contexts free of a constraint such as (4.15).

Mechanistic Model Building

The kinds of equations most easily fit to multi-factor data using standard (least squares) regression analysis techniques are polynomial equations such as (4.6), (4.14), (4.18), and (4.19). These are particularly convenient because they are linear in their parameters, b. But they are probably best thought of as empirical "mathematical French curve" descriptions of the relation of a response, y, to the explanatory variables, x. Polynomial equations function as extremely useful summaries of observed patterns in one's data, but they do not typically provide direct insight into chemical mechanisms that produce those patterns, and the fitted parameters, b, do not often have direct physical meanings. Their use is particularly appropriate where there is little a priori knowledge of mechanisms involved in a process that might aid in its description, and/or no such knowledge is really essential to achieving one's goals.

Sometimes, however, it is desirable (on the basis of possible reaction kinetics or for other reasons) to posit theoretical descriptions of process outputs in terms of explanatory variables. That is, physicochemical principles often lead (through differential or integral equation descriptions of a system) to equation forms for a response that, like

$$y = \frac{K_1 K_A K_B P_A P_B}{(1 + K_A P_A + K_B P_B)^2}$$
$$y = C_0 \exp(-Kt)$$

and

$$y = \frac{K_1 x}{1 + K_2 x}$$

are nonlinear in the parameters. Although such equations or models may be less tractable than empirical polynomial equations, the parameters involved more often than not *do* have direct physical interpretations. Further, when such a model can be verified as being an adequate description of a process (thereby confirming scientific understanding) and the parameters involved are estimated from process data, such mechanistic models can provide much safer extrapolations beyond an experimental region than the cruder empirical polynomial models.

The process of research in chemical systems is one of developing and testing different models for process behavior. Whether empirical or mechanistic models are involved, the discipline of statistics provides data-based tools for discrimination between competing possible models, parameter estimation, and model verification for use in this enterprise. In the case where empirical models are used, techniques associated with "linear" regression (linear least squares) are used, whereas in mechanistic modeling contexts "nonlinear" regression (nonlinear least squares) techniques most often are needed. In either case, the statistical tools are applied most fruitfully in iterative strategies.

Reilly and Blau [26] and Chap. 16 of Box et al. [16] provide introductions to the general philosophy of using statistical methods in mechanistic modeling contexts, as well as a number of useful references for further reading.

Fairly sophisticated and specialized statistical software is needed in the practical application of nonlinear regression methods to mechanistic modeling for industrial chemistry applications. The techniques implemented in such software are discussed in Seber and Wild [27], Bates and Watts [28], Bard [29], and Riley and Blau [26].

Modern Business Process Improvement and the Discipline of Statistics

The modern global business environment is fiercely competitive in all sectors, including the chemical sector. It is by now widely recognized that corporate survival in this environment depends upon constant improvement in all business processes, from billing to production. Companies have adopted a variety of programs and focuses aimed at facilitating that improvement. A decade ago, efforts organized around a *Total Quality Management* banner (with liberal references to emphases of consultants like W. E. Deming, J. M. Juran, and A. Feigenbaum) were popular. More recently, programs keyed to ISO 9000 [30] certification criteria and Malcolm Baldridge Award [31] criteria have become prominent. And currently probably the most visible programs are the so-called *Six Sigma* programs.

In one sense there is nothing new under the sun, and all successful business process improvement programs (including those in the chemical sector) must in the end reduce to organized problem-solving disciplines. So it is not surprising that programs quite different in name are often very alike in fundamental content. And as they must necessarily make use of empirical information (data), they must have significant statistical components. To make this connection to statistics slightly more explicit, we proceed to provide a few additional details on the Six Sigma movement. (Further material on the subject is easy to find using an Internet search engine, as there are many consultants eager to sell their advice and Six Sigma training. The American Society for Quality at http://www.asq.org offers many entries into the subject. And a search at amazon.com for "Six Sigma" books already produced 6,666 hits in May 2004. Fashions change quickly enough that it seems pointless to provide more detailed recommendations for follow up on the subject).

The phrase "Six Sigma" originated at Motorola in the late 1980s. Programs there and at General Electric in the mid-1990s are widely touted as important contributors to company profits and growth in stock values. The name is now commonly used in at least three different ways. "Six Sigma" refers to

- A goal for business process performance
- A discipline for improvement to achieve that performance
- A corporate program of organization, training, and recognition conceived to support the process improvement discipline

As a goal for business process improvement, "Six Sigma" is equivalent to "$C_{\mathrm{pk}} = 2$." What is perhaps confusing to the uninitiated is that this goal has connections (through normal distribution tail area calculations) to small ("parts per million") fractions defective relative to two-sided specifications on y. Six Sigma proponents often move between the "small process variation" and "parts per million" understandings with little warning.

Six Sigma process improvement disciplines are typically organized around the acronym "MAIC." The first step in an MAIC cycle is a Measure step, wherein one finds appropriate process responses to observe, identifies and validates measurement systems and collects baseline process performance (process monitoring) data. The second step is an Analyze step. This involves summarizing the initial process data and drawing appropriate inferences about current process performance. The third step in an MAIC cycle is an Improve step, where process knowledge, experimentation, and more data analysis are employed to find a way to make things work better. Finally, the four-step cycle culminates in a Control (process monitoring) effort. The object here is to see that the newly improved performance is maintained after a project team moves on to other problems.

Six Sigma corporate organization, training, and recognition programs borrow from the jargon and culture of the martial arts. People expert in the process improvement paradigm are designated "black belts," "master black belts," and so on. These people lead project teams and help train new initiates ("green belts") in the basics of the program and its tools (including statistical tools). The emphasis throughout is on completing projects with verifiable large dollar impact.

Having made the point that improvement in all business activities is of necessity data-driven, it is hopefully obvious that the emphases and methods of the subject of statistics are useful beyond the lab and even production. Of course, for broad implementation, it is the most elementary of statistical methods that are relevant.

Conclusion

We have tried in this chapter to give readers the flavor of modern applied statistical methods and to illustrate their usefulness in the chemical industry. Details of their implementation have of necessity been reserved for further more specialized reading, for which the interested reader is encouraged to consult the references given in this chapter.

References

1. Chambers J, Cleveland W, Kleiner B, Tukey P (1983) Graphical methods of data analysis. Duxbury, Boston, MA
2. Devore J (1991) Probability and statistics for engineering and the sciences, 3rd edn. Brooks/Cole, Pacific Grove, CA
3. Vardeman SB, Jobe JM (2001) Basic engineering data collection and analysis. Duxbury/Thomson Learning, Pacific Grove, CA
4. Vardeman SB (1994) Statistics for engineering problem solving. PWS-Kent, Boston, MA
5. Neter J, Kutner M, Nachsteim C, Wasserman W (1996) Applied linear statistical models, 4th edn. McGraw-Hill, New York
6. Hahn G, Meeker W (1991) Statistical intervals: a guide for practitioners. Wiley, New York
7. Wadsworth H, Stephens K, Godfrey B (1986) Modern statistical methods for quality control and improvement. Wiley, New York
8. Burr I (1976) Statistical quality control methods. Dekker, New York
9. Vardeman S, Jobe JM (1999) Statistical quality assurance methods for engineers. Wiley, New York
10. Duncan A (1986) Quality control and industrial statistics, 5th edn. Irwin, Homewood, IL
11. Grant E, Leavenworth R (1996) Statistical quality control, 7th edn. McGraw-Hill, New York
12. Ott E, Schilling E (1990) Process quality control. McGraw-Hill, New York
13. Vander Wiel S, Tucker W, Faltin F, Doganaksoy N (1992) Algorithmic statistical process control: concepts and an application. Technometrics 34(3):286–297
14. Hicks C, Turner K (1999) Fundamental concepts in the design of experiments, 5th edn. Oxford University Press, New York
15. Mason R, Gunst R, Hess J (1989) Statistical design and analysis of experiments. Wiley, New York
16. Box G, Hunter W, Hunter JS (1978) Statistics for experimenters. Wiley, New York
17. Box G (1990) George's column. Qual Eng 2:497–502
18. Snee R (1981) Developing blending models for gasoline and other mixtures. Technometrics 23:119–130
19. Koons G, Wilt M (1985) Design and analysis of an ABS pipe compound experiment. In: Snee R, Hare L, Trout JR (eds) Experiments in industry: design, analysis and interpretation of results. ASQC Quality Press, Milwaukee, WI, pp 111–117
20. Cornell J (1986) A comparison between two ten-point designs for studying three-component mixture systems. J Qual Technol 18:1–15
21. Kurotori I (1966) Experiments with mixtures of components having lower bounds. Ind Qual Control 22:592–596
22. Cornell J (1983) How to run mixture experiments for product quality, vol 5 in the ASCQ "Basic References in Quality Control" series, vol. American Society for Quality Control, Milwaukee, WI
23. Cornell J (1990) Experiments with mixtures: designs, models, and the analysis of mixture data, 2nd edn. Wiley, New York
24. MacLean R, Anderson V (1966) Extreme vertices design of mixture experiments. Technometrics 8:447–454
25. Piepel G (1989) Mixsoft and Misoft user's guide, Version 1.0. Mixsoft-Mixture Experiment Software, Richland, WA
26. Reilly P, Blau G (1974) The use of statistical methods to build mathematical models of chemical reacting systems. Can J Chem Eng 52:289–299
27. Seber G, Wild C (1989) Nonlinear regression. Wiley, New York
28. Bates D, Watts D (1988) Nonlinear regression analysis and its applications. Wiley, New York
29. Bard Y (1974) Nonlinear parameter estimation. Academic, New York
30. International Organization for Standardization (http://www.iso.ch)
31. National Institute of Standards and Technology (http://www.quality.nist.gov)
32. Albin S (1990) The lognormal distribution for modeling quality data when the mean is near zero. J Qual Technol 22:105–110
33. Odioso R, Henke A, Stauffer H, Frech K (1961) Direct hydration of olefins. Ind Eng Chem 53(3):209–211
34. Wernimont G (1989) Statistical quality control in the chemical laboratory. Qual Eng 2:59–72

Green Engineering: Integration of Green Chemistry, Pollution Prevention, and Risk-Based Considerations[*]

5

David Shonnard, Angela Lindner, Nhan Nguyen,
Palghat A. Ramachandran, Daniel Fichana, Robert Hesketh,
C. Stewart Slater, and Richard Engler

Acronyms

MEN	Mass exchange network
MSA	Mass separation agent
P2	Pollution prevention
PM_5	Particulate matter <5 μm in diameter
POTW	Publicly owned (wastewater) treatment works
RCRA	Resource conservation and recovery act

Overview

Literature sources on green chemistry and green engineering are numerous. The objective of this chapter is to familiarize readers with some of the green engineering and chemistry concepts, approaches, and tools. In order to do this, the chapter is organized into five sections as follows.

Section "Introduction to Green Chemistry and Green Engineering" provides an introduction to green chemistry and green engineering. Section "Pollution Prevention Heuristics for Chemical Processes" provides examples of pollution prevention heuristics for chemical processes. Heuristics of the two most important unit operations, reactors and separators, are covered. Section "Understanding and Prediction of the Environmental Fate of Chemicals" introduces readers to the concept of environmental fate and transport and prediction of environmental fate properties. Understanding of environmental fate and transport is important for exposure assessment and also is essential for evaluating environmental performance of processes and products during process development and design.

Section "Environmental Performance Assessment for Chemical Process Design" covers the environmental performance assessment for chemical processes design and introduces a three-tier approach to green engineering design of processes incorporating green chemistry, pollution prevention, environmental fate and transport, and life-cycle approach. Finally, Section "Life-Cycle Assessment" provides more background and examples on life-cycle assessment which is an essential principle of green engineering.

Introduction to Green Chemistry and Green Engineering

The Pollution Prevention Act of 1990 (42 U.S.C. 13101–13109) established a national policy to prevent or reduce pollution at its source whenever feasible. The Pollution Prevention Act also provided an opportunity to expand beyond traditional EPA programs and devise creative strategies to protect human health and the environment. The pollution prevention (P2) hierarchy established by this act is illustrated later in Table 5.1. Shortly after the passage of the Pollution Prevention Act of 1990, the EPA's Office of Pollution Prevention and Toxics (OPPT) launched a model research grants program called "Alternative Synthetic

---

[*] US EPA. The chapter does not represent the views of the US EPA or the US Government.

D. Shonnard
Department of Chemical Engineering, Michigan Technological University, 1400 Townsend Drive, 2021, Houghton, MI 49931, USA
e-mail: drshonna@mtu.edu

A. Lindner
Department of Environmental Engineering Sciences, University of Florida at Gainesville, 312 Weil Hall, Stadium Road, Gainesville, FL 32611-6450, USA

N. Nguyen • R. Engler
US EPA, Washington, DC, USA

P.A. Ramachandran
Department of Chemical Engineering, Washington University, St. Louis, Campus Box 1198One Brookings Drive, St. Louis, MO 63130-4899, USA

D. Fichana
Clinical Supply Operations, Bristol Myers Squibb, New Brunswick, NJ, USA

R. Hesketh • C.S. Slater
Department of Chemical Engineering, Rowan University, 201 Mullica Hill Road, Glassboro, NJ 08028, USA

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_5,

Table 5.1 P2 hierarchy according to Pollution Prevention Act of 1990

1.	Source reduction
2.	In-process recycling
3.	On-site recycling
4.	Off-site recycling
5.	Waste treatment to render the waste less hazardous
6.	Secure disposal
7.	Direct release to the environment

Pathways for Pollution Prevention." This program provided grants for research projects that include pollution prevention in the design and synthesis of chemicals. The grant program eventually resulted in the establishment of the Green Chemistry Program around 1991 and 1992. Over the years, this program has catalyzed the development of many green chemistry approaches and environmentally benign chemical syntheses.

Green chemistry is the use of chemistry to reduce pollution at the source. More specifically, green chemistry is the design of chemical products and processes that reduce or eliminate the use or generation of hazardous substances. It is an overarching philosophy of chemistry, rather than a discipline of chemistry such as organic or inorganic. In fact, green chemistry is usually multidisciplinary, drawing on a broad range of expertise. Green chemistry, also called *sustainable chemistry*, offers lower-hazard alternatives to traditional technologies. Green chemistry and green engineering are partners in achieving sustainability.

Green engineering research was begun in the early 1990s to support environmental risk assessment of new chemicals. OPPT staff realized that the risk-based tools used to assess human health and environmental risk of new chemicals, when combined with traditional engineering design, could result in "greener" processes. This recognition eventually resulted in the establishment of the green engineering program which has catalyzed the green engineering movement, including many educational initiatives, research programs, and development of environmentally beneficial processes and products.

Environmental risk or environmental impact is an essential concept of green engineering (http://www.apa.goviogpugreenengmeaning) [1]. There are numerous literature references on environmental risk assessment and some are included in the list of references in this chapter. Risk assessment considers the extent of harm a chemical and its uses pose to human health and the environment. Mathematically, it is expressed as a function of hazard and exposure:

$$\text{Risk} = f(\text{hazard and exposure})$$

A hazard is anything that will produce an adverse effect on human health and the environment. In environmental risk assessment, the hazard component generally refers to toxicity. Exposure is the quantitative or qualitative assessment of contact to the skin or orifices of the body by a chemical. Traditional pollution prevention techniques focus on reducing waste as much as possible; however, risk assessment methods used in pollution prevention can help quantify the degree of environmental impact for individual chemicals. This approach provides a powerful tool that enables engineers to better design processes and products by focusing on the most beneficial methods to minimize all aspects of risk.

By applying risk assessment concepts to processes and products, one can accomplish the following:

- Estimate the environmental impacts of specific chemicals on people and ecosystems.
- Prioritize chemicals that need to be minimized or eliminated.
- Optimize design to avoid or reduce environmental impacts.
- Assess feed and recycle streams based on risk and not volume within a chemical process.
- Design "greener" products and processes.

Historically, scientists, engineers, and policy makers have focused their energy on minimizing exposure as the easiest way to minimize risk to humans or to the environment. Green chemistry focuses on the hazard or toxicity component of the equation instead of exposure. In reducing hazard, risk is reduced in a more reliable and frequently cost-effective manner. Exposure controls, such as personal protective equipment, thermal oxidizers, or treatment plants, are frequently expensive. Also, there is always a risk that an exposure control can fail. Such failure can be mitigated and minimized, but it cannot be eliminated. If instead, hazard is reduced, risk is minimized more reliably. To be sure, there are not yet technologies to eliminate the use or generation of all hazardous substances, so exposure controls are still needed, but to the extent that future technological developments can minimize hazard, the need for treatment or remediation will also be minimized.

Chemistry is an inherently creative discipline. Chemists routinely create new molecules and new methods to make molecules. Green chemistry taps this creativity. The 12 Principles of Green Chemistry [2], originally published in *Green Chemistry: Theory and Practice* [3], provide a roadmap for implementing green chemistry. Green chemistry techniques and principles are very powerful, especially for development of new chemicals and processes.

Green engineering (see Fig. 5.1) provides a system-based framework for evaluating and improving the environmental performance of chemical processes and products (both new and existing) by integrating consideration of health and environmental risk, green chemistry, and pollution prevention approaches into traditional engineering design.

As defined previously, risk is a function of hazard and exposure. In the environmental risk context, hazard is a

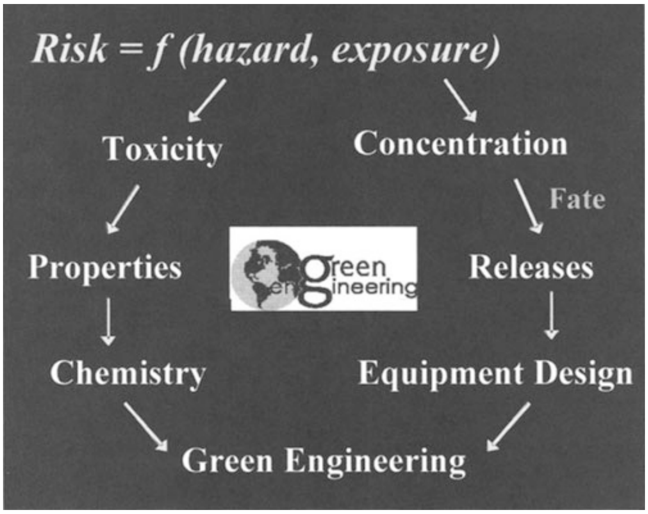


Fig. 5.1 Green engineering is a holistic approach to green process development and design

function of toxicity which is affected by physical/chemical and environmental fate properties and hence chemistry. Green chemistry reduces the risk or environmental impact of processes or products by focusing on the hazard component.

Exposure is a function of concentration and also is affected by the environmental fate properties of a chemical. Concentration is a function of the environmental release that is affected by the equipment or unit operations.

Environmental fate is an important subject area of environmental engineering discipline, especially in the modeling and design of treatment systems. It is also an important component of green engineering. Environmental fate properties can be useful for designing greener chemicals and performing environmental performance assessment of processes.

Green engineering minimizes risk or environmental impact by addressing both the exposure component or unit operations and the hazard or chemistry of the process or product (Fig. 5.2).

Figure 5.3 provides an example for a green engineering process. The left-hand side is a conventional process for making methyl acetate which involves a reactor and a series of distillation steps. The right-hand side is a greener process which involves a combined unit operation, reactive distillation. The green process is superior to the conventional process in many aspects, including chemistry, exposure, release, energy, and economics.

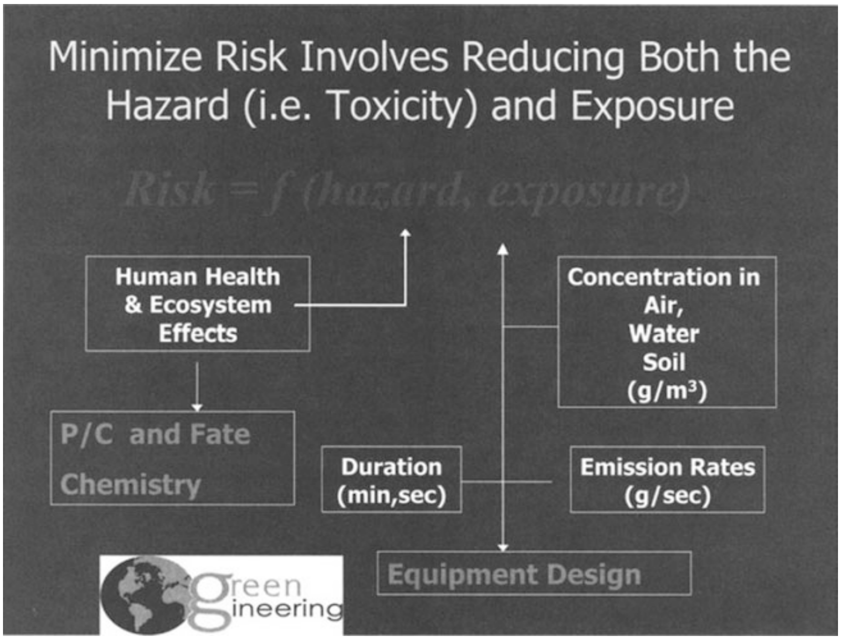
There have been a number of green engineering concepts, approaches, and tools that have been developed since the introduction of green engineering in the mid-1990s. Many of the green engineering approaches, concepts, and tools have been compiled into a standard textbook, *Green Engineering: Environmentally Conscious Design of Chemical Processes* by Allen and Shonnard [5]. These concepts, approaches, and tools can also be found via accessing the EPA Green Engineering Web site at www.epa.gov/oppt/greenengineering. The Web site contains links to many computerized green design tools that can be downloaded free of charge.

Twelve Principles of Green Chemistry[1]

1. Prevent waste: *Design chemical syntheses to prevent waste, leaving no waste to treat or clean up*. As commonsensical as this sounds, many chemists do not consider the waste generated by the syntheses they design. Coproducts (substances that are formed in stoichiometric amounts during a reaction) and by-products (substances that are formed unintentionally, usually in side reactions) are an inconvenience with which others are expected to deal.
2. Design safer chemicals and products: *Design chemical products to be fully effective, yet have little or no toxicity*. Although much is understood about the hazard of various substances, chemists frequently neglect toxicity as a design criterion when they are evaluating the performance of a chemical substance. Because hazardous materials are so routinely used in the laboratory, hazard becomes a trivial matter. Hazard should be explicitly considered and minimized during chemical design.
3. Design less hazardous chemical syntheses: *Design syntheses to use and generate substances with little or no toxicity to humans and the environment*. A holistic view of a synthetic pathway often allows a chemist to change factors and minimize hazard in a number of steps simultaneously. A safer synthesis may simply reduce the number of isolations and purifications, or it may allow a cascade of changes where a new first step fosters changes down the line.
4. Use renewable feedstocks: *Use raw materials and feedstocks that are renewable rather than depleting*. Renewable feedstocks are often made from agricultural and forest products or are the wastes of other processes; depleting feedstocks are made from fossil fuels (petroleum, natural gas, or coal) or are mined. Also the oxidation state of renewable feedstocks is often close to that of the desired product, thus simplifing and reducing the number and extent of chemical transformations necessary in a synthetic pathway.
5. Use catalysts, not stoichiometric reagents: *Minimize waste by using catalytic reactions*. Catalysts are used in small amounts and can carry out a single reaction

---

[1] Adapted from *Green Chemistry: Theory and Practice* [3].



Minimize Risk Involves Reducing Both the Hazard (i.e. Toxicity) and Exposure

Risk = f (hazard, exposure)

Human Health & Ecosystem Effects

Concentration in Air, Water Soil (g/m³)

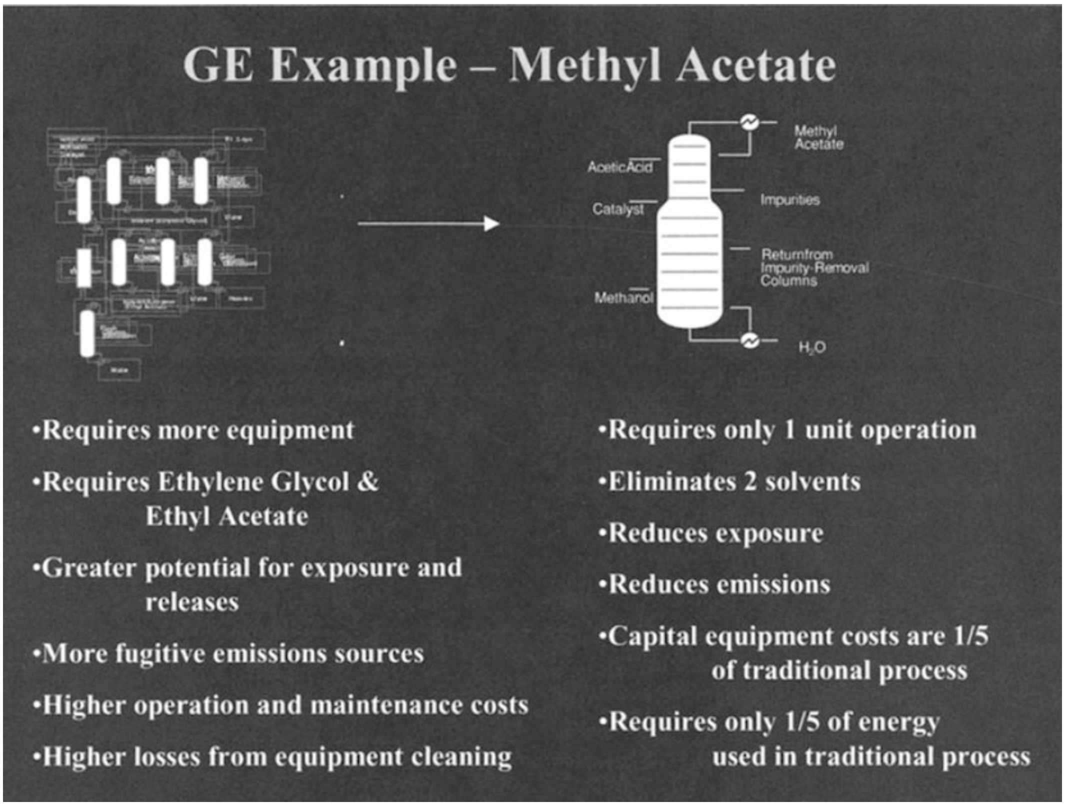
P/C and Fate Chemistry

Duration (min,sec)

Emission Rates (g/sec)

Equipment Design

Fig. 5.2 Green engineering addresses both the hazard and exposure components of the risk equation (P/C = physical chemical properties)



GE Example – Methyl Acetate

- Requires more equipment
- Requires Ethylene Glycol & Ethyl Acetate
- Greater potential for exposure and releases
- More fugitive emissions sources
- Higher operation and maintenance costs
- Higher losses from equipment cleaning

- Requires only 1 unit operation
- Eliminates 2 solvents
- Reduces exposure
- Reduces emissions
- Capital equipment costs are 1/5 of traditional process
- Requires only 1/5 of energy used in traditional process

Fig. 5.3 Green engineering example: conventional vs. greener process using reactive distillation technologies for the production of high-purity metal acetate (adapted from Malone and Russ [4] and modified) many times. They are preferable to stoichiometric reagents, which are usually used in excess and work only once. Even oxidations and reductions, which require a change in oxidation state and therefore an ultimate electron sink or source can be improved by using a catalyst. If reduction using hydrogen gas or oxidation with oxygen can be selectively catalyzed, the hazard and amount of waste generated is minimized.

6. Avoid chemical derivatives: *Avoid using blocking or protecting groups or any temporary modifications if possible*. Derivatives use additional reagents and generate waste. By definition, any group that is added and removed is waste. Reuse of protecting groups is uneconomical, leading to substantial amounts of waste when they are used. Reactants or reaction conditions may be tailored to maximize selectivity for the desired moiety.
7. Maximize atom economy: *Design syntheses so that the final product contains the maximum proportion of the starting materials. There should be few, if any, wasted atoms*. Reactions should also be designed to minimize the *E*-factor (ratio of the mass of all reaction waste to the mass of the desired product). Being aware of all the reaction inputs and outputs helps the designer maximize the benefit of the new chemistry. Materials efficiency not only minimizes environmental impact, but maximizes cost efficiency.
8. Use safer solvents and reaction conditions: *Avoid using solvents, separation agents, or other auxiliary chemicals. If these chemicals are necessary, use innocuous chemicals*. More and more research demonstrates reactions may not require solvents to proceed in clean and quantitative yields. When solvents are required, water, CO_2, ethanol, or other low-toxicity alternatives are preferred to traditional organic solvents.
9. Increase energy efficiency: *Run chemical reactions at ambient temperature and pressure whenever possible*. Shorter reaction times also help minimize energy use. Designers should be cognizant of the conditions necessary to carry out their transformations.
10. Design chemicals and products to degrade after use: *Design chemical products to break down to innocuous substances after use so that they do not accumulate in the environment*. This is arguably one of the hardest principles to apply, despite the existence of sound experimental evidence of what groups degrade well in the environment. The trick, of course, is how the substance “knows” it is at the end of its useful life. Products must be stable long enough to be available before and during use. Degradation may be triggered by a change in the conditions: presence or absence of water, light, oxygen, microorganisms, or other environmental factors.
11. Analyze in real time to prevent pollution: *Include in-process real-time monitoring and control during syntheses to minimize or eliminate the formation of by-products*. It is possible to monitor reactions and to adjust the reaction parameters, such as feedstock ratios, temperature, pressure, etc., to maximize the selectivity for the desired product.
12. Minimize the potential for accidents: *Select chemicals and their physical forms to minimize the potential for chemical accidents, including explosions, fires, and releases to the environment*. Whenever possible, select however energy forms of substances to minimize the energy needed as well as the energy content in case of accident. Avoid substances that are corrosive, highly reactive, or acutely toxic. Inherently safer chemistry reduces the risk from accidents as well as intentional harm.

Principles of Green Engineering

Green engineering is the design, commercialization, and use of processes and products that are feasible and economical yet, at the same time, minimize generation of pollution at the source and risk to human health and the environment. Green engineering embraces the concept that decisions to protect human health and the environment can have the greatest impact and cost-effectiveness when applied early to the design and development phase of a process or product.

There exist a number of green principles, including green engineering principles that have been developed. Most of these principles have overarching themes and objectives. For example, some were developed to help with design. The Green Engineering Principles[2] presented below were developed at the Green Engineering: Defining the Principles Conference held in Sandestin, Florida in May 2003 and attended by more than 65 individuals, primarily engineers of various disciplines and scientists. The attendees used as a starting point various existing compilations of green and sustainability related principles in developing these nine Principles of Green Engineering. Examples of existing principles used in this exercise were the Hannover Principles, Twelve Principles of Chemistry, CERES Principles, and Twelve Principles of Green Engineering proposed (Anastas and Zimmerman 2003). This list, along with the previously described Green Chemistry Principles, is intended

---

[2] The preliminary principles forged at this multidisciplinary conference are intended for engineers to use as a guidance in the design or redesign of products and processes within the constraints dictated by business, government, and society such as cost, safety, performance, and environmental impact.

to be used in tandem as guidance or rules for design and development of green processes and products. The Green Engineering Principles, described below, parallel many of the concepts and approaches covered in the textbook *Green Engineering: Environmentally Conscious Design of Chemical Processes* [5].

1. *Engineer processes and products holistically, use systems analysis, and integrate environmental impact assessment tools.* These concepts resonated in a number of green and sustainable principles and are addressed at length in Allen and Shonnard [5]. Evaluate and reduce the environmental health and safety impacts of designs, products, technologies, processes, and systems on ecosystems, workers, and communities continually and "holistically." Avoid risk shifting (e.g., reducing releases to one environmental medium may increase risk to another medium and/or worker exposure and safety).
2. *Conserve and improve natural ecosystems while protecting human health and well-being.* Reduce, reuse, and recycle the materials used in production and consumption systems, and ensure that residual waste can be assimilated by ecological systems. Rely on natural energy flows. Design processes and products to create cyclical material flows.
3. *Use life-cycle thinking in all engineering activities.* This is an important concept of green engineering. Life-cycle approaches have been widely used by many companies to assess and improve the environmental performance of their products and processes. It is essential that one consider the environmental impacts throughout the product or process life cycle from extraction through manufacturing, use, and disposal.
4. *Ensure that all material and energy inputs and outputs are as inherently safe and benign as possible.* This principle addresses the hazard component of the risk equation. Risk will be minimized if material and energy sources are inherently benign and safe.
5. *Minimize depletion of natural resources.* This principle resonates in many existing lists of principles. It reminds the engineer of the need to reduce and recycle the materials used in production and consumption systems. Conserve energy and improve the energy of internal operations and of the goods and services. Make every effort to use environmentally safe and sustainable energy sources.
6. *Strive to prevent waste.* It is always better and more economical to prevent waste from occurring in the first place.
7. *Develop and apply engineering solutions, while being cognizant of local geography, aspirations, and cultures.* Successful implementation of green engineering solutions can be affected by such factors as availability of resources and geography. It is important that these factors are considered in selecting the green engineering solutions that are most effective for certain localities or regions.
8. *Create engineering solutions beyond current or dominant technologies; improve, innovate, and invent (technologies) to achieve sustainability.* This principle encourages engineers to be creative and innovative in design and development of green engineering solutions. Think "outside the box" for development of green and sustainable technologies.
9. *Actively engage communities and stake-holders in development of engineering solutions.* This is important to ensure that stakeholders are supportive of the engineering solutions.

Engineers and chemists, as designers of products and processes, have a central role in designing chemical processes that have a minimal impact on the environment. Green chemistry and green engineering approaches and tools should be used to design new processes and modify existing processes. Green engineering broadens the scope of engineering design to encompass critical environmental issues and is an important framework for achieving goals of sustainable development.

Pollution Prevention Heuristics for Chemical Processes

Introduction

This section discusses pollution prevention (P2) guidelines and heuristics in chemical process industries useful in process/product development and design stages. The heuristics can also be used for analysis of existing processes but it should be noted that the number of available options is somewhat limited at this stage because the design decisions are already locked in. Therefore, for existing processes, these rules provide some guidelines for maintenance procedures and for making retrofitting decisions.

The outline of this section is as follows: First is a listing and discussion of the general hierarchical rules to be evaluated during the process design stage. Then, because reactors and separators form the heart of a chemical process, specific aspects of P2 heuristics as applied to reactors and separators are discussed in somewhat more detail.

It is appropriate at this point to define what is meant by a heuristic and indicate the pattern of discussions in the other sections of this chapter. A heuristic is commonly defined as a general "rule of thumb," or procedure to arrive at a solution in the absence of a detailed analysis. More specifically, it is a set of rules, often based on common knowledge, used to guide thinking towards a final solution. In our case, the final goal is P2, and heuristics guide us towards achieving this goal.

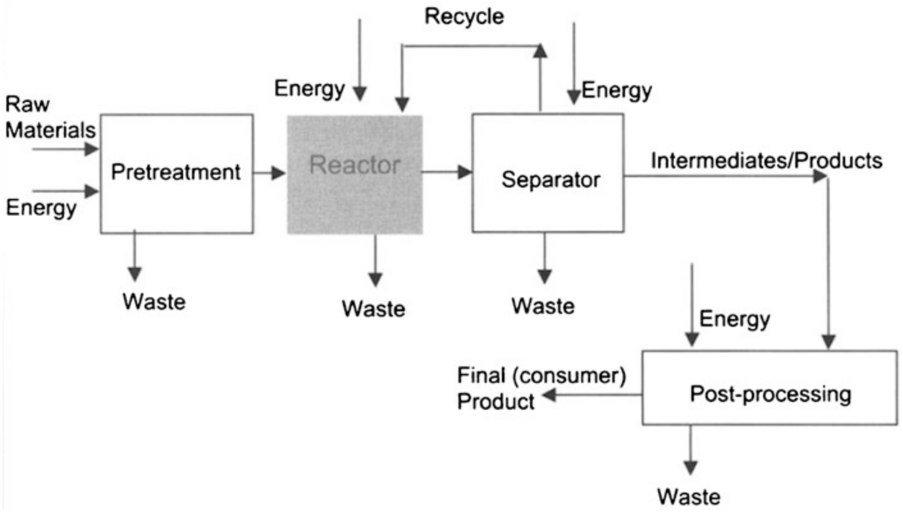


Fig. 5.4 Overall flowsheet showing the various sections for P2 analysis

In this section, the heuristics are often presented as questions rather than as affirmative rules. The rules are then often obvious answers to the questions, and, hence, often no explanation is necessary nor is any provided. However, if an elaboration or clarification is needed, this is provided immediately following the question. The question format of presenting the heuristics is very useful, for instance, in brainstorming sessions during process evaluation and design.

Hierarchical Rules for Waste Minimization

General rules for P2 follow the Pollution Prevention Act of 1990 (42 U.S.C. 13101–13109) and are applicable to any manufacturing activity. This act clearly identifies the waste management hierarchy, and we list them in Table 5.1 for completeness. The P2 Hierarchy specific to a chemical process is derived from this set of principles.

A simplified overall schematic of a chemical process is shown in Fig. 5.4 and forms a block diagram to analyze pollution prevention strategies. A flowsheet can be broken down into the following sections: (1) Pretreatment section, (2) reactor section, (3) separation section, and (4) postprocessing section. The reactor and separator are usually coupled, and, hence, the P2 decisions are often to be made together for this section. Also it may be noted that the blocks cannot be treated in isolation and any design changes made in one section may affect the other. The blocks are, therefore, only convenient partitions for outlining a general strategy. Any decision made for one section has to be reexamined in the overall context, and reconsideration may often be needed. The rule "think outside the box" is very relevant here. With this background setting, we can discuss the hierarchical approach to evaluate P2 opportunities.

Table 5.2 P2 hierarchy for flowsheet analysis or process design

1.	Batch or continuous mode
2.	Input–output structure of the flowsheet
3.	Recycle structure of the flowsheet
4.	Reaction systems
5.	Separation systems
6.	Postprocessing and product sections (e.g., tableting, product drying, etc.)
7.	Energy systems (boilers, cooling towers, heat exchangers, etc.)
8.	Auxiliary equipment (piping, storage tanks, etc.)

The hierarchical rules for analyzing a flowsheet or designing a new process were first formulated by Douglas [6] and then modified by Rossiter and Klee [7], and these form a starting point in this section. The steps are outlined in Table 5.2, which is an extended version of the original suggestions by Rossiter and Klee [7]. The key idea in forming such a hierarchy is that decisions made in earlier steps do not generally affect the design/changes to be made in the later steps and, hence, the rules can be followed on a "top to bottom" basis. This prunes the decision tree and makes the process synthesis task somewhat simpler. Each of the hierarchical steps is now discussed in detail together with the heuristics to be followed for each step.

Batch or Continuous?

The general guidelines are as follows:

1. *What is the annual rate of production?* For capacities of the order of 500 t/year or less, use batch. For capacities of more than the order of 5,000 t/year, use a continuous process.
2. *Does the system involve resource conservation and recovery act* (*RCRA*) *hazardous chemicals?* If so, try to use continuous processes even if the scale is relatively small. Continuous processes generate waste at a constant rate.

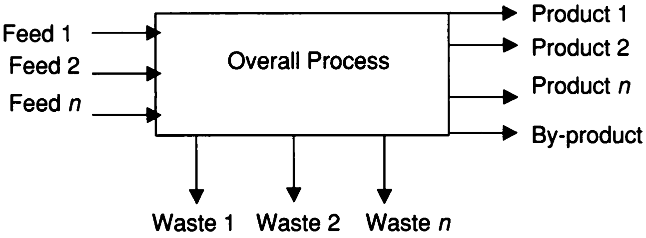


Fig. 5.5 Input–output structure of flowsheet

The composition of the waste is also constant with time and, hence, the treatment is easier. In contrast, batch processes generate wastes intermittently with a large rate of generation for a short period (during the peak point in production). As a result, continuous processes are easier to monitor and control. Batch processes also generate additional waste due to the need for cleaning the equipment between two batches. Hence, the capacity alone may not be the deciding factor in choosing between batch and continuous operation.

It may be noted that raw material substitution is easier in batch processes. Therefore, if one anticipates a varying range of feedstocks then batch processes may have some merit. In a similar note, if the same equipment is used to make a wide range of products (e.g., dyestuffs) then a batch process may be more suitable.

Input–Output Structure

Input–output analysis focuses on the overall structure of the flowsheet, and the recycle streams are not considered here in as much as they do not appear in the overall flow streams. Hence, this structure is essentially an overall mass balance for the entire process. A simplified representation of the input–output structure is shown in Fig 5.5. The main purpose of this analysis is to identify the amount of raw materials used, useful products, and waste formed.

At this stage one can define a process efficiency parameter as:

Process efficiency = MMass of raw ma/Mass of raw materials used which can be used as an indicator of the P2 success scorecard. The goal is to improve the process efficiency, and any changes can be benchmarked against the level of existing efficiency.

The heuristics to be considered at this stage are as follows:

Are any impurities in the feed tied to the waste streams? If so, try source reduction.

Is there a scope for better raw material selection? This would achieve a source reduction which is number 1 in the P2 list given in Table 5.1.

Do the products have a toxic component? Can this be minimized by proper raw material selection?

Can value be added to the waste? Are there any raw materials that are suitable for in-process recycling? In-process recycling is covered under item 2 in the P2 list.

Can wastes be used in the process recycling? Are there waste agencies nearby to take these away? This concern is addressed in item 3 of the P2 hierarchy.

Are wastes generated due to reactor inefficiency? Some guidelines for improving the reactor performance are discussed below under "Heuristics for Green Reactor Design."

Are wastes generated as a result of separation inefficiency? Some guidelines for improving the separation processes are discussed in a later section titled "The P2 Rules for Separations Devices."

Recycle Structure of the Flowsheet

At this level of analysis, the focus is on the recycle streams. The pertinent questions to be addressed at this stage are as follows.

Is there a scope to redirect any output waste streams back to the reactor feed to reduce waste? Can the recycle to reactor be fed directly to the reactor or is a separate reactor needed for this purpose? Often the streams leaving the reactor have trace impurities (formed as by-product in a side reaction) that can be harmful to the catalyst. Hence a separate reactor may be more suitable for such cases. In some cases, the streams can be recycled back to the existing reactor but only after a pretreatment step to remove the trace contaminants and catalyst poisons.

Can any impurity be recycled to extinction? This is possible if the impurities are formed by a reversible reaction while the main reaction is irreversible.

Is the recycle affecting the purge streams? If so consider raw material pretreatment. Nonreactive materials in the feed stream are responsible for the purge and reducing their quantity would reduce the purge. For example, if oxygen-enriched air is used instead of air in a typical partial oxidation process, the quantity of the purge and the associated waste can be reduced.

Can any waste from separation units be reprocessed? Can any waste from separation be reused in any other part of the process?

Has the scope for water recycle been examined completely? Often the possibility of reuse of stripped sour water, wastewater from utility blowdown, etc. exists, and these are often overlooked.

Reaction Systems

Having examined the recycle nature of the flowsheet and the relevant decisions made, the focus shifts to the reaction section. Some of the common issues related to reactor design to be addressed at this stage are as follows:

Are wastes formed from the side reactions? Often the side reactions can be minimized by simple changes in operational procedure. For example, temperature sensitivity of the side reactions can be examined and if the rate of the side reaction increases more with temperature than the main reaction, then lowering the reactor temperature will reduce the waste.

Are the side reactions leading to waste formed in a reversible step of the reaction? If so, there is scope to recycle wastes to extinction.

Is the catalyst used the best available or is there a scope for catalyst replacement?

Is there a scope for solvent substitution? If so, how would it affect the other steps? Additional information on solvent selection is discussed later.

Separation Systems

Having addressed issues related to the overall flowsheet and the reaction system, attention is now focused on the separation systems. Separation systems offer considerable scope for waste minimization. They can be generally classified into the following categories.

1. Gas–liquid separation
2. Gas–solid separation
3. Liquid–liquid separation
4. Liquid–solid separation

General heuristics related to separation systems are indicated here. Detailed heuristics with respect to each individual type of separation are addressed in "Separations Devices."

Some of the key questions to be addressed are the following.

1. *Are any waste streams produced as a result of poor separation?*
2. *Is the separation sequence optimum?* Separations that are easy are done first, followed by the less energy-intensive ones. The final separation step is usually assigned to the most difficult step because the overall quantity of stream to be treated would be minimum by then.
3. *Is the current separation the most suitable or should other methods of separation be considered?* For example, distillation is commonly the workhorse in the chemical industry and is often the immediate choice. Other separation methods, such as pervaporation, can be sometimes more effective especially if the volatility differences are small. Pervaporation is a membrane-based process with the difference that the permeate appears as a vapor, thus permitting solute recovery and recycle. For example, benzene can be recovered from hydrocarbon streams using this method in fairly high concentrations and in a usable form ready for recycle. Many alternative separation methods must be considered, and one should not simply bank on past experience or expertise.
4. *Is the current separation most efficiently designed or are there alternative designs that can reduce energy consumption and lead to less waste?* Examples may be in solvent extraction where a simple switch of solvent may lead to a nonhazardous waste. Also, conventional distillation is often used, but, currently, new design concepts, such as divided wall distillation [8], may prove to be more beneficial for certain processes. Replacing a conventional packing by a more efficient, structured packing may improve the treatment of waste gases in absorption columns.
5. *Is the choice of the mass separation agent (MSA) both cost-effective and environmentally benign? Can any improvements be possible in this choice?* Caffeine extraction from coffee beans is an example. Chlorinated solvents were used as mass separating agents, which posed both a health risk to workers and global harm due to ozone depletion. Supercritical CO_2 and, more recently, supercritical water have replaced these as MSAs.
6. *Have mass integration possibilities been evaluated? Are mass exchange networks (MENs) in place in the flowsheet?* Mass exchange network (MEN) synthesis has been addressed in the book by El-Halwagi [9], and application of this method to water management has been addressed in a paper by Liu et al. [10]. These aspects should be looked at carefully during the process synthesis stage, as they offer considerable scope for solvent or MSA reduction.
7. *Can the separation and reaction sections be suitably combined?* Combining reaction and separation in one unit can have some advantages. It often removes the equilibrium barrier, and, in some cases, complete recycle of unreacted feed is achieved in the same vessel. This question is more important in the design stage of new processes and is somewhat difficult to address in retrofitting existing processes. Membrane separations provide another area where reaction and separation can be combined.
8. *Can any hazardous waste be removed before discharging into a POWT (publically owned water treatment) facility?*

Postprocessing and Product Section

Postprocessing forms an additional important step in the manufacture of many consumer-oriented products, in contrast to commodity chemicals. Examples include tablets, pills, toothpaste, creams, and a wide variety of common products. The chemical industry has traditionally focused on commodity chemicals (process engineering), but, recently, the emphasis has shifted to product engineering. The postprocessing of chemicals to make consumer-used products can often be a major source of pollutants. Hence, careful attention has to be paid to this part of the process as well. Also, this part of the overall plant may involve batch processing, whereas the rest of the process may be continuous processing. Some heuristics are listed below.

Are fines created as a result of product drying? If so, select a different type of dryer. High air velocities or thermal degradation of solids lead to fine formation in drying equipment. These can be avoided by choosing the proper type of dryer depending on the properties of the solid.

Is the blending operation optimized or is it creating some wastes? Can any waste from this operation be recycled?

Is the cleaning operation in between batches optimized? Is there a possibility of wash-solvent recycle here without affecting the overall product quality?

Can process equipments be coordinated to minimize vessel cleaning and to reduce the associated wastes?

How will the products be packaged? Will containers be available for recycle? Are any additional wastes generated as a result of the container reuse policy?

Energy Systems

Energy costs are a major part of the operating costs in chemical processes and, hence, any reduction in energy consumption leads to an increased profitability. Further increased energy use has a direct bearing on the greenhouse gas emissions and, thus, the optimization of the energy systems is also important from a P2 point of view. Some heuristics to be discussed at this stage of analysis are the following:

Is the process fully optimized for energy use?

Are the heat integration networks in place and is there further scope in reducing the energy costs?

Can cleaner fuels be used to reduce SO_2 emissions? If so, at what cost?

Do the steam systems operate at the needed pressure rather than at the available boiler pressure? Too large a pressure leads to a higher condensation temperature for steam. This causes an unnecessarily large temperature differential for heat transfer. Fouling and other problems arise as a consequence, reducing the energy efficiency. In other words, use utilities at the lowest practical temperature. A simple solution to reduce the steam temperature is the use of a thermocompressor. This device uses high-pressure steam to increase the pressure of low-pressure steam to form steam at a desired intermediate pressure.

Are the heat exchangers routinely maintained to reduce fouling? Are the state-of-the art cleaning methods used to reduce sludge formation?

Is the wastewater from high-pressure cleaning of heat exchangers treated separately and not with all the other water streams? These often carry fine particles that provide a large surface area for oil and water to stabilize creating an oily sludge in the wastewater which is difficult to separate.

Can the cooling tower blowdown be reduced? Often pretreatment of fresh water to the cooling tower to reduce calcium salts can be helpful to reduce the scaling and thereby reduce the blowdown. Ion-exchange or even more expensive options, such as reverse osmosis, may prove to be beneficial in this regard depending on the quality (hard or soft) of the feed water.

Auxiliary Equipments

These include pumps, compressor, and storage tanks. The pumps, valves, flanges, etc. can often be a major source of pollution due to fugitive emissions (unintentional release of process fluids). Storage tanks also contribute to pollution by breathing and standing losses. Hence, sufficient attention should be paid to the design and maintenance of this equipment from a P2 standpoint, and this auxiliary equipment should not be taken for granted. The following guidelines are useful.

Can welded pipes be used instead of flanged pipes? This will reduce the fugitive emissions.

Can the total amount of equipment and number of connections be reduced?

Are tanks properly painted? A tank freshly painted will reduce the breathing losses by 50% as shown in an example by Allen and Rosselot [11].

Has the EPA recommended estimates for fugitive emissions done on a periodic basis? Are any improvements noted over a period of time?

Heuristics for Green Reactor Design

Green design of a reactor can be approached in the following hierarchical manner with the top of the hierarchy being the least cost-effective solution.

- Minor modifications in operating conditions and better "housekeeping" practices in existing processes.
- Additional "end-of-the-pipe" clean-up of wastes. (The goal here is to recycle and recover.)
- Major retrofitting of existing processes. Waste reduction and enhanced recycle. Goal is zero emissions and "total recycle."
- Development and installation of more efficient new process technologies that minimize waste and pollution.
- Process intensification concepts and new reactor design concepts.

The last two items in the above list involve considerable R&D and hence present a long-term strategy for waste minimization in reactors. These items should be looked into carefully at the early stage of process design. The first three can be attempted in existing processes.

The minor modifications can be addressed by looking at some of these issues.

A. Source reduction
   - Use nonhazardous raw materials.
   - Use renewable resources.
   - Use benign solvents.
   - Reduce use of solvents.

B. Routine maintenance
   - Temperature control.
   - Pressure control.
   - Vent and relief system tuning.
   - Routine calibration of instruments.
   - Routine cleaning of steam jackets, coils, and other reactor cooling auxiliaries.

C. Operational changes
   - Reduce by-products and generate less waste.
   - Produce products easy to separate.
   - Use heat integration.
   - Evaluate the effect of temperature on side-product formation rate. Can an increase in temperature be beneficial or should one decrease the temperature?
   - Evaluate the role of mass transport for multiphase systems. In some cases an increased mass transfer may favor by-product formation, and reducing the degree of agitation, for instance, may be helpful.
   - Evaluate the role of mixing of reactants. Is a premixed feed better?
   - Evaluate the role of catalyst. Is there a better or more stable catalyst?
   - Is the temperature profile in batch or semi-batch the optimum? Can the temperature peak be minimized?

D. Design changes
   - Recycle unreacted materials.
   - Consider a separate reactor for recycle to minimize catalyst deactivation.
   - Do not overdesign. Overdesign is not only capital consuming but also leads to increased energy costs that translate into greenhouse gases in the environmental context.

The above rules can be applied to an existing process. For new plants it may be necessary to revamp the entire process and use alternative benign production technologies. Some approaches to investigate here are as follows:

Improve atom efficiency.

Improve energy efficiency.

Novel chemistry, e.g., solid-catalyzed routes to replace liquid-phase routes or simple quantitative organic chemistry-based routes.

Novel solvents, e.g., CO_2 expanded systems, ionic liquids, etc.

Novel reactor concepts, e.g., periodic operation, membrane reactors, etc.

Reaction + separation combination, e.g., catalytic distillation, extractive reactions. Biphasic catalysis.

A more detailed description of the role of chemical reaction engineering is available in the paper by Tunca et al. [12], along with some examples and recent trends. Another useful source is the EPA Green Chemistry Web site (www.epa.gov/oppt/greenchemistry).

The P2 Rules for Separations Devices

Some heuristics related to specific methods of separation are as follows.

Distillation Columns

Distillation columns are the workhorse of separations in the chemical industry, and these are very common and offer considerable scope for P2. Some important questions to be addressed in this context are listed below.

1. *Does the system form an azeotrope?* If so, is an entrainer used? Can one replace the entrainer by a more benign agent? Can the azeotrope be broken by other methods such as pervaporation (see paper by Wynn 2001) or membrane separation and then continue with distillation to get the final purity?
2. *Is the reflux ratio in distillation column optimized?* Increase in reflux ratio increases the product purity. However, this causes a larger pressure drop in the column, and increases the reboiler temperature and the reboiler heat duty. Hence, there is a delicate balance among these factors, and often the reflux ratio has to be continuously adjusted to meet the change in feed composition and other day-to-day variations.
3. *Is the feed location optimum*? Simply relocating the feed may cause a more pure product and may be a simple strategy to reduce waste formation. It has both economic and environmental benefits.
4. *Is there a scope for combining reaction and distillation in one piece of equipment*?

Gas–Liquid Separation

These include absorption columns to remove a gaseous impurity as well as stripping columns to remove a VOC from a liquid (usually wastewater). Note that these involve an MSA. For example, for stripping of VOC from wastewater, steam or air is used as the MSA. Hence, appropriate choice of MSA is an important consideration in the context of P2 for gas–liquid separations. Some important heuristics are as follows.

Are the off-gas specifications within the regulations? If not try to switch to more efficient structured packings.

Are the desorption systems at the optimum pressure?

Can the solvent losses be minimized? Should one use alternate solvents that have better heat stability over the repeated absorption–desorption cycle?

Gas–Solid Separations

Gas–solid systems are encountered in a number of processes, such as fluid bed dryers, fine capture from gaseous effluents to meet the PM_5 air quality criteria, etc. Efficient design of these systems will improve the air quality.

Another example of gas–solid separation is the adsorption process. The regeneration of adsorbent is often not complete due to pore diffusion limitations and other factors. Furthermore, the eventual replacement of spent adsorbent leads to solid wastes. Optimization of adsorbent pore structure is one option that can be examined here. The process of regeneration also leads to waste formation and needs to be set at optimum conditions.

Liquid–Liquid Separations

These systems are common in liquid extraction and also in a multiphase reactor with an organic and an aqueous phase. Common sources of pollution are incomplete separation and contamination due to trace organics in the aqueous phase. An example is in alkylation reactions (e.g., *n*-butane reaction with olefins to form isooctanes). Strong acids, such as sulfuric and hydrofluoric acids, are used as catalysts, and the recovery and the recycle of acid need to be optimized in order to reduce the waste generation.

Liquid–Solid Separations

These are encountered in processes such as filtration that involve flow through a membrane. The prevalence of this in the chemical industry has not, however, diminished the challenges in proper design of these systems. A wide range of membranes is available for separation of a solid from a liquid. Fouling is the major challenge in membrane processes and is the most common source of pollution. Membrane systems need periodic cleaning and are another source for pollution generation in terms of waste water. Some heuristics to consider at this stage of design are listed below.

Is the membrane hydrophobic or hydrophilic?

Has the proper selection of membrane tied to the property of the slurry?

Can fouling be reduced by maintaining a higher shear rate at the interface?

In summary, the hierarchical process review and the various heuristics presented here provide a systematic method for identification of waste formation and the appropriate strategies to minimize waste in process industries. Maximum benefit of using this approach is realized at the process development and process synthesis stage followed by the process design stage. Some of the guidelines can also be used in the retrofitting state for existing processes. Additional P2 methods, technologies, and practices are found in the compilation by Mulholland and Dyer [13].

Understanding and Prediction of the Environmental Fate of Chemicals

Introduction

To understand the fate of a chemical once released into the environment requires knowledge of not only how the environment is modeled but also the intimate connection among a chemical's structure, its physicochemical properties, and its behavior under a given set of conditions. The behavior of a chemical in the environment falls under one of two categories: *translocation*, resulting in the movement of the chemical either within a compartment or between compartments, and *transformation*, resulting in the alteration of the chemical's structure. In predicting the persistence of a chemical in the environment, three broad questions are asked:

1. Will it move, and, if so, where will it go?
2. Will its structure be altered?
3. How long will it persist?

The first question is answered by understanding the susceptibility of the chemical to translocation, whereas the second question is answered by understanding the chemical's susceptibility to transformation. Once these two questions are answered, then the third question, requiring rates of translocation and/or transformation, can be answered, thus providing a foundation upon which subsequent assessments of risk to the environment and public health can be performed.

The goals of this section are to introduce methods of modeling chemical movement within and between environment compartments, to define specific translocation and transformation processes, to provide a basic understanding of the association among chemical structure, physicochemical properties, and susceptibility to specific translocation and transformation processes, and to provide methods of accessing and estimating physicochemical properties and environmental fate of chemicals.

Translocation of Chemicals in the Environment

Modeling the Environment

Before assessing how a chemical moves in the environment, the relevant media, or compartments, must be defined. The environment can be considered to be composed of four broad compartments—air, water, soil, and biota (including plants and animals)—as shown in Fig. 5.6. Various approaches to modeling the environment have been described [14–16]. The primary difference in these approaches is the level of spatial and component detail included in each of the compartments.

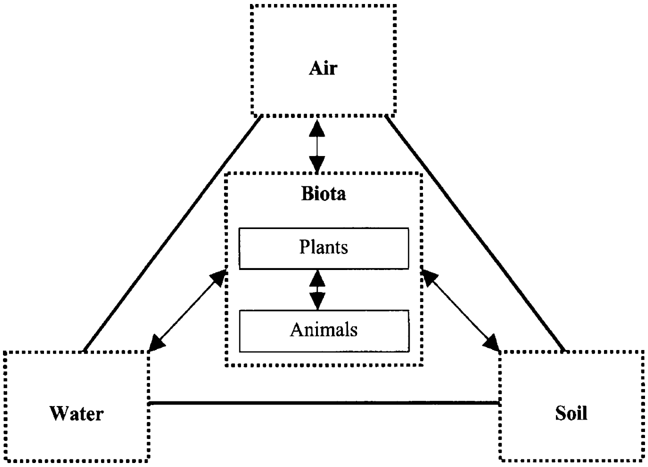


Fig. 5.6 Potential exchange routes of chemicals among the environmental compartments

For example, the most simplistic model considers air as a lumped compartment. A more advanced model considers air as composed of air and aerosols, composed of species such as sodium chloride, nitric and sulfuric acids, soil, and particles released anthropogenically [17]. A yet more complex model considers air as composed of air in stratified layers, with different temperatures and accessibility to the earth's surface, and aerosols segmented into different size classes [16]. As the model complexity increases, its resolution and the data demands also increase. Andren et al. [16] report that the simplest of models with lumped air, water, and soil compartments is suitable for chemical fate evaluations at very large scales (global), whereas the more complex models, with greater levels of segmentation and spatial differences within each compartment, are best suited for problems involving regional to site-specific studies.

Regardless of the model complexity chosen, the transport and partitioning of a chemical between water and air, air and soil, and biota and all compartments must be assessed. Furthermore, each of these compartments may contain different air, liquid, and vapor phases, and chemical partitioning among these phases must be determined. The driving force of transport of chemicals within and between compartments is a difference between chemical potential, the tendency of a chemical to undergo physical or chemical change, in one region compared to the other. When thermal, mechanical, or material equilibrium has been upset in the compartment(s) in question, the chemical moves in response. Despite the awareness that nonequilibrium conditions are at the heart of net chemical transport, modeling of partitioning of chemicals is commonly performed by assuming equilibrium conditions using estimates of liquid–vapor, solid–liquid, and solid–vapor partition coefficients [18]. These coefficients and other properties estimating environmental partitioning, along with properties enabling prediction of environmental transport of chemicals, are briefly discussed in the following sections. For greater detail on transport and fate of chemicals in the environment, the reader is encouraged to refer to numerous resources that are dedicated solely to this topic [19–21].

Translocation Processes in Air

As characterized in Fig. 5.7, once a chemical is released into the atmosphere, it is rapidly transported by the average wind and subjected to dispersion, defined as spreading as a result of thermal or density gradients and/or turbulence, and advection, defined as movement as a result of mass flow in the wind. These processes occur within the planetary boundary layer up to approximately 1 km above ground level [16]. Individual contaminant molecules are also prone to move within air along high-to-low concentration gradients. This type of transport is known as diffusion. These three processes combined result in the same effects: movement of the contaminant away from the source by dilution of the contaminant concentration in the release plume and spreading of the contaminant over a larger area.

While in the air compartment, the contaminant "solubilizes" in the vapor–liquid phase or is associated with aerosol particles by adsorption. It is also prone to desorption from the aerosol particles into the vapor phase. Relevant properties of the air used to model transport of partitioning of a contaminant in the air compartment include temperature, turbulence, wind speed, size and composition of aerosol particles, etc. [16, 19]. Relevant properties of the contaminant that measure its tendency to partition among the vapor, liquid, and solid phases in the air include its aqueous solubility (S_{aq}), vapor pressure (VP), Henry's (K_H), and a variety of coefficients measuring sorption on solids. These properties will be discussed in more detail in subsequent sections.

The net dispersive, advective, and diffusive transport of contaminants in the air is greatly influenced by the degree of deposition to the soil, water, and biota. Deposition occurs in three steps: (1) turbulent diffusion through the surface layer of the atmosphere, (2) diffusion through a laminar sublayer just above the surface, and (3) the ultimate disposition of the chemical on the surface [17]. As discussed previously with the first step, the second step is also affected by properties of the atmosphere and terrain, including turbulence, wind speed, and temperature, along with the size and composition of the particles composing the aerosols [16, 19]. Dry deposition of the contaminant occurs in the absence of precipitation and involves contaminants with and without association with particles. Wet deposition results from condensation of aerosol particles or equilibrium partitioning of the "dissolved" contaminant from the air to the liquid phase. Relevant chemical properties that enable prediction of the tendency of the

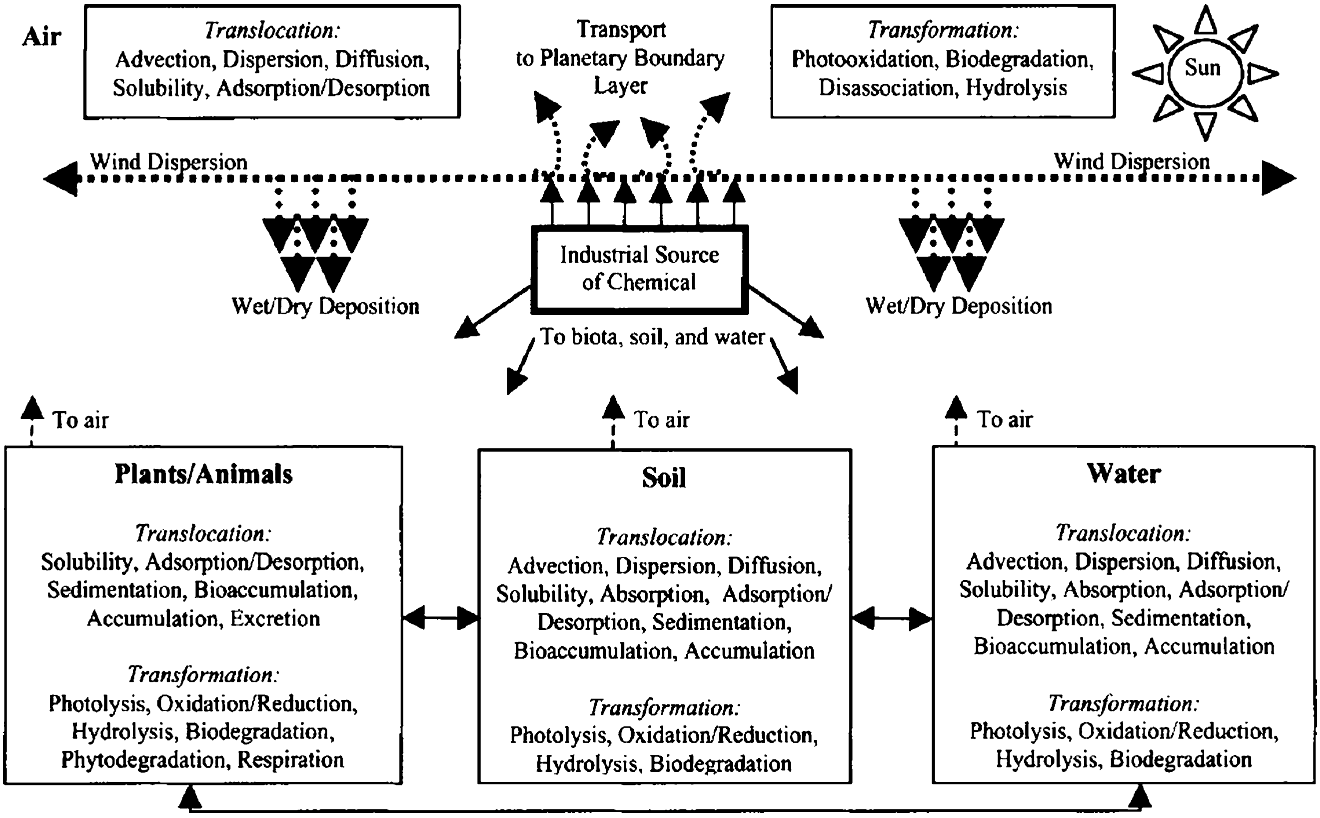


Fig. 5.7 Translocation and transformation processes possible in air, soil, water, and biota (adapted from Bishop [21])

contaminant to undergo deposition include aqueous solubility (S_{aq}), vapor pressure (VP), Henry's constant (K_H), and various sorption coefficients on solids. The last step of deposition is the behavior of the chemical once it has reached the water, soil, or biota surfaces, where it may return to the atmosphere in its original form or in an altered structure.

Translocation Processes in Water

Like the transport of a chemical in air, movement of a chemical in water is governed by the flow characteristics of the water itself (advection) and by the degree of diffusion within the water body. Two different types of diffusion can exist in water bodies. Eddy diffusion (or eddy dispersion) results from the friction caused by the water flow over the sediment or soil bottom surfaces. The vertical and horizontal flow resulting from eddy diffusion is more random and temporal than advective flow and thus extends over a smaller region. Horizontal eddy diffusion tends to be greater than vertical flow, and the contaminant is therefore transported to a greater extent horizontally from a point of discharge [22, 23]. A contaminant is also prone to transport by molecular diffusion, generated by concentration gradients. As a general rule, unless the water body is stagnant and uniform in temperature, molecular diffusion plays a minor role in transporting contaminants in comparison to eddy diffusion [23, 24]. Temperature plays a significant role in determining the degree of eddy and molecular diffusion of a contaminant, particularly in water bodies with stratified layers, such as lakes. As described previously in air translocation processes, the net result of the advective and diffusive transport of a contaminant in water is dilution away from the point of discharge and spreading of the contaminant plume into regions of greater area.

A contaminant can partition from the aqueous phase to solid, air, and biota media, and the presence of each of these media can greatly influence the extent of transport of the contaminant. With a porous solid phase with which the contaminant does not interact, transport of the contaminant is governed by the same laws of mass transport, involving advection and diffusion, that apply in aqueous media free of solids. However, interaction of the contaminant with the solid medium greatly inhibits its movement in the aqueous phase. When contaminant mixtures are present, those with less interaction with the solid medium move along with the aqueous medium, whereas those with greater attraction to the solid medium are retained in proportion to their degree of interaction [23].

As shown in Fig. 5.7, partitioning between the aqueous and solid phases may result from absorption, adsorption/desorption, and sedimentation processes. The contaminant may be taken up into the interior of a solid by means of diffusion in a process known as absorption. As previously described in air–solid partitioning, the contaminant may also be taken up by the surface of the solid, known as adsorption, and its release from the surface of the solid is known as desorption. Adsorption of a chemical to soil or sediment particles may be a result of electrostatic or hydrophobic attraction between the contaminant and the solid surface [23]. Once associated with solid particles, the contaminant may also settle to the sediment surface. This process, known as sedimentation, typically occurs in water bodies with laminar flow, such as a wetland. The increase in contaminant concentration in the solid phases as a result of any of these partitioning processes is known as accumulation [14].

The contaminant's aqueous solubility and density greatly influence its final disposition in water–solid systems. Dense nonaqueous phase liquids (DNAPLs) are chemicals with densities greater than water and typically low aqueous solubilities. DNAPLs naturally partition away from the aqueous phase and towards the solid phase, often pooling on top of an impermeable solid layer. On the other hand, light nonaqueous phase liquids (LNAPLs) possess densities less than water and have a tendency to pool on the water's surface where they may be prone to volatilization and photolysis reactions. Various liquid–solid partition coefficients have been developed to predict a contaminant's tendency to associate with the solid phase. These include the octanol–water partition coefficient (K_{ow}), the soil–water distribution ratio (K_{d}), the organic carbon partition coefficient (K_{oc}), the organic matter–water partition coefficient (K_{OM}). These will be discussed in detail in sections to follow.

Transport of a contaminant from water to air is influenced primarily by wind velocity [16]. The contaminant's density, vapor pressure, and aqueous solubility also factor into its tendency to be introduced into the air phase, and its Henry's constant (K_{H}) provides a good indication of this tendency. Biota have a strong attraction to hydrophobic contaminants, and, as a result, uptake of contaminants, by partitioning into plants and animals, known as bioaccumulation, has been reported to be a dominant mechanism of removal [16, 25]. The tendency of a chemical to be taken up into biota is quantified by the bioconcentration factor (BCF), as measured by the ratio of its concentrations in biota and water.

Translocation Processes in Soil

Contaminants in the soil compartment are associated with the soil, water, air, and biota phases present. Transport of the contaminant, therefore, can occur within the water and air phases by advection, diffusion, or dispersion, as previously described. In addition to these processes, chemicals dissolved in soil–water are transported by wicking and percolation in the unsaturated zone [26]. Chemicals can be transported in soil air by a process known as barometric pumping that is caused by sporadic changes in atmospheric pressure and soil–water displacement. Relevant physical properties of the soil matrix that are useful in modeling transport of a chemical include its hydraulic conductivity and tortuosity. The diffusivities of the chemicals in air and water are also used for this purpose.

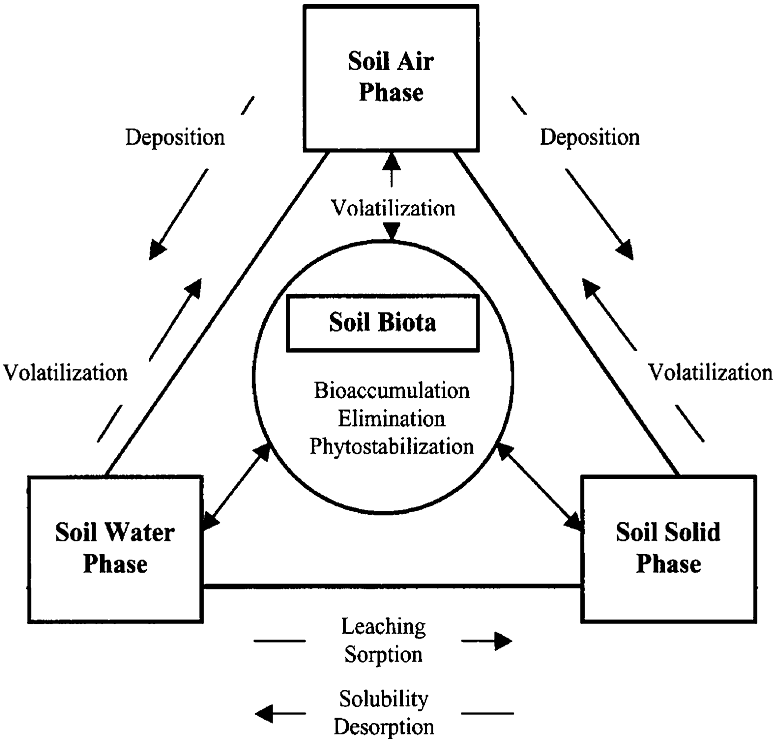
As shown in Fig. 5.7, any of the translocation processes described previously in air and soil compartments may also occur in soil. Figure 5.8, adapted from Baum [26], shows the potential intermedia exchanges that occur among the soil, air, water, and biota phases within the soil compartment. Exchange between air and water and air and solid phases may involve volatilization or deposition. Interchange between the water and solid phases may involve leaching (movement of the chemical in water through the soil column), absorption into the solid matrix, adsorption onto the solids, or desorption into the aqueous phase. Partitioning from any of the solid, water, or air phases to biota results in the bioaccumulation of the contaminant. Biota can use various means of elimination of the contaminant or its metabolites. Volatilization of the contaminant or its metabolites from plants to the air phase may occur, as well as excretion of the contaminant or its metabolites from roots or foliage to the surrounding soil and water phases in the root zone. The root zone may also serve to "stabilize" the contaminant in the soil, a process known as phytostabilization [27], resulting in accumulation of the contaminant in the soil phase. Relevant physicochemical properties of the contaminant that provide an indication of the degree of these partitioning processes include aqueous solubility (S_{aq}), vapor pressure (VP), Henry's constant (K_{H}), soil partitioning coefficients, and the BCF.

Translocation Processes Involving Biota

As mentioned previously and shown in Fig. 5.8, movement of a chemical in a system containing plants or animals may involve exchange with the air, soil, and water phases. Bioaccumulation results when the plant or animal uptakes the contaminant. In plants, the contaminant may be, in turn, released to the air by means of volatilization or to the soil with subsequent accumulation by phytostabilization, adsorption, and sedimentation or transport to the aqueous phase by advective, diffusive, or dispersive processes. Excretion of the contaminant or a metabolite from animals may also be received by any compartment.

Partitioning between a plant and air depends on the properties of the chemical, such as vapor pressure (VP), properties of the plant, and the temperature. Andren et al. [16] report that partitioning from the air to plants occurs primarily through the foliage, and only compounds with mobility in the phloem can partition to the stem and trunk

Fig. 5.8 Intermedia translocation processes involved in the soil compartment (adapted from Baum [26])



of the plant. As a general rule, hydrophilic compounds have higher phloem mobility, and, therefore, the octanol–water partition coefficient (K_{ow}) is a good indicator of the tendency of a chemical to be transported into a plant via its foliage. Uptake and release to/from the water and soil phases and the relevant parameters used to measure the tendency for these processes were discussed in the previous section. Of all the compartments of the environment, the least information is known concerning the translocation processes involving biota [16]. Also, the BCFs cannot be considered equilibrium constants as the other partitioning coefficients are because biota are "reactive sorbents" where the contaminant may be transformed upon uptake. These and other transformation processes involved in each of the environmental compartments are discussed in the following sections.

Transformation of Chemicals in the Environment

The primary mechanisms of degradation of chemicals in soil, water, sediment, air, and biota environments are classified as biotic (biodegradation, phytodegradation, and respiration) or abiotic (hydrolysis, photolysis, and oxidation/reduction), as shown in Fig. 5.7. Biodegradation, the transformation of chemicals by microorganisms, has potential to occur in any environmental compartment that contains moisture to support microbial processes; therefore, it can occur in all compartments, including in the root zone of plants. Phytodegradation, the transformation of chemicals by green plants, and respiration, the metabolism of chemicals by animals, occur wherever these biota reside, and, as discussed previously, they interact with each of the other compartments of the environment. Hydrolysis, the reaction with water, can occur in any compartment containing water, whereas nonbiological oxidation/reduction reactions are most dominant in the air phase [21]. Photolysis, the reaction of a chemical with ultraviolet radiation from the sun, occurs on surfaces exposed to sunlight, including soil, water, air, and plants. Compounds not susceptible to any of these degradation processes will have a tendency to be persistent in the environment. These reactions and their kinetics are discussed in more detail below.

Biotic Transformation Processes

Biodegradation. Because of the abundance and species diversity of environmental microorganisms and their ability to adapt to many different conditions and to degrade a wide range of substrates, biodegradation plays a major role in the transformation of contaminants [28]. The rate of biodegradation is influenced by numerous factors, including availability of nutrients, pH, temperature, level of oxygen, moisture content, and chemical structure. The degradation of the contaminant is catalyzed by enzymes, and most biodegradation processes are composed of a series of steps mediated by different enzymes. If the chemical is completely degraded to carbon dioxide, methane, water, and other inorganic compounds, it is said to be mineralized.

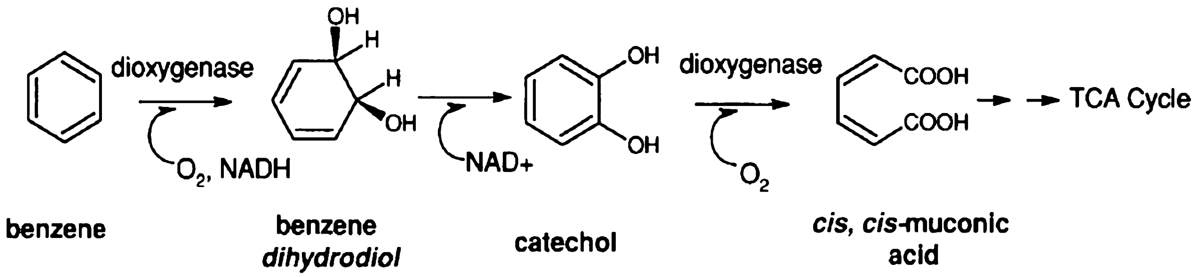


Fig. 5.9 The *ortho*-cleavage pathway of benzene oxidation by dioxygenase enzymes [29]

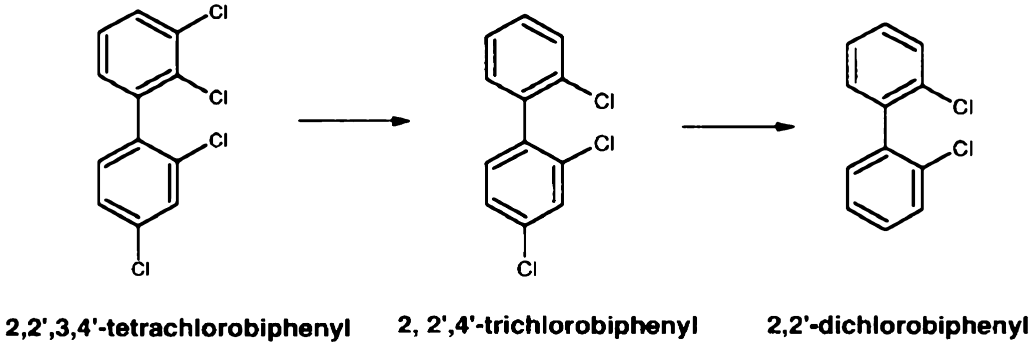


Fig. 5.10 Pathway of reductive dehalogenation of PCBs [31]

However, the contaminant is often converted to a chemical structure more complex than these mineralization products, and, in many cases, these incomplete degradation products, called daughter products or metabolites, are more toxic than the parent compound.

Some chemical compounds serve as a source of carbon and/or energy for the degrading microorganism, whereas in other cases they are not but are, rather, “cometabolized.” For example, methanotrophs, or methane-oxidizing bacteria, are known to oxidize aliphatic and aromatic compounds by cometabolism as long as methane is available to provide a source of carbon and energy. These bacteria are obligately aerobic because oxygen is necessary for their growth and metabolism; however, many other types of microorganisms cannot function in the presence of oxygen, and they are known as obligately anaerobic bacteria. Furthermore, bacterial populations exist that can function with or without oxygen, and they are known as facultative aerobes or facultative anaerobes.

Biodegradation reactions can be classified as aerobic or anaerobic, with oxidations and reductions dominating within these two classes, respectively. Oxidations occur in aerobic pockets in groundwater aquifers, near the soil surface, and aerobic zones of lakes and streams, for example. Dioxygenase and monooxygenase enzymes mediate oxidation reactions, yielding hydroxylated metabolites and possibly ring-cleavage products resulting from the activity of ring-cleaving diooxygenase enzymes. Figure 5.9 shows one of the known pathways of oxidation of benzene involving a consortium of microorganisms expressing diooxygenase enzymes that cleave the benzene ring between the two hydroxyl groups of the catechol, called *ortho*-cleavage. As implied by the TCA cycle endpoint in Fig. 5.9, the microorganisms involved in this pathway gain energy from the degradation of benzene. It is important to note that typical intermediates of biologically mediated oxidation reactions are more polar than the parent compound, thus with different environmental behavior.

More oxidized compounds, such as chlorinated benzenes, are susceptible to biologically mediated reduction in environments under anaerobic conditions, such as in lake and river sediments. It is known that highly polychlorinated biphenyl (PCB) congeners, for example, are susceptible to reductive dehalogenation, the result of the interaction of syntrophic microbial communities that are active under methanogenic and sulfate-reducing conditions [30]. Previous studies of Hudson River sediments have reported that the more chlorinated PCB congeners (with more than 3–4 chlorine atoms) are transformed to lesser chlorinated congeners by a series of dechlorination steps where the chlorine atoms are replaced by hydrogen atoms [31]. As shown in Fig. 5.10, preferential removal of the chlorine atoms in the *meta*- and *para*-positions on the ring occurs before the *ortho*-chlorines are removed, resulting in accumulation of the *ortho*-substituted congeners in the environment [32].

As a result of a large body of studies, heuristics of biodegradation have been developed. Table 5.3 provides

Table 5.3 Chemical characteristics conferring susceptibility to biodegradation

Nonaromaticity
Unsaturated bonds
Straight chains of alkanes (greater than 9 carbons)
Soluble in water
Increased number of halogens (anaerobic conditions)
Decreased number of halogens (aerobic conditions)
Alcohol, aldehyde, and carboxylic acid functional groups
Ortho- or *para*-substitution patterns on a benzene ring

Adapted from Bishop [21]

general guidelines for prediction of the tendency of a chemical to be biodegraded [21]. More detailed information about the specific biodegradation pathways and kinetics for individual groups of compounds and microorganisms can be found in numerous references [33–35].

Phytodegradation. The transformation of contaminants by plants is believed to play a major role in contaminant removal from the biosphere, particularly considering that plants cover approximately 146,000,000 km^2 of terrestrial surface [28]. Plants generally degrade chemicals by first uptaking them and subsequently transforming them to products that are conjugated or bound to the cell wall or stored in vacuoles [36]. Some plants use reductive dehalogenation and oxidation transformation pathways of degradation similar to those pathways followed by microorganisms in biodegradation. Compound classes that have been reported to be degraded in aqueous systems by plants include halogenated hydrocarbons [37, 38], nitroaromatics [39, 40], organophosphate pesticides [41], and polycyclic aromatic hydrocarbons [42, 43]. At this time, there are few rules of thumb to enable prediction of phytodegradation of contaminants because the study of phytodegradation is relatively recent.

Abiotic Transformation Processes

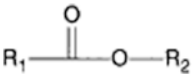
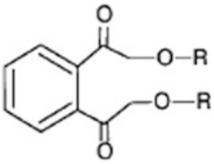
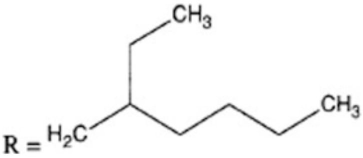
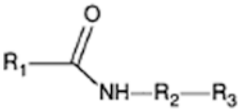
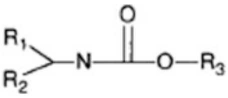
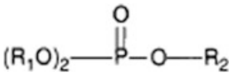
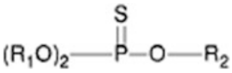
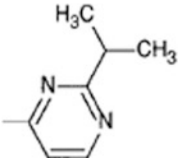
Hydrolysis. Compounds that possess an electrophilic atom (electron-poor) have a tendency to undergo hydrolysis reactions with electron-rich water or hydroxide ion (OH^-). The reaction forms a product with a hydroxyl group replacing a group that leaves the parent compound. The products are more polar than the parent compound, thus indicating different environmental behavior that must be evaluated, as also reported for products of biological oxidation. Hydrolysis is observed anywhere water is present, thus indicating that it can occur in any environmental compartment, given the presence of compounds susceptible to this reaction. This is considered a dominant pathway of transformation of compounds with hydrolyzable groups in aquatic systems [28]. These reactions are extremely pH^- sensitive, as they are often catalyzed by H^+ or OH^- ions. For example, Bishop [21] reports that the rate of hydrolysis of insecticide carbaryl increases logarithmically with pH, where the rate at pH 9 is ten times greater than at neutral pH and 100 times greater than at pH 6.

Numerous functional groups have been reported to be hydrolyzable, and examples of compound classes that are susceptible to hydrolysis are provided in Table 5.4. The chemical structure of the compounds greatly affects the rates of hydrolysis as shown in Table 5.4, with half-lives measured at pH 7 and 25°C ranging from seconds to thousands of years. A detailed discussion of degradation of contaminants via hydrolysis can be found in Schwartzenbach et al. [44], Larson and Weber [23], and Wolfe and Jeffers [45]. Ney [14] has set a low range of hydrolysis half-lives falling below 30 days and a high range above 90 days. As shown in Table 5.5, compounds considered to be hydrolyzed rapidly are not prone to bioaccumulation, accumulation, food-chain contamination, or adsorption, and are not generally considered persistent, whereas the opposite is true for compounds not hydrolyzed rapidly. Half-lives of various contaminants undergoing hydrolysis reaction have been reported to be significantly higher than those undergoing microbial transformation and phytotransformation [27]. For example, tetrachloroethylene (PCE), a common solvent used in dry cleaning, possesses a half-life of over 10 years if subjected to hydrolytic conditions at neutral pH. In contrast, PCE's half-lives due to microbial transformation (using a density of 10^5 organisms per liter) and phytotransformation (using *Spyrogyra* spp. at a density of 200 g wet weight per liter) have been reported to be 35 and 5 days, respectively.

Photolysis. Photochemical transformations of chemicals, also known as photolysis, result from the uptake of light energy (quanta) by organic compounds. These reactions can occur in the gas phase (troposphere and stratosphere), the aqueous phase (atmospheric aerosols or droplets, surface waters, land–water interfaces), and the solid phase (plant tissue exteriors, soil and mineral surfaces) [23]. Photolysis may be direct, where the structure of the chemical absorbing the light energy is "directly" transformed, or indirect, where photosensitizers, such as quinones and humic acids, absorb light energy and then transfer it to a contaminant whose structure is altered in the process [46]. Chemical reactions that affect structural change of photochemically excited contaminants include fragmentation, intramolecular rearrangement, isomerization, hydrogen atom abstraction, dimerization, and electron transfer from or to the chemical [44].

Functional groups such as unsaturated carbon–carbon bonds and aromatic rings lend greater susceptibility to photolytic reactions, and compound classes such as nitrosamines, benzidines, and chlorinated organics are more readily photolyzed [28, 46]. In order for photolysis to be

Table 5.4 Comparison of hydrolysis half-lives of representative chemicals in classes susceptible to hydrolysis at pH 7 and 25 °C [14, 23, 44]

Hydrolyzable compound class	Representative compound(s)	Half-life, $t_{1/2}$
Monohalogenated hydrocarbons R–X	CH_3Cl $(CH_3)_2CHCl$ $(CH_3)_3C–Cl$	340 days 38 days 23 seconds
Polyhalogenated hydrocarbons $R–X_n$	CH_2Cl_2 $CHCl_3$ $CHBr_3$ $BrCH_2–CH_2Br$	700 years 3,500 years 700 years 4 years
Carboxylic acid esters 	$R_1 = CH_3–, R_2 = –CH_2CH_3$ $R_1 = CH_3–, R_2 = –C(CH_3)_3$ $R_1 = CH_3–, R_2 = C_6H_5–$	2 years 140 years 38 days
Dialkyl phthalates 		100 years (pH 8)
Amides 	$R_1 = CH_3–, R_2 = –H, R_3 = –H$ $R_1 = CH_3–, R_2 = CH_3–, R_3 = –H$ $R_1 = –CH_2Cl, R_2 = –H, R_3 = –H$	4000 years 4,000 years 40,000 years
Carbamates 	$R_1 = CH_3–, R_2 = CH_3–$ $R_3 = –CH_2CH_3$ $R_1 = –H, R_2 = C_6H_5–, R_3 = –C_6H_4NO_2$	1.5 year 50,000 years 25 seconds
Phosphoric acid triesters 	$R_1 = CH_3–, R_2 = CH_3–$ $R_1 = C_6H_5–, R_2 = C_6H_5–$	1.2 years 320 days
Thiophosphoric acid triesters 	$R_1 = CH_3CH_2–, R_2 = C_6H_5NO_2–$ $R_1 = CH_3CH_2–, R_2 =$ 	89 days 23 days

considered to be a significant gas-phase destruction mechanism for a chemical, it must absorb light energy beyond a wavelength of 290 nm [28]. Specific compounds that do absorb greater than 290 nm include 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (2,3,7,8-TCDD), 1,2,3,7,8-pentachloro-*p*-dioxin, octachloro-*p*-dioxin, dibenzo-*p*-dioxin, hexachlorobenzene, and various polychlorinated dibenzofurans, and polychlorinated and polybrominated biphenyls [40, 47–51]. The kinetics of photolysis of herbicides in natural waters and soil surfaces were compared by Konstantinou et al. [52], who reported faster rates in soils than in lake, river, marine, and ground waters and significant enhancement of rates in the presence of increasing dissolved organic matter. Applying the same general rule as described for hydrolysis (Table 5.5), if the photolytic half-life of a contaminant is less than 30 days, it is not considered to be persistent; however, if the photolysis half-life is greater than 90 days, the contaminant poses risk to accumulation, bioaccumulation, and food-chain contamination [14].

Oxidation/reduction reactions. Reactions of chemicals via abiotic oxidation or reduction involve a transfer of electrons and result in a change in oxidation of the state of the product compared to its parent compound. As a general rule, reduction reactions are prevalent in soil sediments, while oxidation reactions are more important in surface waters and in the atmosphere [28].

The contaminant that is oxidized serves as the donor of electrons that are transferred to oxidizing agents including O_2, ferric (III) iron, and manganese (III/IV) in aquatic/soil systems and ozone (O_3), hydrogen peroxide (H_2O_2), and free radicals, such as the hydroxyl radical, in the atmosphere. Reduction reactions result in the transfer of electrons from an electron donor to the contaminant. Examples of electron

Table 5.5 Heuristics for hydrolysis half-lives and susceptibility of contaminants to translocation and transformation processes [14]

Process	Rapid hydrolysis ($t_{1/2} < 30$ days)	Slow hydrolysis ($t_{1/2} > 90$ days)
Accumulation	Not likely	Yes
Bioaccumulation	Not likely	Yes
Food-chain contamination	Not likely	Yes
Persistence	Negligible	Yes
Adsorption	Negligible	Maybe
Dissipation	Yes	Negligible to slowly

Table 5.6 Useful properties of the chemical, the medium, and transport/kinetics for assessment of environmental fates of contaminants [14, 18, 53]

Chemical	Medium	Transport and kinetics
Traditional Boiling point, T_b Melting point, T_m Density, ρ Surface tension, v Vapor pressure, VP Aqueous solubility, S_{aq}	*Temperature* *pH* *Biota present* *Light present* *Air*: Wind velocity, turbulence, stratification, composition, etc.	Diffusivity in air and water Phase transfer coefficients (air–water, air–soil) Half-lives of biotic and abiotic reactions
Specialized Henry's constant, K_H Octanol–water partition coefficient, K_{ow} Bioconcentration factor, BCF Soil–water distribution ratio, K_d Organic matter–water partition coefficient, K_{om} Organic carbon–water partition coefficient, K_{oc}	*Water*: Number of active cells, redox potential, activity, velocity, etc. *Soil*: Hydraulic conductivity, tortuosity, composition, etc.	

donors present in the environment include pyrite (FeS), ferrous carbonates, sulfides, and natural organic matter. As described previously for biologically mediated reduction reactions, highly oxidized contaminants, such as PCE and highly chlorinated PCBs, can be reduced abiotically provided that they are in intimate contact with electron donors. Unfortunately, the kinetics of abiotic oxidation/reduction reactions have not been described for many contaminants [28].

The Connection Between Chemical Properties and Environmental Fate

As indicated previously, the final disposition of a chemical in the environment is dependent on the environmental conditions, characteristics of the media involved, and the various physicochemical properties of the contaminants. Table 5.6 provides a listing of properties describing the chemical, medium, and potential for translocation and transformation. This section focuses on chemical properties that are frequently used to assess the fate of a contaminant in air, water, soil, or biota. The goals of this section are to provide brief descriptions of relevant properties of a contaminant and to directly link specific ranges of these properties to predicted fate in the environment.

Traditional Chemical Properties

Traditional chemical properties including boiling point, T_b, melting point, T_m, density, ρ, surface tension, γ, vapor pressure, VP, and aqueous solubility, S_{aq}, have use in many systems beyond environmental applications; however, they provide strong indication of partitioning among the air, water, soil, and biota compartments in the environment. When combined with more specialized physicochemical properties to be discussed below, these traditional chemical properties can provide a very powerful means of answering the basic questions of where will a chemical go, what reactions will it undergo, and how long will it persist in the environment.

A chemical's T_b, the temperature at which its vapor pressure equals the ambient pressure, and T_m, the temperature at which its solid and liquid forms are in equilibrium at ambient pressure, are easily located in references and databases. As a result, many of the correlations that have been constructed for property estimations use these parameters as independent variables. The T_b of a chemical can nonetheless provide an indication of the partitioning between gas and liquid phases [53], with the higher values denoting a lower tendency to exist in the vapor phase. The surface tension, γ, of a chemical, the ratio of the work done to expand the surface divided by the increase in the surface area, is often used to estimate the VP of liquids in aerosols and in soil capillaries [28]. The VP of a chemical is the pressure of a pure chemical vapor that is in equilibrium with the pure liquid or solid, and it provides an indication of the tendency of a chemical in its pure liquid or solid phase to volatilize. Chemicals with high vapor pressures will be likely to escape to air and thus exist in higher concentrations in the air phases of environmental compartments than those with low vapor pressures. Ney [14] suggests that the VP of a chemical is considered high if greater than 0.01 mmHg, whereas low VP is considered to be below 10^{-6} mmHg. The S_{aq} of a chemical is another property that can be directly used to assess translocation and transformation potential of a chemical. By definition, the S_{aq} of a chemical is its concentration in a saturated water solution. A chemical possessing high S_{aq}, defined by Ney [14] as greater than 1,000 ppm, is more likely to be mobile in the aqueous environment than a chemical with a low S_{aq}, defined as less than 10 ppm.

Specialized Chemical Properties

Chemical properties that reflect the tendency of a chemical to partition between phases have been constructed specifically for environmental fate applications. These specialized properties are known as partition coefficients. The parameter that is a measure of a chemical's tendency to partition between water and air is known as the dimensionless Henry's constant (K_H), determined by the ratio of the equilibrium concentrations of the chemical in air and in water, respectively. The Henry's law constant (H), describing the ratio of the chemical's partial pressure in solution to its concentration in solution at equilibrium, is often used to describe chemical partitioning between air and water. K_H is simply the ratio of H to the product RT, where R is the gas constant and T is the temperature. Compounds with high H values (greater than 10^{-1} atm m^3/mol) will prefer the air phase and tend to volatilize from the aqueous phase in water, soil, air, and biota compartments, whereas compounds with low H values (less than 10^{-7} atm m^3/mol) prefer to escape the air phase into the aqueous phase [53].

Partitioning of a chemical between the aqueous and nonaqueous phases is frequently measured by the octanol–water partition coefficient, K_{ow}. K_{ow} of a chemical is the ratio of its concentration in *n*-octanol, frequently used to represent a model phase of living and nonliving natural organic material, to its concentration in water. This parameter which ranges from 10 to 10^7 is a measure of the partitioning of a chemical between hydrophobic and hydrophilic phases in all environmental compartments. A high K_{ow} value (characterized as greater than 1,000) indicates that a chemical has a tendency to be less mobile in the aqueous phase, more greatly adsorbed to solids, and bioaccumulate [14, 16, 53]. K_{ow} values are frequently used to determine another parameter, the BCF, which is the ratio of a chemical's concentration in the tissue of a living organism to its concentration in water. A chemical with a high BCF (greater than 1,000) will be more likely partitioned from water into the fatty tissues of fish and humans, for example, and thus is more likely to bioaccumulate and yield food-chain contamination effects. Chemicals with low potential for bioaccumulation typically have BCF values less than 250 [53].

Various coefficients are helpful in measuring the potential of a chemical to partition between the aqueous and solid phases. These parameters are valuable in predicting the potential of a chemical to adsorb to the solid phase. The soil–water distribution ratio, K_d, of a chemical is the ratio of its equilibrium concentration sorbed onto a solid phase to its equilibrium concentration in solution. Because most sorption of neutral, nonpolar chemicals occurs primarily on the organic matter in soil and sediments, the organic matter–water partition coefficient, K_{om}, is frequently used as a measure of a chemical's tendency to partition from water to solid phases. The useful parameter of K_{oc}, the organic carbon–water partition coefficient, has been introduced to describe the ratio of the equilibrium concentration of a chemical associated with the organic carbon content of soil to that associated with the water phase. K_{oc} is roughly 1.724 times greater than K_{om} because the organic matter sediment is roughly 1.724 times larger than the organic carbon content [28]. With all of these partition coefficients, higher values indicate that a chemical has a greater tendency to leave the aqueous phase and sorb to soils and sediments. Log K_{oc} values of chemicals are considered high when greater than 4.5 and low when less than 1.5 [53].

Sources of Chemical Property and Fate Data

Handbooks. Tabulations of traditional properties of chemicals (T_b, T_m, ρ, S_{aq}) are often available in publications by fine-chemical manufacturers (e.g., Fisher Scientific, Web page: www.fishersci.com). Other collections of these parameters, surface tension, the specialized partition coefficient values, and kinetics of transformations include Windholz et al. [54], Howard et al. [55], Dean [56], Lide [57], Mackay et al. [58], Howard and Meylan [59], Tomlin [60], Yaws [61], and Verschueren [62].

Software and online sources of chemical properties. Larson et al. [28] and Andren et al. [16] provide a thorough listing of electronic databases and online search engines for environmental fate properties. Very useful online databases for these parameters include Syracuse Research Corporation's (SRC's) (Syracuse, NY) Environmental Fate Database (EFDB) and the Hazardous Substances DataBank available at http://toxnet.nlm.nih.gov/. The SRC Web site also provides access to its DATALOG database that provides a literature search engine for numerous chemical properties. In addition, direct photolysis rates and half-lives of contaminants in the aquatic environment can be obtained from the US EPA's GCSOLAR software, available at http://www.epa.gov/ceampubl/swater/gcsolar/. A very useful software for estimating physical/chemical and environmental fate properties is the EPI Suite software [1].

Estimation methods. Numerous references focus on the theory and application of estimating physicochemical properties of chemicals. Lyman et al. [63], Neely and Blau [64], Howard and Meylan [59], Baum [26], and Allen and Shonnard [53] provide thorough descriptions of the methods available for manually calculating chemical properties.

Heuristics for Predicting Environmental Fate

Once the properties and fate data are obtained for a chemical, its general disposition in the environment can be estimated. Many of the properties described previously have been divided into ranges from low to high that enable estimation of tendencies of a chemical to undergo various translocation and transformation processes. Table 5.7 summarizes expected behavior of chemicals using values of VP, S_{aq}, K_{ow}, and K_{oc}. These findings are summarized below with general heuristics for predicting environmental fate of chemicals. These general trends in environmental fate can be combined with transformation kinetics to provide an assessment of exposure for any chemical.

Environmental Performance Assessment for Chemical Process Design

Introduction

The chemical industry contributes significantly to economic development, yet faces many environmental and societal challenges that require a rethinking of traditional approaches in the commercialization of processes and products. The generation of toxic, hazardous, and global change by-products of chemical processes is but a few examples of these challenges. A more comprehensive evaluation of economic, environmental, and societal consequences, at times spanning the entire product or process life cycle, is needed to achieve sustainable growth. Coincident with this will be an enlargement in the data and computational requirements for these assessments. Computer-aided analysis tools will therefore be needed to efficiently link process/product design with critical environmental and societal impacts in a larger systems analysis.

Fortunately, much progress has been made recently in developing environmental assessment methods and computer-aided tools to accomplish these goals. Table 5.8 shows a description of key factors for the environmentally conscious design of chemical processes for which computer-aided tools are useful. These tools fall into two categories: those providing information on environmental fate and impacts and those intended to improve process environmental performance. As will be shown, some tools provide information early in design whereas others are employed at later stages. These environmental factors and assessment tools should be incorporated into the design of chemical processes and products as illustrated in Fig. 5.11. The computer-aided tools in Fig. 5.11 include what has thus far been employed in traditional design: process simulation, design heuristics, and optimization. Added to these conventional design tools is a set of environmental evaluation methods and tools that inform the design activity on a range of potential impacts. The Green Engineering approach uses these environmental assessments in a hierarchical fashion during process and product design. A hierarchical approach for evaluating environmental performance during process design will be described in a later section.

Overview of Environmental Issues

Before beginning a series of case study evaluations, we will need to establish a set of environmental performance measures to use in design. These performance measures must reflect societal attitudes on the importance of several environmental impact categories. This is naturally a subjective judgment, but a consensus is emerging in the professional literature and in regulatory agencies on a set of environmental "midpoint" indicators for this purpose [53, 65–69]. Table 5.9 features several environmental impact categories, a description of the causes, and midpoint/endpoint effects.

Environmental Impact Assessment: Tier I–Tier III

Figure 5.12 illustrates a design hierarchy in the synthesis and evaluation of chemical processes, integrating the design approach of Douglas [70] with environmental risk assessment methods [53]. Environmental evaluation progresses from simple ("Tier I") assessments early in design when process information is limited to detailed flowsheet evaluations later in the design process. Different assessment approaches are needed in early design as opposed to those applied later in flowsheet evaluation. In early design, a large number of design choices are considered, and a simple and

Table 5.7 Rules of thumb of chemical property ranges guiding prediction of environmental fates [14, 53]

Property	Activity predicted	Heuristic
VP (mm Hg) Low: $<10^{-6}$ Medium: 10^{-6}–0.01 High: 0.01	Translocation Volatilization Accumulation Bioaccumulation Adsorption	1. Chemicals with high VP are more likely to volatilize into the air phase and less likely to accumulate, bioaccumulate, and adsorb to solids
S_{aq} (ppm) Low: <0.1 Medium: 0.1–10,000 High: $>10,000$	Translocation Mobility in water and soils Accumulation Bioaccumulation	2. High S_{aq} confers a greater tendency for a chemical to be mobile in the aqueous environment and is less likely to accumulate, bioaccumulate, volatilize, and be persistent
	Volatilization Adsorption	3. Chemicals with high S_{aq} are prone to biodegradation and respiration processes
	Transformation Biodegradation Respiration	4. Chemicals with low S_{aq} have a greater tendency to be immobilized via adsorption and less likely to be leached in soil
H (atm m^3/mole) Low: $<10^{-7}$ Medium: 10^{-7}–10^{-1} High: $>10^{-1}$	Translocation Mobility in air/water Volatilization	5. Chemicals with high H values have a greater tendency to escape the aqueous phase into the air phase, are less mobile in the aqueous environment, and less biodegradable but more mobile in the air environment
	Transformation Biodegradation	6. Chemicals with higher molecular weight tend to experience a decrease in both VP and S_{aq}. Thus, VP is a better measure of tendency to volatilize for these chemicals
K_{ow} Low: $<3,000$ Medium: 3,000–20,000 High: $>20,000$	Translocation Accumulation Bioaccumulation Adsorption Transformation Biodegradation Respiration	7. A chemical with a low K_{ow} indicates that it has high water solubility, aqueous mobility and is not susceptible to bioaccumulation, accumulation, or sorption to solids but is susceptible to biodegradation and metabolism by plants and animals
K_{oc} Low: <30 Medium: 30–32,000 High: $>32,000$	Translocation Accumulation Bioaccumulation Adsorption Transformation Biodegradation Respiration	8. Chemicals with high K_{oc} will adsorb to organic carbon and is likely to bioaccumulate and accumulate and less likely to biodegrade or to be metabolized by plants and animals
BCF Low: <250 Medium: 250–1,000	Translocation Accumulation Bioaccumulation Adsorption	9. Chemicals with high BCF values will bioaccumulate and yield food-chain contamination
High: $>1,000$	Transformation Biodegradation Respiration	10. A high BCF value implies a high K_{ow} value, and thus the same heuristics applying for K_{ow} apply for BCF
Half-lives of hydrolysis (days) Low: <30 Medium: 30–90 High: >90	Translocation Accumulation Bioaccumulation Adsorption	11. Compounds considered to be hydrolyzed rapidly are not prone to bioaccumulation, accumulation, food-chain contamination, or adsorption, and are not generally considered persistent, whereas the opposite is true for compounds not hydrolyzed rapidly
	Transformation Hydrolysis	12. In general, for all transformation reactions, if half-lives are high, the chemical is not considered to be susceptible to bioaccumulation, accumulation, food-chain contamination, or adsorption, and are not generally considered persistent

efficient methodology is needed. Design choices in early design might include the selection of raw materials, solvents, reaction pathways, and heat or mass separating agents, and the number of choices may be very large. Later in design, we consider the environmental performance of a small number of process flowsheets. The purpose of this section is to present applications of this tiered approach in the comparison of alternative design choices, emphasizing the methodologies and computer-aided tools.

Early Process Design Evaluations: "Tier I" Assessment

Early design is arguably the most important stage in the hierarchy for conducting environmental assessment. Environmental burdens of chemical processes are largely "locked in" by choices made during early design and attempts to improve environmental performance after a chemical process is operational are expensive and disruptive. Given the importance of early design assessment and the need to

Table 5.8 Description of environmental factors and prediction/analysis methods

Environmental factors	Description
Environmental properties of chemicals	Equilibrium distribution of chemicals among air, water, solid phases in the environment Degradation rates in air, water, and soil/sediment Toxicological properties Structure–activity relationships based on chemical structure Online databases
Emission estimation from process units	Emission factors for major process units Emission correlations for fugitive sources, storage tanks, material transfer and handling Emissions for process heating and utilities
Environmental fate	Fate models for wastewater treatment Fate models in a multimedia environment
Environmental performance metrics	Models to characterize environmental impacts
Process intensification and integration	Integration of heat and power Mass integration to prevent waste Pinch analysis, source sink diagrams
Process optimization	Mixed integer nonlinear programming Multi-criteria optimization

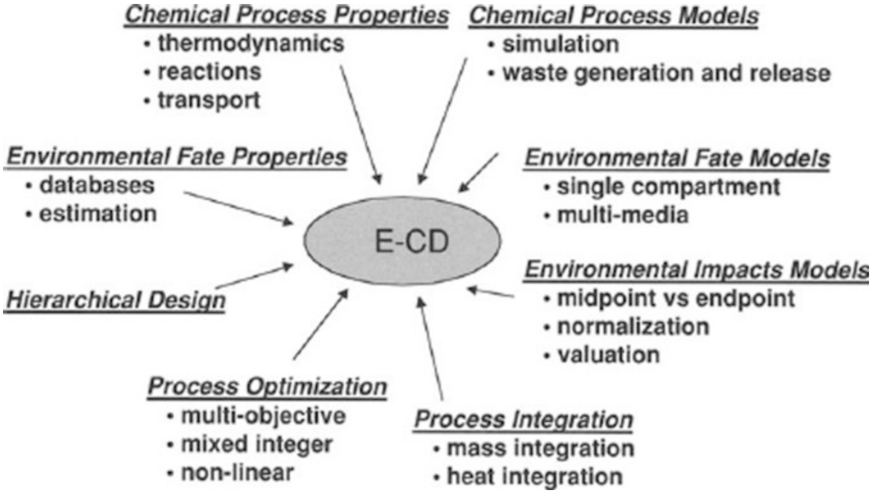


Fig. 5.11 The linking of environmental assessment methods and tools to the design of environmentally conscious chemical processes. *E-CD* environmentally conscious design

provide information on a large number of choices, streamlined yet accurate assessment methods are a high priority.

Solvent selection. Solvent selection is often conducted in early design of chemical processes. A method to match desirable solvent properties (solubility parameters, for example) while simultaneously avoiding undesirable environmental impacts (persistence, toxicity, volatility, etc.) would improve design performance. PARIS II is a program combining such solvent design characteristics. Solvent composition is manipulated by a search algorithm aided by a library of routines with the latest fluid property prediction techniques, and by another library of routines for calculating solvent environmental performance requirements [71].

Reaction pathway selection. Reaction pathway selection is another very important early design activity. High conversion of reactants, selectivity to desired products, and avoiding by-product reactions are among the 12 Green Chemistry principles from Anastas and Warner [3] described previously. Atom and mass efficiency are Green Chemistry performance measures that aid in early design assessment. Atom efficiency is the fraction of any element in the starting material that is incorporated in the product. The rationale is that reactions with high atom efficiencies will be inherently less wasteful. For example, in the reaction of phenol with ammonia to produce analine ($C_6H_5 - OH + NH_3 \rightarrow C_6H_5 - NH_2 + H_2O$), atom efficiencies of C, H, O, and N are 100%, 77.8%, 0%, and 100%, respectively. Mass efficiency for this reaction is defined as the ratio of mass in the product with mass in the reactants. For this example mass efficiency is

$$\begin{aligned}\text{Mass in Product} &= (6\text{C})(12) + (7\text{H}) \times (0\text{O}) \times 16) + (1\text{N}) \\ &\quad \times (14) \\ &= 93 \text{ g}\end{aligned}$$

Table 5.9 Environmental causes and midpoint/endpoint effects

Impact category	Initiating event	Environmental processes	Midpoint effects	Endpoint effects	Impact indicator(s)
Global climate change	Emission of greenhouse gases (CO_2, CH_4, N_2O, CFCs)	Absorption of infrared radiation	Increase in atmospheric temperature, sea level increase	Increases in human diseases and mortality Climate warming Ecological damage	Global warming potential (GWP)
Stratospheric ozone	Emission of ozone depleting compounds (CFCs, halons)	Chemical reaction release of C1 and cBr in stratosphere	Catalytic destruction of ozone in stratosphere	Skin and crop damage, damage to materials	Ozone depletion potential (ODP)
Acidification	Release of oxides of sulfur and nitrogen (SO_x, NO_x)	Chemical oxidation and deposition of acid compounds	Acidic precipitation and lowering of pH in lakes and soils	Ecosystem changes and plant/animal death	Acidification potential (AP)
Smog formation	Emission of smog precursors	Photochemical conversion to ozone and other species	Elevated ozone and aerosols in urban atmospheres	Reduced plant growth, human morbidity and mortality	Photochemical ozone creation potential (POCP)
Human cancer and noncancer effects	Release of toxic and carcinogenic compounds	Environmental fate and transport, exposure pathways	A dose to humans in excess of acceptable levels	Increases in human mortality and morbidity	Human toxicity potential (HTP)
Ecological toxicity	Release of toxic and carcinogenic compounds	Environmental fate and transport, exposure pathways	A dose to animals in excess of acceptable levels	Increases in animal mortality and morbidity	Ecological toxicity potential (ETP)
Resource consumption	Utilization of natural resources from the Earth	Extraction processes	Disruption of natural environments	Unsustainable depletion of resources	Fossil energy, land use area, mineral resource use

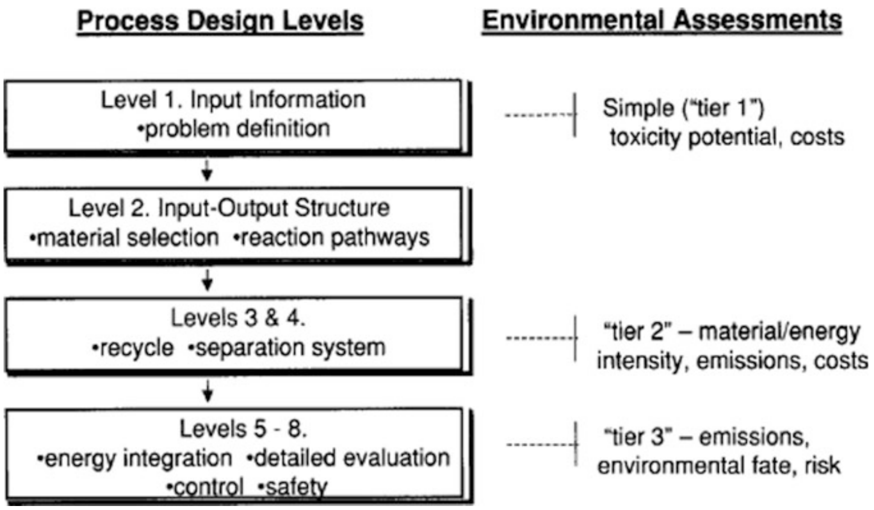


Fig. 5.12 A hierarchical structure to environmental evaluation of chemical process design

$$\begin{aligned}\text{Mass in Reactants} &= (6\text{C})(12) + (9\text{H}) \times (1) + (1\text{O}) \times 16) \\ &\quad + (1\text{N}) \times (14) \\ &= 111\ \text{g}\end{aligned}$$

$$\text{Mass Efficiency} = 93/111 \times 100 = 83.8\%$$

A higher atom and mass efficiency is desirable when comparing alternative reaction choices. A clear example of this is the production of maleic anhydride (MA) starting from either benzene or *n*-butane [72]. Benzene or *n*-butane is partially oxidized in the vapor phase in the presence of air and a solid catalyst at high temperature and pressure.

$$\text{Benzene route: } 2C_6H_6 + 9O_2 \rightarrow 2C_4H_2O_3 + H_2O + 4CO_2$$

$$n\text{-Benzene route: } C_4H_{10} + 7/2O_2 \rightarrow C_4H_2O_3 + 4H_2O$$

Mass efficiencies for these routes are 44.4% and 57.6% for the benzene and *n*-butane routes, respectively.

This simplistic analysis indicates that *n*-butane is the "greener" reaction route. But further analysis at this early design stage can shed more light on the differences in environmental performance. We will focus on raw material cost and CO_2 generation in this screening comparison because of concerns of economics and global climate change. It is necessary to bring in differences in conversion and selectivity at this stage. The benzene route has typical conversions of 95% and selectivity to MA of 70%, with approximately equimolar amounts of CO and CO_2 generated as by-products. For the *n*-butane route these values are 85% and 60%, respectively. Assuming 1 mole of MA produced, 1/0.70 mole of benzene or 1/0.60 mole of *n*-butane is needed. The raw material costs are as follows:

$$\begin{aligned}&\text{Benzene: } (1\ \text{mole}/0.70\ \text{mole}) \times (78\text{g/mole})\\&\quad\times (0.000280/\text{g}) = 0.0312\ \$/\text{mole of MA}\end{aligned}$$

$$\begin{aligned}&n\text{-Butane: } (1\ \text{mole}/0.60\ \text{mole}) \times (58\ \text{g/mole})\\&\quad\times (0.000214/\text{g}) = 0.0207\$/\text{mole of MA}\end{aligned}$$

This result shows that the *n*-butane route costs less for the raw material than the benzene route due to the lower price of *n*-butane, even though the molar yield of MA from *n*-butane is less. This simplistic economic analysis is of relevance because raw material costs often dominate. The generation of CO_2 in the reactor is estimated as follows:

$$\begin{aligned}&\text{Benzene: } [1\ \text{mole MA} \times (2\ \text{mole CO}_2/\text{mole MA})]\\&\quad + [(1\ \text{mole Benzene}/0.70\ \text{mole MA}) \times (0.965 - 0.7)\\&\quad \times 6/2] = 3.071\ \text{mole CO}_2/\text{mole of MA}\end{aligned}$$

$$\begin{aligned}&n\text{-Butane: } (1\ \text{mole } n\text{ - Benzene}/0.60\ \text{mole MA})\\&\quad \times (0.85 - 0.6) \times 4/2 = 0.833\ \text{mole CO}_2/\text{mole of MA}\end{aligned}$$

The first [] term in the Benzene calculation accounts for the two carbons that are liberated when MA is formed, and the second [] term is for conversion to by-products, with 1/2 going to CO_2 and the other to CO. MA from benzene generates almost four times as much CO_2 in the reactor than the *n*-butane route. Additional CO_2 is generated when unreacted benzene or *n*-butane and by-product CO is incinerated in the pollution control equipment from the process. The total CO_2 emission including pollution control is calculated as shown next.

$$\begin{aligned}&\text{Benzene: } 3.071 + (1.071)(0.99) + (0.0714)(0.99)(6)\\&\quad = 4.595\ \text{moles CO}_2/\text{mole of MA}\end{aligned}$$

$$\begin{aligned}&n\text{-Butane} : 0.833 + (0.833)(0.99) + (0.25)(0.99)(4)\\&\quad = 2.688\ \text{mole CO}_2/\text{mole of MA}\end{aligned}$$

The second term in each summation is CO converted to CO_2 in the pollution control device; the third is due to unconverted feedstock. These screening calculations verify that the *n*-butane route emits approximately 1/2 the CO_2 compared to the benzene pathway. Based on the economic and environmental screening, the benzene route would be excluded from further consideration. More detailed calculations based on optimized flowsheets confirm these screening calculations are accurate [72].

Table 5.10 provides another example of the type of assessment that chemical engineers will need to perform with limited information [73]. A process engineer evaluating two alternative synthesis routes for the production of methyl methacrylate can use data on persistence, bioaccumulation, toxicity, and stoichiometry to quickly evaluate potential environmental concerns. These data can be estimated using group contribution methods when measured values are not available. The estimates of persistence, bioaccumulation, toxicity, and stoichiometry can then be combined to provide preliminary guidance. In this case concerns about the health and safety issues associated with sulfuric acid dominate, and the isobutylene route appears preferable because it requires less acid. Although more detailed data are available for these two processes, this level of data is typical of what might be available for new process chemistries.

Evaluations During Process Synthesis: "Tier II" Assessment

"Tier II" environmental assessment is employed for flowsheet synthesis on a smaller number of design alternatives. This provides an opportunity to evaluate the impacts of separation and other units in the process in addition to the reactor. Identification of emission sources and estimation of release rates are also part of this assessment. Inclusion of additional environmental and sustainability metrics (energy intensity [energy consumption/unit of product], water intensity, toxic release intensity, etc.) are hallmarks of "Tier II" assessment. More information on "Tier II" assessment is provided in the text by Allen and Shonnard [53].

Detailed Evaluation of Process Flowsheets: "Tier III" Assessment

After a process flowsheet has been established, it is appropriate for a detailed environmental impact evaluation to be performed. The end result of the impact evaluation will be a set of environmental metrics (indexes) representing the major environmental impacts or risks of the entire process.

Table 5.10 Stoichiometric, persistence, toxicity and bioaccumulation data for two synthesis routes for methyl methacrylate [53, 73]

Compound	Pounds produced or pounds required per pound of methyl methacrylate[a]	Atmospheric half-life or aquatic half-life[b]	1/TLV[c] (ppm)[c]	Bioconcentration factor[d] (conc. in lipids/conc. in water)
Acetone-cyanohydrin route				
Acetone	−0.68	52 days/weeks	1/750	3.2
Hydrogen cyanide	−0.32	1 year/weeks	1/10	3.2
Methanol	−0.37	17 days/days	1/200	3.2
Sulfuric acid[e]	−1.63		1/2 (est.)	
Methyl methacrylate	1.00	7 h/weeks	1/100	2.3
Isobutylene route				
Isobutylene	−1.12	2.5 h/weeks	1/200 (est.)	12.6
Methanol	−0.38	17 days/days	1/200	3.2
Pentane	−0.03	2.6 days/days	1/600	81
Sulfuric acid[e]	−0.01		1/2 (est.)	
Methyl methacrylate	1.00	7 h/weeks	1/100	2.3

[a] A negative stoichiometric index indicates that a material is consumed; a positive index indicates that it is produced in the reaction. A screening environmental index is used for comparison in "Tier I" assessment; Environmental index $= \sum |\nu_i| * (\mathrm{TLV}_i)^{-1}$

[b] The atmospheric half-life is based on the reaction with the hydroxyl radical; aquatic half-life via biodegradation is based on expert estimates. From: EPISuite software (http://www.epa.gov/oppt/greenengineering/software.html) or ChemFate Database (http://www.syrres.com/eswc/chemfate.htm)

[c] TLV is the threshold limit value, and the inverse is a measure of inhalation toxicity potential for a chemical. Values taken from NIOSH Pocket Guide to Chemical Hazards (http://www.cdc.gov/niosh/npg/npg.html), and the Specialized Information Service of the National Library of Medicine (NLM) (http://sis.nlm.nih.gov/)

[d] Bioconcentration factor is an indicator of a chemical's potential to accumulate through the food chain. From the EPISuite software (http://www.epa.gov/oppt/greenengineering/software.html) or ChemFate Database (http://www.syrres.com/eswc/chemfate.htm)

[e] The lifetime of sulfuric acid in the atmosphere is short due to reactions with ammonia

A number of indexes are needed to account for potential damage to human health and to several important environmental compartments.

In quantitative risk assessment [74], it is shown that impacts are a function of dose, dose is a function of concentration, and concentration is a function of emission rate. Therefore, emissions from a process flowsheet are a key piece of information required for impact assessment during process design. A number of computer-aided tools are available to generate pollutant emissions to air using process flowsheet information. Emission factors are used to calculate emission rates from major units (distillation columns, reactors, other columns, furnaces, boilers, etc.) based on process flows and utilities [75]. Fugitive sources (valves, pumps, flanges, fittings, sampling valves, etc.) number in the thousands for a typical chemical process and together contribute to facility air emissions. A compilation of emission estimation methods for chemical processes is found in Chap. 8 of Allen and Shonnard [53].

As previously discussed, the concentrations in the relevant compartments of the environment (air, water, soil, biota) are dependent upon the emission rates and the chemical/physical properties of the pollutants. A fate and transport model transforms emissions into environmental concentrations. Although single compartment fate models are often used to predict concentrations downwind or downgradient from emissions sources—for example, an atmospheric dispersion model, a groundwater fate model, or river model—most applications of risk assessment to the design of chemical processes employ a multimedia compartment model (MCM). MCMs predict regional pollutant transport and fate, with a typical scale of 100 × 100 km of the earth's surface. Mechanisms of diffusive and convective transport as well as degradation reactions are applied to separate well-mixed environmental compartments (air, water, soil, sediment). Steady-state concentrations in each compartment are expected to match actual environmental concentrations within an order of magnitude. An illustrative example calculation using the "Level III" model of Mackay [76] is presented in Chap. 11 of Allen and Shonnard [53].

Finally, information regarding toxicity or inherent impact is required to convert the concentration-dependent doses into probabilities of harm (risk). Based on this understanding of risk assessment, the steps for environmental impact assessment are grouped into three categories, (1) estimates of the rates of release for all chemicals in the process, (2) calculation of environmental fate and transport and environmental concentrations, and (3) the accounting for multiple measures of risk using toxicology and inherent environmental impact information. Computer-aided software packages integrating these three calculation steps are available and have been linked to commercial process simulation packages [72, 77, 78]. Figure 5.13 shows the information flows occurring in the software tool EFRAT (Environmental Fate and Risk Assessment Tool), which links with the simulation package HYSYS. Flowsheet stream and utility information from

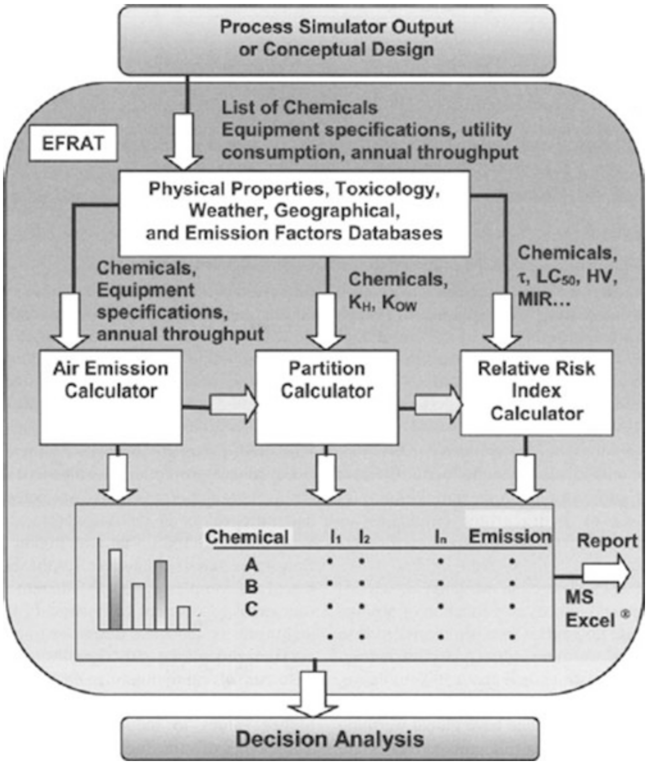


Fig. 5.13 Information flow diagram for EFRAT and relation to HYSYS™ (Shonnard and Hiew [77], Copyright American Chemical Society and reproduced by permission)

HYSYS is automatically transferred to EFRAT once the flowsheet is “synchronized” with EFRAT. Emission factors and correlations within EFRAT estimate release rates to the air, an MCM predicts environmental partitioning, and finally a relative risk assessment module in EFRAT generates nine environmental risk indicators. A version of EFRAT is available free of charge for education purposes [79].

Figure 5.14 provides an example of the type of decision that a process designer would face in flowsheet evaluation. In this simple example, absorption with a regenerable solvent is used to capture (and recycle or sell) toluene and ethyl acetate, which might otherwise be emitted into the atmosphere. To increase the fraction of the hydrocarbons that are absorbed, the circulation rate of the solvent can be increased, but this will increase the duties of the reboiler, condenser, and pumps in the system, increasing energy use and increasing atmospheric emissions (primarily of criteria pollutants, sulfur and nitrogen oxides, particulates, carbon monoxide, and of the greenhouse gas CO_2). The process engineer will need emission estimation tools to evaluate such trade-offs, and will need to evaluate the potential environmental and economic costs associated with different types of emissions, in this example the relative costs of hydrocarbon emissions as opposed to emissions of criteria pollutants and CO_2.

Figure 5.15 shows the variation of several environmental indices as a function of absorption solvent flow rate. There is a sharp decrease in the global warming index (I_{GW}) with increasing solvent flow rate until about 50 kgmoles/h, due mostly to toluene recovery. (Note that toluene and ethyl acetate have global warming impacts assuming that all emitted VOCs are oxidized to CO_2 in the environment and that the VOCs are of fossil origin.) Thereafter, increasing utility-related emissions of greenhouse gases (primarily CO_2) drive this index up faster than its rate of decrease by further recovery of ethyl acetate. The smog formation index (I_{SF}) decreases sharply with absorber solvent flow rate in the range of 0–50 kgmoles/h, again due to recovery of toluene. Afterwards, there is a slow decline in I_{SF} with increasing solvent flow rate above 50 kgmoles/h as ethyl acetate is recovered. The acid rain index I_{AR} increases in nearly direct proportion to solvent flow rate. Utility consumption and its associated sulfur and nitrogen oxide emissions (precursors to acid rain) drive the I_{AR} up with higher solvent flow rates. These flowsheet results begin to reveal the complex trade-offs inherent in environmental assessment of chemical processes and products. Does the design engineer operate the process at 50 kgmoles/h for the absorber solvent to minimize emission of CO_2 and global warming impacts or operate at higher values to reduce smog formation? These types of value judgments are commonly encountered in the decision-making process. Notwithstanding the complexity, environmental metrics inform the design activity on potential environmental impacts of the process.

Hybrid Screening Evaluations: Combining “Tier I”–“Tier III” Life-Cycle Assessment

Early design assessments similar to those presented in previous sections of this chapter have the following limitations: (1) they tend to focus on the reaction step and neglect the impacts of downstream units, (2) the assessment includes one or a small number of environmental indicators, and (3) the early assessment typically does not consider impacts beyond the process boundary, for example, the environmental burdens associated with the life cycle of materials used in the process. In this section we will explore some approaches to address these limitations.

Combining “Tier I” with Tier “III.” In the next example, the maleic anhydride process from an earlier section is reexamined by including the effects of units downstream from the reactor. Figure 5.16 shows the major emission sources from the process: the reactor, absorption unit, and distillation. Emission factors along with stream flow rates in the reactor are used to estimate releases of benzene, *n*-butane, CO, and MA. The emission from the reactor is estimated

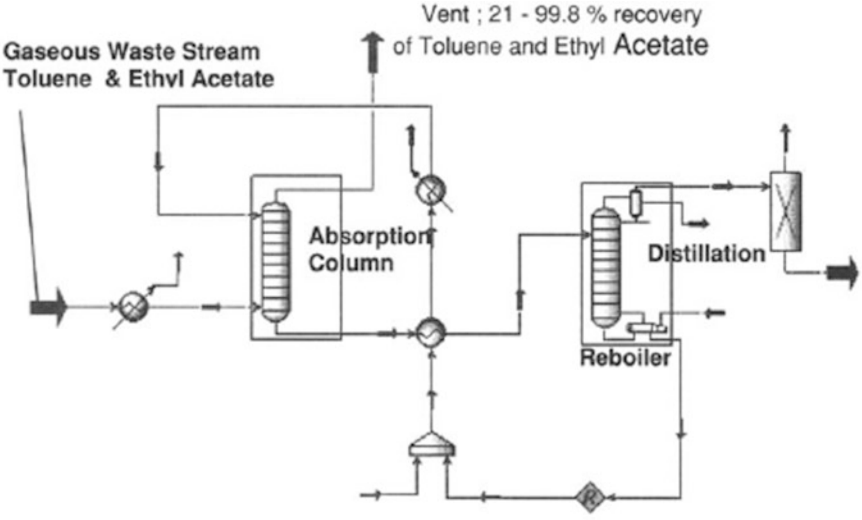


Fig. 5.14 Absorption with a regenerable solvent is used to capture toluene and ethyl acetate (Shonnard and Hiew [77], Copyright American Chemical Society and reproduced by permission)

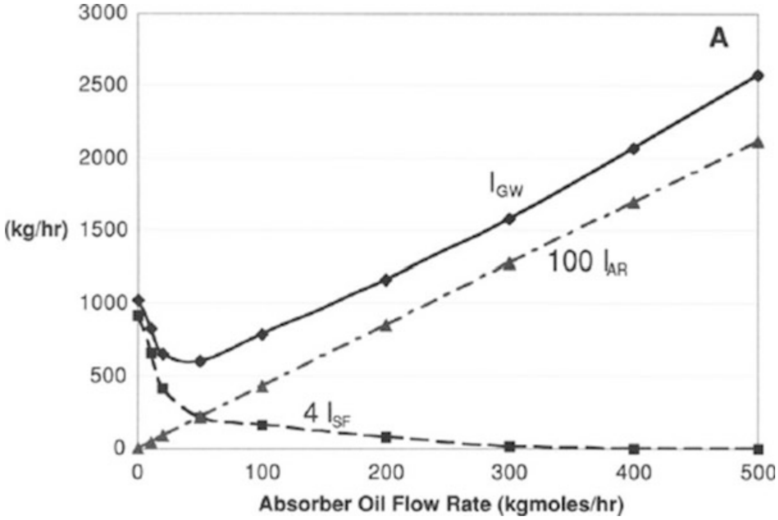


Fig. 5.15 Variation of environmental indexes with absorber oil flow rate for global warming (L_{GW}), smog formation (I_{SF}), and acid rain (I_{AR}) (Shonnard and Hiew [77], Copyright American Chemical Society and reproduced by permission)

using an average emission factor (EF_{av}, 1.50 kg emitted/10^3 kg throughput [53], chapter 8, eqn. 8-4) using the equation

$$E = M_{voc} EF_{av}$$

In this equation, M_{voc} is the mass flow rate (kilograms per unit time) of the volatile organic compound in the reactor, and it is taken as the average mass flow rate through the reactor. The emissions from the absorber column originate from the offgas vent and this stream contains unreacted raw material, by-products, and product. Raw material, especially benzene, and one of the by-products, CO, are toxic. To include the effects of pollution control at this early design stage, it is assumed that unreacted raw material, by-products, and unrecovered product exiting the reactor are incinerated to CO_2 with 99% destruction efficiency with the remaining 1% released to the environment. The recovery of MA in the separation system is assumed to be 99%, with the remaining 1% going to pollution control.

Based on the process description above, Table 5.11 can be generated which contains the total emission of each pollutant. Emissions are converted to impact indicators using the environmental fate and impact assessment tool EFRAT [79]. A comparison of environmental indicators for MA production from either benzene or *n*-butane is shown in Table 5.12. All of the environmental indices in the *n*-butane process are less than or equal to those in the benzene process. This table reveals that the benzene route is estimated to have greater environmental impacts than the *n*-butane route. This early assessment is more rigorous than the approach in the previous section titled “Early Process Design Evaluations” because more indicators were used and the effects of units downstream of the reactor were included.

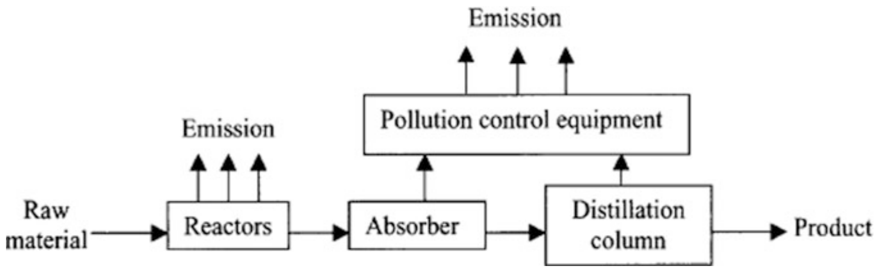


Fig. 5.16 Major emission sources from the maleic anhydride process

Table 5.11 Stream flow rates and emission rates for screening evaluation of environmental impacts from two alternative reaction routes for maleic anhydride production

	Benzene process (kg/mole of MA)				
	Benzene	CO	CO_2	Maleic anhydride	Total
Total process emission rate	1.4×10^{-4}	3.23×10^{-4}	2.03×10^{-1}	8.33×10^{-5}	2.04×10^{-1}
	n-Butane process (kg/mole of MA)				
	n-Butane	CO	CO_2	Maleic anhydride	Total
Separation unit (w/pollution control)	2.3×10^{-4}	2.51×10^{-4}	1.18×10^{-1}	8.33×10^{-5}	1.18×10^{-1}

Table 5.12 Environmental indices of both routes (all the values are in units of kg/mole of MA)

Chemical	I_{FT}	I_{ING}	I_{INH}	I_{CING}	I_{CINH}	I_{OD}	I_{GW}	I_{SF}	I_{AR}
Benzene	6.8×10^{-6}	3.3×10^{-3}	4.6×10^{-2}	1.4×10^{-4}	1.4×10^{-4}	0.0	2.0×10^{-1}	2.5×10^{-5}	0.0
n-Butane	3.0×10^{-6}	3.1×10^{-3}	3.8×10^{-2}	0.0	0.0	0.0	1.2×10^{-1}	4.4×10^{-6}	0.0

Including life-cycle assessment at the early design stage. In this example, a comparison is made between a new wood pulp bleaching process and a conventional process in order to uncover advantages and limitations of the "Green Chemistry" innovation [80]. In the production of most paper, wood is chemically treated with NaOH and Na_2S at high temperature (170°C [81]) to remove lignin and some hemicellulose, yielding pulp with a small residual of lignin that must be removed by bleaching to achieve a high brightness paper. Conventional bleaching utilizes ClO_2 and NaOH in sequential processes (Fig. 5.17). This process generates approximately 0.5 kg of chlorinated organics per ton of bleached pulp, and even though these water pollutants are less persistent, bioaccumulative, and toxic than those generated using Cl_2 (earlier bleaching agent), they are still of concern. Furthermore, the need to raise and lower the stream temperature consumes a large amount of energy and cooling water.

The 1999 Green Chemistry Challenge Award was given to Professor Terrence J. Collins of Carnegie Mellon University for the development of "Tetraamido-macrocyclic ligand" TAML™ activators [82]. These organic amine compounds contain a catalytic iron center that activates hydrogen peroxide at minute concentrations in the bleaching process (50 g/ton pulp). Although information is lacking about the toxicity of the TAML™ Oxidant Activators, it is known that the temperature of the TAML™ peroxide process is 50°C compared to 70°C for chlorine dioxide bleaching (Fig. 5.18). We will use these temperature differences as the basis of the green engineering analysis in comparing conventional ClO_2 bleaching with TAML™ process, with energy over the life cycle being the sustainability indicator.

Let us assume that the chlorine dioxide bleaching process has the sequence *DEDED* ($D =$ chlorination with ClO_2, $E =$ extraction with NaOH) in which the pulp is cold-water drum washed between *DE* stages giving a drum exit temperature of 50°C. TAML™ peroxide bleaching process is assumed to follow a sequence *PPPD* ($P =$ peroxide). Thus, two of the *DE* stages are replaced with two *PP* stages in which there exists a 20°C difference between the stream temperatures. Furthermore, assuming a 10% consistency for the pulp (9 t water/ton pulp), we can estimate an energy savings of Energy Saved/ton pulp $= 2$ stages $\times$ $(20 \times 9/5)°$ F/stage $\times$ (2,200 Btu/ton water $\times$ 9 t water/ton pulp) $= 1.426 \times 10^6$ Btu/ton pulp. (Note, the contribution of the pulp to the energy analysis is neglected, being a minor component).

According to Department of Energy statistics the average energy consumed per ton of pulp is 39.4×10^6 Btu/ton pulp

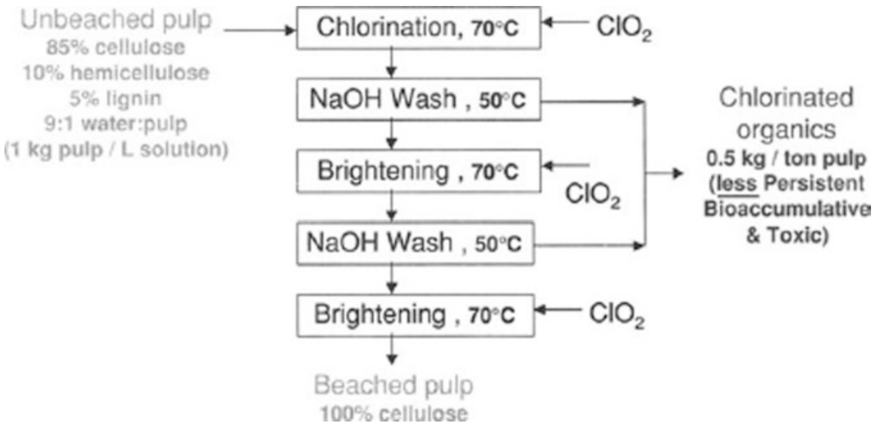


Fig. 5.17 Chlorine dioxide process for bleaching of unbleached wood pulp

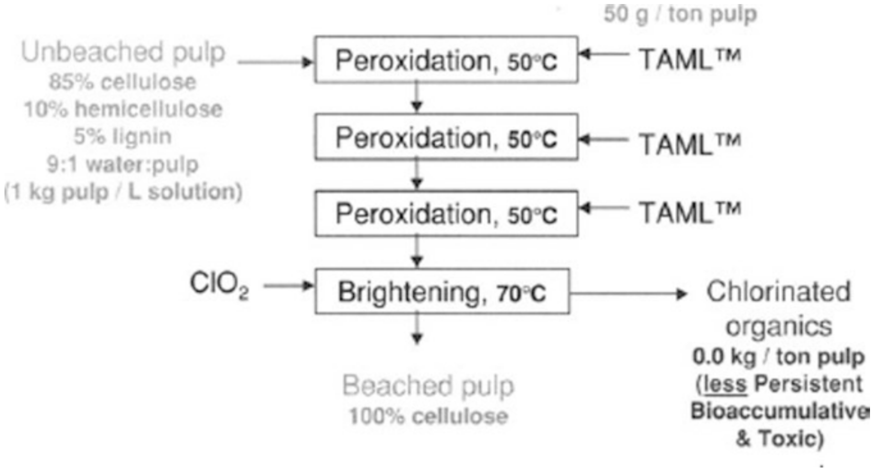


Fig. 5.18 Wood pulp bleaching process using TAML™ activators and hydrogen peroxide

[81]. Given that roughly 1/2 of the pulp in the United States is bleached, this energy savings is about 2% of the annual energy consumed for production of pulp. This significant energy savings was cited in the application for the 1999 Green Chemistry Award [82]. This analysis included the energy differences between these bleaching processes in terms of the stream heating requirements. The next section will present a more complete energy efficiency analysis of the TAML™ peroxide bleaching processes by including the effects of producing and delivering the bleaching agents to the process. Figure 5.19 shows the chain of material flows in the production of ClO_2, H_2O_2, NaOH, and TAML activator. Table 5.13 shows the primary energy intensity for each of the materials in the chain of materials for ClO_2 and H_2O_2 bleaching of pulp.

The overall energy change for substituting H_2O_2 bleaching for ClO_2 bleaching is obtained by combining the net energy change from producing the chemicals and then subtracting the in-process energy savings. This yields a total decrease in energy consumption for the TAML™ H_2O_2 bleaching process of $(1.15 \times 10^7 - 1.87 \times 10^6$ Btu/ton pulp$) - 1.426 \times 10^6 = -8.63 \times 10^6$ Btu/ton pulp.

The energy decrease as a percentage of the energy consumption rate of the pulp and paper industry is $0.5 \times 8.63/39.394 \times 100 = 11.0\%$.

This Green Engineering screening level energy analysis of the Green Chemistry award for TCF (total chlorine free) bleaching of pulp points out a potential energy benefit by implementing the TAML™ process in the US pulp and paper industry of approximately 11.0% over the "cradle-to-gate" life cycle for pulp. This conclusion is based upon a screening-level analysis that is uncertain, and more detailed analysis of the life-cycle energy burdens must be completed before the energy implications are fully understood. One step in the life cycle for ClO_2 production was not included for lack of data, the process energy for converting $NaCl0_3$ to ClO_2. This screening assessment is expected to underestimate the energy savings for implementing this new technology. However, the analysis points to the need for including all relevant life-cycle stages in the evaluation, not only the in-process changes, as was done previously. The energy burdens of the chemicals in the supply chain for each bleaching alternative are more significant than the in-process energy improvements.

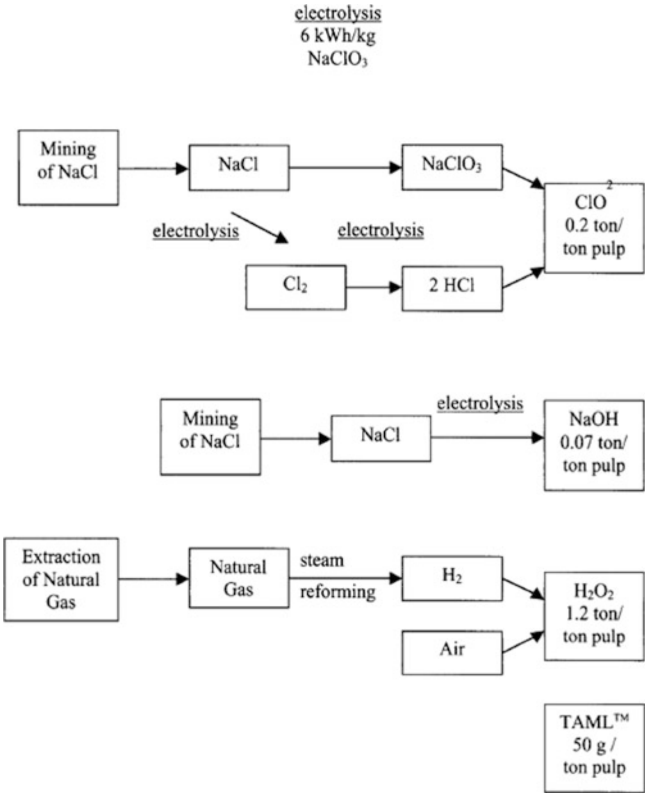


Fig. 5.19 Material flow chain for pulp bleaching chemicals

Conclusions

Environmental performance assessment is a powerful tool for moving the chemical industry toward sustainability. It provides important insights to the process designer on potential environmental impacts of alternative choices at several stages in the design process. Computer-aided tools further enhance the environmentally conscious engineering design activity. These tools for solvent selection, reaction pathway analysis, emission estimation, environmental property estimation, fate modeling, and life-cycle assessment will find increased use in the chemical industry. They will be particularly effective in decision-making when coupled with economic and societal criteria. Application of these tools for comparison of traditional processes and products to innovative alternatives is already taking place in fields such as bio-based products, recycle of wastes, elimination of toxic compounds, green chemistry, benign solvents, and material flow analysis.

Life-Cycle Assessment

Introduction

Products and processes all have a natural life cycle. For example, the life cycle of a product starts from the extraction of raw materials for its production and ends when the product is finally disposed. In the production, use, and disposal of this product, energy is consumed and wastes and emissions are generated. A life-cycle assessment is an analysis in which the use of energy and materials are quantified and the potential

Table 5.13 The subitems in calculating the energy requirement in ClO_2 and H_2O_2 bleaching

	Energy in original units	Energy (Btu/ton ClO_2 or ton NaOH)
ClO_2 bleaching		
Mining NaCl	1.3 MJ/kg NaCl	9.69×10^5 Btu/ton ClO_2
Electrolysis of NaCl to $NaClO_3$	17.1 kWh/kg $NaClO_3$	8.35×10^7 Btu/ton ClO_2
HCl production	2.5 MJ/kg HCl	2.32×10^6 Btu/ton ClO_2
HCl production electricity	0.043 kWh/kg HCl	1.44×10^5 Btu/ton ClO_2
Total from ClO_2		8.69×10^7 Btu/ton ClO_2
NaOH production	21.3 MJ/kg NaOH	1.83×10^7 Btu/ton NaOH
Total ClO_2 bleaching		(1.87×10^7 Btu/ton pulp)
H_2O_2/TAML bleaching		
H_2O_2 production		9.60×10^6 Btu/ton H_2O_2
TAML production		1.0×10^8 Btu/ton TAML™
Total H_2O_2/TAML bleaching		1.15×10^7 Btu/ton pulp

Energy intensities are taken from the SimaPro 6.0 software and other sources [80]

1 Btu = 0.0010551 MJ

1 Btu = 0.0002931 kWh

1 kWh = 3.6 MJ

Example calculation: The energy required to produce 1 t of ClO_2 is therefore (0.867 kg NaCl/kg ClO_2)×(1.3 MJ/kg NaCl) × (2,000/2.205 kg/ton)×(947.8 Btu/MJ + (24,478 kWh/ton of ClO_2) ×(3,411.8 Btu/kWh for electrolysis of NaCl solution (1.08 t HCl/ton ClO_2) × (2.5 MJ/kg HCl) × (2,000/2.205 kg/ton) × (947.8 Btu MJ) + (1.08 t HCl/ton ClO_2) × (0.043 kWh/kg HCl) × (2,000/2.205 kg/ton) × (3,411.8 BtukWh) = $9.69 \times 10^5 + 8.35 \times 10^7 + 2.32 \times 10^6 + 1.44 \times 10^5 = 8.69 \times 10^7$ Btu/ton ClO_2.

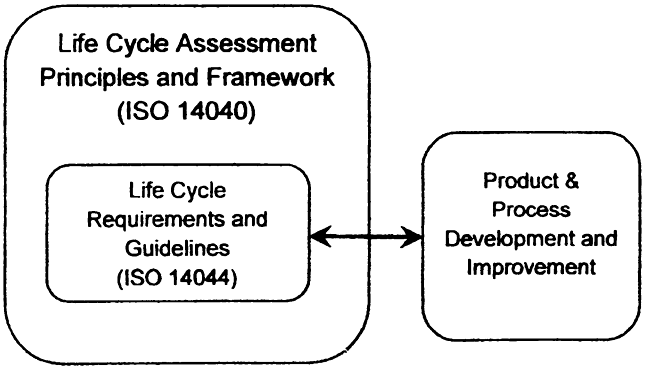


Fig. 5.20 ISO standards for life-cycle assessment [90–95]

environmental and societal impacts are predicted. Life-cycle thinking is progressively being adopted by industry as an essential tool for analyzing processes and products. For example, 3M instituted a life-cycle management policy in 2001, requiring reviews of all new products and prioritized reviews of existing products [83]. BASF is using eco-efficiency analysis of a life-cycle assessment to compare products they produce and examine their economic and ecological characteristics [84, 85]. This section presents a review of methods used for life-cycle assessment and then gives examples of assessments conducted on chemical products and processes.

Life-cycle assessment methodology has been used since the 1960s with early studies that focused solely on energy usage and solid waste issues. This focus continued in life-cycle assessments performed during the oil crisis in the 1970s [86, 87]. The unique aspect of all of these initial studies was the early development and use of life-cycle data inventories with less emphasis on environmental risk impacts of the associated processes studied. A method published by the Royal Commission on Environmental Pollution in 1988, employing the “Best Practicable Environmental Option,” attempted to minimize the environmental burdens at the LCA practitioner’s manufacturing plant, but this method did not analyze the entire lifecycle of the product or process in question [88]. However, even today, life-cycle assessments that focus primarily on mass and energy balances (e.g., Beaver [89]) provide useful life-cycle information.

Life-cycle assessments capturing the full life of a process or product came into fruition in the late 1980s but were performed primarily by private companies, thus leaving into question the specific methodology used and the biased nature of the results. In 1993, the Society of Environmental Toxicology and Chemistry (SETAC) addressed the concern of no common framework for LCA methodology by introducing principles on how to conduct, review, use, and present the findings of a life-cycle assessment. The International Organization of Standardization (ISO) introduced Standards 14040 (principles and framework) [90] and 14041 (goal and scope definition and inventory analysis) [91] on life-cycle assessments in 1997 and 1998, respectively, and Standards 14042 (impact assessment) [92] and 14043 (interpretation) [93] in 1999 and 2000 respectively. These were updated in 2006 to life-cycle assessment Standards 14040:2006 (principles and framework) [94], these were updated in 2006 to life-cycle assessment standard 14040: 2006 (in) [94] and Standards 14041 through 14043 were condensed into the new Standard 14044: 2006 (requirements and guidelines) [95]. Figure 5.20 illustrates the Life-cycle Assessment according to ISO Standards 14040 and 14044.

Life-cycle impact assessments typically consist of four steps [53]. The first step includes the definitions of the system boundary, scope, and the functional unit. The second step consists of an inventory of the inputs and outputs of the system. Figure 5.21 depicts the various components of life-cycle inventory (LCI). As the earliest life-cycle assessments were previously described, many life-cycle assessments end at this stage, and conclusions are made on minimizing mass and energy usage. To increase the information gained by using life-cycle assessment, the relative hazards of the chemicals used in and emitted from the process must be included. The linking of the inputs and outputs determined in step 2 to their inherent risk to the environment and public health is the third step of LCA. The fourth and final step of LCA is interpreting the results and recommending process improvements [96]. SETAC has defined a different, but similar, fourth step which is termed “improvement assessment” which includes ways to improve the impact assessment [86]. There are several good reference texts that give detailed procedures on life-cycle assessment [97–99].

Life-cycle assessments are typically conducted on end products, but can also be used proactively for process selection, design, and optimization. Figure 5.22 illustrates the various steps involved in product and process life cycle. Some debate exists on the relative value of life-cycle assessments applied to products vs. processes. Burgess and Brennan [87] state that a life-cycle assessment conducted on a product is also valid for the processing steps involved in the manufacture of the product and that the purpose of conducting an LCA on a process is different than on a product. On the contrary, Chevalier et al. [100] state that a life cycle conducted on a process is more thorough than a life cycle conducted on a product. A life-cycle assessment on a process is usually conducted in the research and development phase to determine if there are other options, such as replacing a solvent with a less environmentally harmful solvent, for comparison to processes used by competitors, to reduce liability, and for marketing and policy-making purposes [87]. Regardless of whether a process or product is evaluated, the basic steps of LCA remain the same. These steps are presented in more detail to follow.

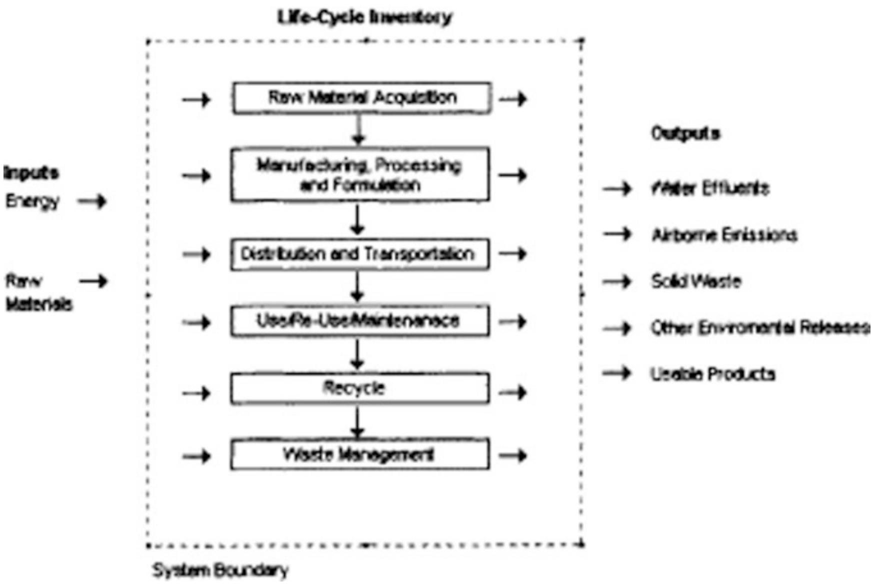


Fig. 5.21 Components of a life-cycle inventory

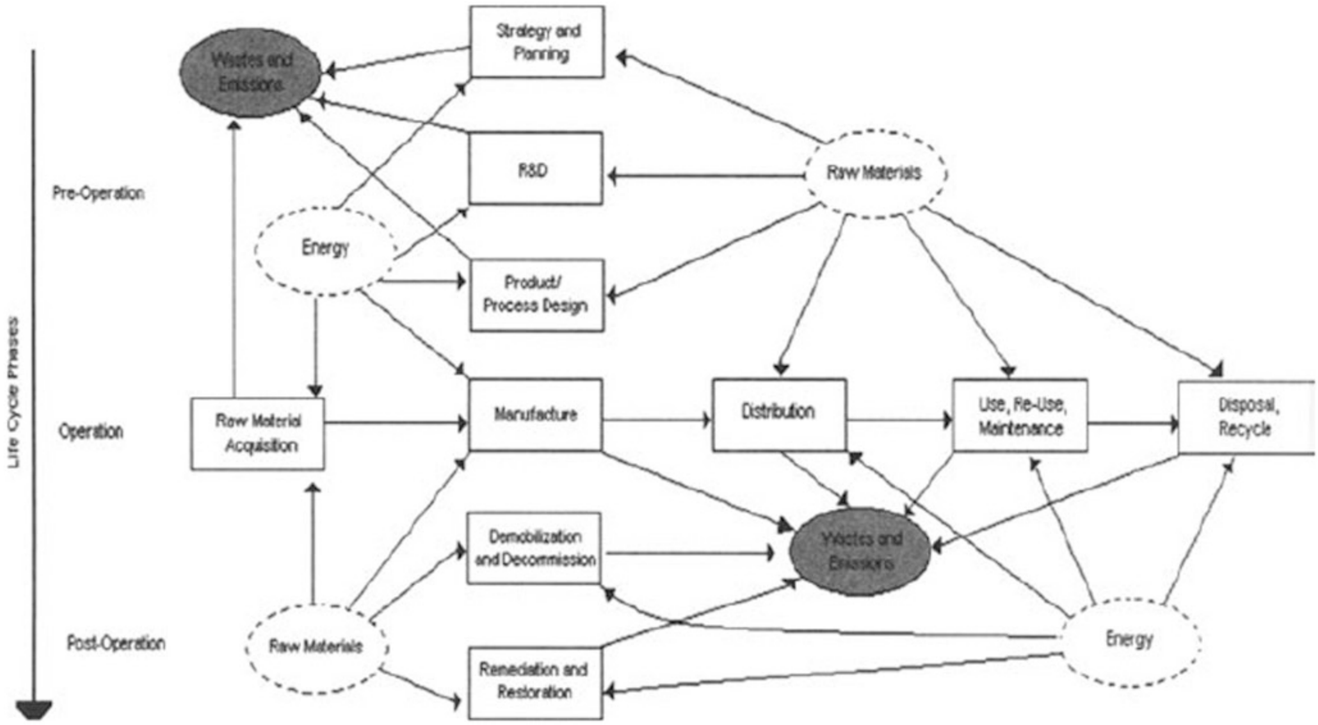


Fig. 5.22 Product and process life cycles [53]

Goal and Scope of LCA

Defining a boundary for a life-cycle assessment varies depending on the methodology used. As previously described, ISO 14040 standardized life-cycle impact assessments in 1997 [101, 102] and established that a life-cycle assessment should be conducted in terms of elementary flows. This approach is typically understood as a cradle-to-grave approach, involving raw material extraction through the disposal of the product, including all of the relevant supply chains [103]. The standard was further refined by ISO 14041 [91], which established a "streamlining" approach to LCA that allowed omission of certain processes, inputs, and outputs if these processes are deemed insignificant. Burgess and Brennan [87] state that assumptions concerning the boundary conditions are necessary to maintain manageability and to ensure that the law of diminishing returns can be observed after three upstream processes. However, according to the ISO standards, these data must be collected before they can be disregarded.

Scoping and goal definition is often thought of as the most important process of an LCA, and care must be taken to establish appropriate boundaries that are consistent with the objectives of a study. Suh et al. [104] warn that including a scientific basis for excluding and including processes in system boundaries is essential and emphasizes the difficulty in doing so, particularly if the practitioner is not familiar with the specific process under study.

The functional unit is defined as the basis upon which all data and impacts are presented and compared and typically relates to the "function" of the product studied. For example, the basis for comparing plastic and paper grocery bags can be the number of bags required to carry a certain amount and type of grocery items. Another example of a functional unit directly reflecting the product function is used when comparing cloth and disposable diapers, accounting for the number of diapers required for each diapering. A common functional unit used in LCAs involving product manufacture is "per mass of product"; however, there are other appropriate functional units that may be more relevant to the specific product effectiveness studied.

Methods of Life-Cycle Inventory

The inventory stage involves the collection of all the data that will be used in the life-cycle analysis. The quality of the data is an important part of the LCI process, and, as with any model, the results of an LCA are only as good as the data inputs. There are two basic sources of data for an LCA, primary and secondary in nature. Primary data are derived directly from the process in question. These are the most accurate data that can be applied to an LCA and, as a result, are the most desirable. However, in many cases, data are proprietary and are not readily available to the public [87], thus necessitating the LCA practitioner to seek secondary sources of data, including databases, peer-reviewed literature, etc., that may not be as accurate and are not often accompanied with error estimates. Because of the frequent use of secondary data, debate within the LCA community has arisen on how to best capture and report error and subsequently interpret results of LCAs.

There are four main methods of conducting an LCI, each differing in its degree of detail. These methods are economy scale, life-cycle scale, equipment scale, and a hybrid LCI [103], listed from the most general to the most detailed in approach. The most general analysis, at the economy scale, contains all the information from cradle-to-grave of a product or process but is not specific for an individual process. On the other hand, the most detailed method, the hybrid LCI approach, is typically used for gate-to-gate (within the grounds of a manufacturing facility) analysis. A brief summary of each LCI method, along with its advantages and disadvantages, is presented below.

The "economy scale" method of LCI typically uses national statistics about resource use and emissions from a specified sector and considers the entire economic system. With this method the boundary is defined from the cradle-to-the-grave of the product. The most well-known method in this category is provided in the EIOLCA software developed at Carnegie Mellon University (2005). There are two disadvantages to this approach. First, an assumption must be made that there is a linear relationship between the dollars spent on the product in question and its environmental impact. In many cases, this is a good assumption, such as in processes involving obtaining a higher purity chemical. This assumption, however, is not valid for other industries with equivalent products having different perceived values, such as automobiles. There is also another drawback to using this type of method because the data typically used are a composite for the entire industry. This assumption may be valid assuming that the company is average in its performance; however, this assumption does not account for industries that are leaders or laggards in areas such as sustainability, operation efficiency, or whether they are using state-of-the-art processes.

The second LCI approach that can be used is at the life-cycle scale, involving focus on a specific sector. Industry uses this approach most frequently in developing LCIs. Although the data using the life-cycle-scale approach are more detailed than the data offered in economic-scale approach, they do not include specific information about individual processes, equipment, or other materials that could be considered negligible.

Another method for conducting an LCI is the equipment-scale approach, most frequently used in a "gate-to-gate" analysis of various processes. Full life-cycle assessments do not often include this LCI approach even though the most accurate data can be obtained from this method. The drawback of using this method is that it is very time and resource intensive. Another issue with this LCI approach is that a comparison with other products is relatively difficult to obtain in as much as a gate-to-gate inventory does not account for how the given raw materials are being produced. This type of method is usually supplemented with data from the other LCI methods.

The hybrid LCI approach combines the features of the economic, life-cycle, and equipment-scale approaches in an attempt to overcome the shortcomings of the previous methods. This approach may involve combining the economy scale with the life-cycle scale to give a cradle-to-grave analysis with more detailed information on the specific industry. A hybrid LCI may also combine the life-cycle scale LCI approach with the equipment-scale approach [104]. The advantage of this method is that it fills the gaps that are left from the life-cycle scale with data from the economic-scale, thus yielding a full "cradle-to-gate" life cycle. This type of life cycle also has the same limitations of the economy-scale approach, as described above.

There are several other methods that can be used for an LCI which include a limited LCI [89]. This method only considers metrics such as mass, water usage, energy, toxics emitted, and overall pollutants emitted. In-depth calculations are used, but all of the common metrics are not used as stated by Allen and Shonnard [53]. Another method for life-cycle assessment was proposed by Lei et al. [105], coined by the authors as "the Most of the Most." This approach consists of finding the most significant impact factors and then selecting the most significant stages of the life cycle. Although this method was designed to consider the whole process life cycle, it severely limits the size of an LCA.

Life-cycle assessment can also be conducted using the concept of exergy, which is defined as the available energy for the specific process. This method reduces the amount of double counting involved in a process [106]. According to Cornelissen and Hirs [107], an exergetic life-cycle assessment is a good tool to use in the area of depletion of natural resources, as life-cycle assessments are weak in evaluating this impact. An exergetic analysis compares the irreversibility of products, and the product with the lower irreversibility is more sustainable [108]. Cornelissen and Hirs [107] compared coal, green wood, and dried wood in an exergetic life-cycle assessment. According to this study, it appears as though using waste wood as chipboard rather than as an energy source results in less depletion. It was also found that using green wood for electricity production rather than for making chipboard gives less depletion of resources compared to using waste wood for electricity production. This difference is because the green wood needs to be dried prior to being used in chipboard production. Ukidwe and Bakshi [103] stated that the exergetic life-cycle assessment does not account for various ecological resources and suggests that a thermodynamic life-cycle assessment can be used in place of an exergetic LCA.

If there is only limited information that can be obtained for a specific process, then this information must be estimated. Jimenez-Gonzalez et al. [109] have proposed several rules of thumb. This methodology includes determining the inlet temperature and pressure, reflux ratio, fugitive losses, and a way to account for any water in contact with other chemicals as contaminated. They mention that using this approach produces results within a 20% error [110].

Streamlined life-cycle analysis has been used in cases such as in the fine chemical and pharmaceutical industry where there may be an extensive list of raw materials consisting of unique compounds where data does not exist. GlaxoSmithKline developed the Fast Life-cycle Assessment for Synthetic Chemistry, FLASC™ software tool to allow quick screening of synthesis routes from drug discovery chemistry through manufacturing [111]. Grinter [112] shows how the method was used to analyze alternative processes used to make radafaxine. The method measures the life-cycle environmental impact of materials used in the processes and provides an overall FLASC™ score.

Impact Assessment and Analysis

There are three steps recommended in a complete impact analysis: classification, characterization, and valuation. Classification involves grouping of chemical emissions into impact categories. Typical impact categories include global warming, ozone depletion, smog formation, human carcinogenicity, atmospheric acidification, aquatic and terrestrial toxicity, habitat destruction, eutrophication, and depletion of nonrenewable resources. Xun and High [113] offer that these categories can be subdivided into four classes of metrics: generic for both chemical and site, chemical specific (not accounting for environmental conditions), chemical specific in a generic environment (such as a chemical's global effects), and site and chemical specific (such as releases into a specific waterway).

The characterization step of impact assessment in LCA involves quantifying the specific impact category chosen. For example, the mass or dollar amount of material emitted is multiplied by the potential for the compound to cause the chosen impact. This provides a weighted factor for the specific impact category. When comparing two similar products and one product is higher for all impact categories, no further analysis is needed. This is rarely the case, however, thus necessitating valuation. The valuation step consists of determining which impact categories are the most significant. An example of a valuation step, the total amount of ocean and land that is needed to "buffer" the various environmental impacts may be calculated [53], the so-called "critical volume" approach. It is also possible to conduct the valuation step in terms of a monetary value. Burgess and Brennan [87] state that developing a standard for assigning relative weights to the categories is difficult because there is no clear consensus on how to carryout the valuation process. One method to approach valuation uses weighting of impact categories based on their relevance to a specific country [66, 114]. These authors also show how weighted factors can be used to determine which factor contributes most using societal factors. They use a process called Eco-Efficiency to determine which process causes the least environmental impact at the lowest normalized cost.

LCA in Practice

Typical life-cycle assessments are conducted during the product review stage of a process, after the plant, prototypes, and detailed designs of the product have been performed. However, Mueller et al. [115] state that life-cycle evaluation should be conducted starting at the planning stage of product development. They illustrate this using an example of multifunctional chip cards that are used in a wide variety of electronics. They determine the amount of material used for each board, how much is recycled, how much is incinerated, the toxic emissions, and the energy required to mine/produce the material. Their results show that the board is nearly 50% PVC, which accounts for two-thirds of the toxic emissions, but only 8.1% of the total energy required for its production. The most energy-intensive material in a multifunctional chip card is silver oxide, comprising only 2.5% of the product by weight, but requiring the most energy to produce. If LCA were applied at the product design stage of a chip card, alternatives to PVC and silver oxide could be explored and compared in terms of their contribution to emissions and energy requirements.

There have been a number of studies on fuels used for electricity generation. One of these studies focuses on the use of natural gas, heavy oil, or coal in cogeneration of electricity [116]. Using a numerical "eco-load total standardized" evaluation system, these authors found that coal had the lowest eco-load of all alternatives considered. In another LCA study, Goralczyk [117] compares hydroelectric, photovoltaic cells, wind turbines, oil, coal, and natural gas and quantifies that electricity from hydropower had the least environmental impact. Schleisner [118] focuses on wind farms in an LCI study that focuses on the materials used to manufacture the windmills and reports that 2% of the electricity generated during the windmill's lifetime is used to manufacture the windmill components.

A life-cycle assessment for various forms of production of hydrogen has also been conducted. Koroneos et al. [119] examine six methods used to manufacture hydrogen: photovoltaic cells, solar thermal energy, wind power, hydroelectric power, biomass degradation, and natural gas steam reforming. These authors show that wind power yields the lowest environmental impacts for greenhouse gas formation, acidification, eutrophication, and smog formation, whereas photovoltaic cells result in the largest total environmental impact. As a result of this study, these authors recommend the use of wind power, hydropower, and solar thermal power to produce hydrogen because these are the "most environmentally friendly methods."

Life-cycle assessments have also been conducted on transportation fuels. Furuholt [120] has conducted a study comparing the production of diesel, gasoline, and gasoline with MTBE in Norway. The factors that were considered in this study were global warming, photo-oxidant formation, acidification, eutrophication, fossil energy, and solid waste. The impacts in the environmental categories listed above were conducted on the basis of 1 MJ of energy. In both analyses it was found that gasoline with MTBE contributes the most to the impact categories. It was also found that diesel fuel and gasoline have approximately the same scaled values for acidification, eutrophication, and solid waste, but gasoline without additives has 1.5 times the global warming, 2.6 times the photochemical oxidant formation, and uses 1.5 times the fossil energy as diesel.

Another study on automobile fuel options was conducted by MacLean and Lave [121]. This study focused on light-duty vehicles and the CO_2 equivalent gases released during manufacture, gasoline refining, operation, maintenance, and other services. The authors cite that 73% of the greenhouse equivalent gases are released during operation and propose viable alternatives to the use of gasoline in vehicles. They emphasize that, although battery-powered vehicles have zero emissions, there are other factors that give this alternative a negative environmental impact, such as the use of heavy metals. Hybrid vehicles are also discussed as an alternative, but they state that the higher sales price of the Toyota Prius is not justified by fuel savings, emissions reductions, or a combination of the two. Diesel fuel, another fuel alternative, has a well-to-tank efficiency of 24%, whereas gasoline only has a 20% efficiency. There are also many drawbacks to the use of diesel, including higher NO_x and particulate matter emissions, possibly carcinogens. MacLean and Lave also cite that fuel cell vehicles are 20 years away from wide-scale use. Ethanol is a viable option as a fuel source, and there are two renewable processes that can be used to obtain ethanol. The first method is from plant cellulosic material. The well-to-tank efficiency for this material ranges from 80 to 95%, and the emissions are 15 g CO_2 equivalent gases/MJ. Ethanol can also be manufactured using corn. This process releases six times as much CO_2 equivalent gases/MJ as the previous process.

The life-cycle for the production of biodiesel from renewable feedstock has been studied. Biodiesel is a renewable fuel that can replace traditional petroleum-derived diesel fuels. It is typically made through a transesterification reaction of triglycerides derived from fats, oils, and alcohols. Narayanan et al. [122] used an analytical hierarchy process method to determine that soybean oil was the most sustainable raw material when compared to rapeseed, sunflower, and beef tallow. Dinh et al. [123] incorporated environmental, economical, safety, raw material, and fuel performance criteria in comparing biodiesel production routes. Algae received better sustainable development scores in their study than did soybean, rapeseed, jatropha, and palm oil. This was due to the lower green house gas emissions, water usage, and land area needed.

Other than fuels, life-cycle assessments have been performed on other sectors of the transportation industry. One study focused on the catalytic converters for passenger cars [124]. The goal of this life-cycle assessment was to compare the life-cycle impacts of a catalytic converter and the environmental benefits in terms of emission reductions through the exhaust pipe. The study on catalytic converters involved a cradle-to-grave approach, but excluded the mining and transportation of raw materials because no data were available. The criteria used for environmental loads were global warming potential, waste, eutrophication, acidification, resource use, and photochemical ozone creation potential. It was found from this study that waste and global warming are drastically increased, but acidification, eutrophication, and photochemical ozone creation potential are drastically decreased as a result of converter use. Auxiliary power units for diesel trucks were compared by Baratto and Diwekar [125]. The environmental criteria used were the same criteria as mentioned by Allen and Shonnard [53], but also included the toxicity to humans, terrestrial species, and aquatic species. It was found the auxiliary power unit had the least impact for all the categories. An economic analysis was also conducted, and it was found that the payback period was a little over 2 years.

Eagen and Weinberg [126] conducted a life-cycle assessment on two different anodizing processes, differing in the mixture of boric and sulfuric acid or chromic acid used. Boric and sulfuric acid are shown to be a better choice than a mixture of boric and chromic acid. Tan et al. [127] have conducted a cradle-to-gate life-cycle assessment of an aluminum billet, which included the mining of bauxite, the processing of the alumina, and the final casting process for three plants located in Australia. Four different scenarios were analyzed—a base case, a reduction of scrap metal, a more sustainable practice for the smelter, and the latter with clean coal technology—and they conclude that implementation of the last case, clean coal technology, decreases all emissions considered and decreases the global warming potential by 21%.

Life-cycle assessments have been conducted on processes involved in the pharmaceutical industry. A study conducted by Jodicke et al. [128] focuses on one processing step of an intermediate using either a metal catalyst or bio-catalyst. These authors show that the solvents used in the extraction of the product play a large role in the environmental impacts. Jimenez-Gonzalez et al. [129] conducted a cradle-to-gate life cycle for a pharmaceutical product of GlaxoSmithKline. The metrics used for this study were eutrophication, acidification, greenhouse gases, photochemical ozone creation, energy, and mass requirements. This study reports that eutrophication, ozone creation, total organic carbon, energy, and raw materials are most affected by this process, whereas greenhouse gas formation and acidification are most affected by energy use requirements. The manufacturing process is broken down further and shows the impact of solvents, chemicals, and internal drug manufacture on the environmental criteria listed above. Solvent selection also contributes significantly to the impact of the manufacture of a pharmaceutical product. Solvents contribute 75% to the energy use, 80% of the mass (excluding water), and 70% of the ozone depletion. Energy also contributes 70% to resource depletion and 90% to greenhouse gas emissions. Jimenez-Gonzalez and Overcash [130] compared two processes for making sertraline: the THF and TOL processes. These two processes were analyzed from the lab scale to the production scale. In comparing energy usage between the lab and production scales energy usage decreased by about 70%. It was also found that there is no significant energy difference during the final production stage of the product with regard to the two different processes. The results of an eco-efficiency analysis on Vitamin B2 production was reported by Wall-Markowski et al. [131]. In this study they compared vitamin B2 produced from three fermentation processes and one chemical process. One of the three bioprocesses outperformed the others in both economic and ecological indicators. Based on the results of this study BASF recently started vitamin B2 production using a one-step fermentation from vegetable oil.

Assessments of green designs for solvent recovery in the pharmaceutical industry have been evaluated by Slater et al. [132, 133]. These studies have evaluated the life-cycle impact of using various separation processes for solvent purification and reuse within active pharmaceutical ingredient (API) synthesis [134]. Membrane pervaporation has been proposed as a “green drying” alternative to entrainer-based constant volume distillation in one step of the synthesis of an investigational Bristol-Myers Squibb oncology drug [135]. The life-cycle assessment for this study compared the two process routes to dehydrate a batch process vessel that contained the solvent tetrahydrofuran. The pervaporation membrane system allows for a 95% reduction in greenhouse gas emissions through purifying and reusing the solvent within the process step. This avoids the emissions associated with purchasing the tetrahydrofuran entrainer used in constant volume distillation and its disposal. In addition, the economics are favorable, since tetrahydrofuran is a relatively costly solvent. Life-cycle assessments have also been performed for various separation strategies in the recovery of isopropanol in the Pfizer celecoxib synthesis [136]. The case study evaluated seven different process routes which included combinations of distillation, molecular sieve adsorption, and membrane pervaporation. A proposed process of using distillation and pervaporation to recover isopropanol reduced CO_2 life-cycle emissions by 95% from the base case of solvent incineration. This study showed that the emissions generated by operating the separations equipment are small when compared to the overall life-cycle emissions reduced from virgin solvent manufacture and waste incineration. This study also evaluated the life-cycle emissions reduced from concentrating waste solvents for sale as a generic solvent.

Life-cycle assessments have also been conducted on seawater desalination technologies. One study was conducted by Raluy et al. [137] used SimaPro 5.0 software [138] for the analysis portion. Three desalination technologies were compared: multi-effect distillation, multistage flash, and reverse osmosis. The study focused on the environmental criteria of CO_2, NO_x, nonmethane volatile organic compounds (NMVOC), and SO_x. The analysis focused on integrating the distillation process and flash process with a cogeneration plant and with reverse osmosis. This study also compared different regions with the primary difference related to the type of fuel used for electricity production. It was found that in models using the average of the European countries, as well as in models of Spain and Portugal alone, that the multistage flash had the least environmental burdens. However, the French and Norwegian models showed that reverse osmosis had the least environmental negative effects using Eco-indicator 99, Ecopoint 97, and CML 2 baseline impact-assessment methods that accompany the software. Raluy et al. [137] also stated that using a hybrid plant cuts down energy usage by 75%.

There have also been a number of LCA studies performed on industrial paint coatings. One study done by Shonnard et al. [114] compared five different coating processes for wooden doors. It appears from this analysis that the UV coating process has the least risk potential, raw material consumption, emissions, and energy consumption. Papasavva et al. [139] also conducted a life-cycle assessment on paints used in the automotive industry. This study focused on three types of coating materials: primer, basecoat, and clear coat. Three primers were investigated, one solvent-based and two powders (acrylic and polyester). The two basecoats that were used were waterborne and white and pewter in color, respectively. The two clear coats were both acrylic, but one was a solvent-based and the other, powder-based. The white basecoat was chosen for the article because the energy required to produce either basecoat was approximately the same. The criteria used by Papasavva et al. were material requirements, energy consumption, atmospheric emissions, water emissions, solid waste emissions, particulate matter, SO_x, NO_x, CO, VOC, and CO_2. They show that there is a trade-off among environmental factors. These authors report that using the combination of powder primer, water basecoat, and powder clearcoat (PP2–WB1–PC2) results in the least energy requirement, water consumption, solid waste, and VOC emissions. However, this combination of surface coatings exceeded the other scenarios in PM, SO_x, and CO_2 equivalent air emissions. An interesting finding by Dobson [140] is that use of water-based and solvent-based paints yielded the same environmental impacts as incineration of the VOCs.

The pulp and paper industry is another industry where life-cycle assessment methodology has been applied. One example is a paper by Lopes et al. [141], who compared the two major fuels used in the pulp and paper industry: fuel oil and natural gas. The environmental categories were the same categories listed by Allen and Shonnard [53]. The use of methane in place of fuel oil decreases all of the environmental parameters except photochemical ozone formation, which does not vary between fuel options.

Other life-cycle assessments have been conducted on postconsumer-use recycling. One of these studies by Rios et al. [142] focused on the end-of-life recycling of plastics used in electronics. The focus of this paper was on the separation and sorting of various types of plastics for recycling. Another study [53] compared PET bottles to glass bottles with and without recycling. It was found that recycling the PET bottles has the least environmental impact and the same normalized cost as the glass bottle. Song and Hyun [143] focused on the recycling aspects of PET bottles for 11 different scenarios and showed that as the percent of bottles collected increases, the energy used for collection also increases. They cited that the recycle pathway which produces the least CO_2, SO_x, and NO_x was the closed loop and landfill pathway, and the pathway which produced the least solid waste was the pathway for pyrolysis and incineration. Ekvall [144] conducted a life-cycle analysis for recycling newsprint in Sweden. Shiojiri et al. [145] conducted a life-cycle assessment on sulfur hexafluoride. There were different ways to use and to recycle the sulfur hexafluoride. From the study conducted on sulfur hexafluoride "energy consumption as well as global warming risk can be reduced by using a mixture of SF_6 with nitrogen as an insulating gas compared to SF_6," but the other environmental impacts will increase due to transportation to the recycling plant.

There have been life-cycle assessments conducted on waste management. Jimenez-Gonzalez et al. [109] conducted a partial LCI on three different waste treatments for pharmaceutical waste: waste-water treatment, incineration, and solvent recovery. Life-cycle assessment has been used to analyze technologies for the treatment of wastewater produced in the extraction of oil and gas in the energy industry [146]. The study evaluated environmental impacts of 20 processes such as dissolved air, rotating biological contactors, air stripping, dual media filtration, membrane filtration, and ion exchange used in successive stages of water treatment. The impact assessment included an analysis of various end-use options for the reused water in both agricultural and industrial scenarios. The study showed that dissolved air flotation, adsorption, dual media filtration, and reverse osmosis produce a low environmental impact for a treatment system. Chevalier et al. [100] compared two flue gas treatment processes for waste incineration using life-cycle assessment.

Other life-cycle assessments have been done on consumer textile products. One example is a life-cycle assessment on the production of clothing for the production of various indigo dyes for dying denim [114]. A 2D plot was used with the axis being cost and normalized environmental impact. It was found that dying the denim electrochemically in a 40% vat solution has the least environmental impact and cost compared to dying the denim using indigo plants.

The consumer products industry has embraced life-cycle assessment in promoting sustainability. Procter & Gamble has performed life-cycle assessment of its product line which includes laundry detergents, paper towels, disposable diapers, shampoos, personal care products, cosmetics, bathroom tissues, liquid dish wash, and liquid fabric softener [147]. Their study was able to determine the stage in the life-cycle that makes the greatest impact on the energy footprint of these products. In the case of laundry detergents, the energy demand is mainly a result of the product use in the home-heating of water to do the laundry. Therefore, a laundry detergent designed to be used in cold water would have the greatest impact on the environment. Tide Cold Water (US market) and Ariel Cool Clean (European market) were formulated as sustainable innovation products. An estimated 34 million tons of CO_2 per year would be saved in the United States if everyone switched to cold water washing. Life-cycle analysis also shows that materials contribute the next highest impact on the carbon footprint. Through production concentration, a.k.a., compaction, savings in packaging and transportation are achieved.

There have also been life-cycle assessments conducted on the food industry. In the study by Cederberg and Mattson [148] organic milk farms were compared to conventional milk farms. They show an increase in global warming in methane due to organic farming, but a decrease in other compounds such as carbon dioxide and N_2O. Most of the acidification potential for these farms was caused by the release of ammonia from manure. The eutrophication parameter was estimated for this study and was based on the manure application rate and a higher phosphorous surplus on conventional farms. They stated that organic farming reduces pesticide use, global warming, acidification, and eutrophication. Zabaniotou and Kassidi [149] conducted a study on two types of egg packaging material, recycled paper and polystyrene. The functional unit for this study was done on a packaging basis instead of a product mass basis. The environmental factors used in the study for egg containers were global warming, ozone depletion, acidification, eutrophication, particulate matter, heavy metals, carcinogenic substances, and photochemical ozone creation potential. This study concluded that polystyrene packages contribute more to acidification potential, winter and summer smog formation, whereas recycled paper egg packages contribute more to heavy metal and carcinogenetic substances impact. Another study on food packaging was conducted by Bohlmann [150], which focused on a comparison of polypropylene and biodegradable packaging. The functional unit for this case study was the packaging required to fill 1,000 kg of yogurt. The environmental criteria used for this comparison were energy and greenhouse gases. It appears as though the biodegradable package consumes less energy, but has slightly higher greenhouse gas emissions. Bohlmann cited that the greenhouse gas emissions are equivalent if the biodegradable packaging is fully decomposed in the landfill.

Anderson et al. [151] report a life-cycle assessment on tomato ketchup, in an attempt to identify “hot spots” in the life cycle of the product. In this study, it was found that the food-processing stage contributes the most to greenhouse gases, human toxicity, and acidification. They also suggest that the major contributor to eutrophication is the agriculture sector of the economy.

Conclusions

Life-cycle assessment is an important green engineering tool for analyzing processes and products. This tool is currently being used by industry to improve the environmental and economic performance of products and processes. Examples have been given in this chapter of life-cycle assessment methods as well as many studies done on chemical processes and products. Through the use of life-cycle assessment tools industry will move to sustainable production of current and new products.

Acknowledgments P. A. Ramachandran would like to thank National Science Foundation for partial support of research in the general area of environmentally benign processing through the grant NSF-ERC center grant EEC-0310689.

References

1. EPA Green Engineering Web site. www.epa.gov/oppt/greenengineering
2. EPA Green Chemistry Web site, 12 principles of green chemistry. www.epa.gov/oppt/greenchemistry/principles.html

3. Anastas PT, Warner JC (1998) Green chemistry: theory and practice. Oxford University Press, New York
4. Malone MF, Huss RS (2003) Green chemical engineering aspects of reactive distillation. Environ Sci Technol 37(23):5325–5329
5. Allen DT, Shonnard DR (2004) Green engineering: environmentally conscious design of chemical processes, 3rd edn. Prentice-Hall, Englewood Cliffs, NJ
6. Douglas JM (1988) Conceptual design of chemical processes. McGraw-Hill, New York
7. Rossiter AP, Klee H (1995) Hierarchical process review for waste minimization. In: Rossetier AP (ed). Waste Minimzation through process Design McGraw-Hill, New York
8. Schultz MA, Stewart DG, Harris JM, Rosenblum SP, Shakur MS, O'Brien DE (2002) Reduce costs with dividing-wall columns. Chem Eng Process May:64–71
9. El-Halwagi MM (1997) Pollution prevention through process integration. Academic, San Diego, CA
10. Liu YA, Lucas B, Mann JM (2004) Up-to-date tools for water system optimization. Chem Eng 111(1):30–41
11. Allen DT, Rosselot KS (1997) Pollution prevention for chemical processes. Wiley, New York
12. Tunca C, Ramachandran PA, Dudukovic MP (2004) Role of chemical reaction engineering in sustainable development. In: Paper presented at AIChE session 7d, Austin, TX, 10 Nov 2004
13. Mulholland KL, Dyer JA (1999) Pollution prevention: methodology, technologies, and practices. American Institute of Chemical Engineers, New York, p 214
14. Ney RE Jr (1990) Where did that chemical go? Van Nostrand Reinhold, New York
15. Mackay D, Patterson S, diGuardo A, Cowan CE (1996) Evaluating the environmental fate of a variety of types of chemicals using the EQC model. Environ Toxicol Chem 15:1627–1637
16. Andren AW, Mackay D, Depinto JV, Fox K, Thibodeaux LJ, McLachlan M, Haderlein S (2000) Intermedia partitioning and transport. In: Klečka G, Boethling B, Franklin J, Grady L, Graham D, Howard PH, Kannan K, Larson B, Mackay D, Muir D, van de Meent D (eds) Evaluation of persistence and long-range transport of organic chemicals in the environment. SETAC Press, Pensacola, FL, pp 131–168
17. Franklin J, Atkinson R, Howard PH, Orlando JJ, Seigneur C, Wallington TJ, Zetzsch C (2000) Quantitative determination of persistence in air. In: Klečka G, Boethling B, Franklin J, Grady L, Graham D, Howard PH, Kannan K, Larson B, Mackay D, Muir D, van de Meent D (eds) Evaluation of persistence and long-range transport of organic chemicals in the environment. SETAC Press, Pensacola, FL, pp 7–62
18. Thibodeaux LJ (1996) Environmental chemodynamics: movement of chemicals in air, water, and soil, 2nd edn. Wiley, New York
19. Clark MM (1997) Modeling for environmental engineers and scientists. Wiley, New York
20. Schnoor JL (1997) Environmental modeling. Wiley, New York
21. Bishop PL (2004) Pollution prevention: fundamentals and practice. Waveland Press, Long Grove, IL
22. Thomas RG (1990) Volatilization from water. In: Lyman WJ, Reehl WF, Rosenblatt DH (eds) Handbook of chemical property estimation methods. American Chemical Society, Washington, DC, pp 15-1–15-34
23. Larson RA, Weber EJ (1994) Reaction mechanisms in environmental organic chemistry. Lewis Publishers, Boca Raton, FL
24. Lerman A (1971) Time to chemical steady states in lakes and oceans. In: Hem JD (ed) Nonequilibrium systems in natural water chemistry. Advances in chemistry series #106. American Chemical Society, Washington, DC, pp 30–76
25. Thibodeaux LJ, Valsaraj KT, Reible DD (1993) Associations of polychlorinated biphenyls with particles in natural waters. Water Sci Technol 28(8):215–221
26. Baum EJ (1998) Chemical property estimation: theory and application. Lewis Publishers, Boca Raton, FL
27. McCutcheon SC, Schnoor JL (2003) Overview of phytotransformation and control of wastes. In: McCutcheon SC, Schnoor JL (eds) Phytoremediation: transformation and control of contaminants. Wiley, New York, pp 3–58
28. Larson R, Forney L, Grady L Jr, Klečka GM, Masunanga S, Peijnenburg W, Wolfe L (2000) Quantification of persistence in soil, water, and sediments. In: Klečka G, Boethling B, Franklin J, Grady L, Graham D, Howard PH, Kannan K, Larson B, Mackay D, Muir D, van de Meent D (eds) Evaluation of persistence and long-range transport of organic chemicals in the environment. SETAC Press, Pensacola, FL, pp 63–130
29. Gibson DT, Subramanian V (1984) Microbial degradation of aromatic compounds. In: Gibson DT (ed) Microbial degradation of organic compounds. Marcel Dekker, New York
30. Bedard DL, Quensen JF III (1995) Microbial reductive dechlorination of polychlorinated biphenyls. In: Young LY, Cerniglia CE (eds) Microbial transformation and degradation of toxic organic chemicals. Wiley, New York
31. Fish KM, Principe JM (1994) Biotransformations of Arochlor 1242 in Hudson River Test Tube microcosms. Appl Environ Microbiol 60(12):4289–4296
32. Ye D, Quensen JF III, Tiedje JM, Boyd SA (1995) Evidence for para-dechlorination of polychlorobiphenyls by methanogenic bacteria. Appl Environ Microbiol 61:2166–2171
33. Chakrabarty AM (1982) Biodegradation and detoxification of environmental pollutants. CRC Press, Boca Raton, FL
34. Alexander M (1994) Biodegradation and bioremediation. Academic, San Diego
35. Young LY, Cerniglia CE (1995) Microbial transformation and degradation of toxic organic chemicals. Wiley-Liss, New York
36. Burken JG (2004) Uptake and metabolism of organic compounds: green liver model. In: McCutcheon SC, Schnoor JL (eds) Phytoremediation: transformation and control of contaminants. Wiley, New York, pp 59–84
37. Jeffers PM, Wolfe NL (1997) Degradation of methyl bromide by green plants. In: Seiber JN (ed) Fumigants: environmental fate exposure and analysis. American Chemical Society, Washington, DC
38. O'Neill W, Nzengung V, Noakes J, Bender J, Phillips P (1998) Biodegradation of tetrachloroethylene and trichloroethylene using mixed-species microbial mats. In: Wickramanayake GB, Hinchee RE (eds) Bioremediation and phytoremediation. Batelle, Columbus, WA, pp 233–237
39. Hughes JB, Shanks J, Vanderford M, Lauritzen J, Bhadra R (1997) Transformation of TNT by aquatic plants and plant tissue cultures. Environ Sci Technol 31:266–271
40. Vanderford M, Shanks JV, Hughes JB (1997) Phytotransformation of trinitrotoluene (TNT) and distribution of metabolic products in myriphyllum aquaticum. Biotechnol Lett 199:277–280
41. Gao J, Garrison AW, Hoehamer C, Mazur C, Wolfe NL (1999) Phytotransformations of organophosphate pesticides using axenic plant tissue cultures and tissue enzyme extract. In situ and on-site bioremediation. In: The fifth international symposium, San Diego, 19–22 April 1999
42. Cunningham SD, Berti WR (1993) The remediation of contaminated soils with green plants: an overview. In Vitro Cell Dev Biol Plant 29:207–212
43. Banks MK, Schwab AP, Govindaraju RS, Kulakow P (1999) Phytoremediation of hydrocarbon contaminated soils. In: Fiorenza S, Oubre LC, Ward CH (eds) Phytoremediation. CRC Press, New York
44. Schwartzenbach RP, Gschwend PM, Imboden DM (1993) Environmental organic chemistry, 1st edn. Wiley, New York
45. Wolfe NL, Jeffers PM (2000) Hydrolysis. In: Boethling RS, Mackay D (eds) Handbooks of property estimation methods for

chemicals: environmental and health science. CRC Press, Boca Raton, FL, pp 311–334

46. Zepp RG (1982) Experimental approaches to environmental photochemistry. In: Hutzinger O (ed) The handbook of environmental chemistry, Vol. 2, Part B. Springer-Verlag, Berlin, pp 19–41
47. Alebić-Juretić A, Güsten H, Zetzsch C (1991) Absorption spectra of hexachlorobenzene adsorbed on SiO_2 powders. Fresenius J Anal Chem 340:380–383
48. Bermen JM, Graham JL, Dellinger B (1992) High temperature UV absorption characteristics of three environmentally sensitive compounds. J Photochem Photobiol A Chem 68:353–362
49. Tysklind M, Lundgrenand K, Rappe C (1993) Ultraviolet absorption characteristics of all tetra-to octachlorinated dibenzofurans. Chemosphere 27:535–546
50. Kwok ESC, Arey J, Atkinson R (1994) Gas-phase atmospheric chemistry of dibenzo-p-dioxin and dibenzofuran. Environ Sci Technol 28:528–533
51. Funk DJ, Oldenborg RC, Dayton DP, Lacosse JP, Draves JA, Logan TJ (1995) Gas-phase absorption and later-induced fluorescence measurements of representative polychlorinated dibenzo-p-dioxins, polychlorinated dibenzofurans, and a polycyclic aromatic hydrocarbon. Appl Spectros 49:105–114
52. Konstantinou IK, Zarkdis AK, Albanis TA (2001) Photodegradation of selected herbicides in various natural waters and soils under environmental conditions. J Environ Qual 30:121–130
53. Allen DT, Shonnard DR (2002) Green engineering: environmental conscious design of chemical processes. Prentice-Hall, Upper Saddle River, NJ
54. Windholz M, Budavara S, Blumetti RF, Otterbein ES (1983) The Merck Index: an encyclopedia of chemicals, drugs, and biologicals, 6th edn. Merck & Co., Rahway, NJ
55. Howard PH, Boethling RS, Jarvis WF, Meylan WM, Michalenko EM (1991) Handbooks of environmental degradation rates. Lewis Publishers, Chelsea, MI
56. Dean JA (1992) Lange's handbook of chemistry, 14th edn. McGraw-Hill, New York
57. Lide DR (1994) CRC handbook of chemistry and physics, 74th edn. CRC Press, Boca Raton, FL
58. Mackay D, Shiu WY, Ma KC (1992–1997) Illustrated handbook of physical chemical properties and environmental fate of organic chemicals, vols 1–5. Lewis Publishers, Boca Raton, FL
59. Howard PH, Meylan WM (1997) Handbook of physical properties of organic chemicals. CRC Press, Boca Raton, FL
60. Tomlin CDS (ed) (1997) The pesticide manual, 11th edn. British Crop Protection Council, Farnham, Surrey
61. Yaws CL (1999) Chemical properties handbook. McGraw-Hill, New York
62. Verschueren K (2001) Handbook of environmental data on organic chemicals, 3rd and 4th edns. Van Nostrand-Reinhold, New York
63. Lyman WJ, Reehl WF, Rosenblatt DH (1982) Handbook of chemical property estimation methods: environmental behavior of organic compounds. McGraw-Hill, New York
64. Neely WB, Blau GE (1985) Environmental exposure from chemicals, vols I and II. CRC Press, Boca Raton, FL
65. Bare JC, Norris GA, Pennington DW, McKone T (2003) TRACI: the tool for the reduction and assessment of chemical and other impacts. J Indust Ecol 6(3–4):49–78
66. Goedkoop M (1995) The Eco-indicator 95, final report. Netherlands Agency for Energy and the Environment (NOVEM) and the National Institute of Public Health and Environmental Protection (RIVM), NOH report 9523
67. Heijungs R, Guinée JB, Huppes G, Lankreijer RM, Udo de Haes HA, Wegener Sleeswijk A (1992) Environmental life cycle assessment of products. Guide and backgrounds. NOH Report Numbers 9266 and 9267, Netherlands Agency for Energy and the Environment, Nov 1992
68. ISO 14040–14049 (1997–2002) Environmental management—life cycle assessment. International Organization for Standardization, Geneva, Switzerland
69. SETAC, Society for Environmental Toxicology and Chemistry (1993) Guidelines for life-cycle assessment: code of practice. Brussels, Belgium
70. Douglas JM (1992) Process Synthesis for Waste Minimization. Ind Eng Chem Res 31(1):238–243
71. PARIS II (2005). http://www.tds-tds.com/
72. Chen H, Shonnard DR (2004) A systematic framework for environmental-conscious chemical process design: early and detailed design stages. Ind Eng Chem Res 43(2):535–552
73. Allen DT, Shonnard DR (2001) Green engineering: environmentally conscious design of chemical processes and products. AIChE J 47(9):1906–1910
74. NRC (National Research Council) (1983) Risk assessment in the federal government: managing the process. Committee on Institutional Means for Assessment of Risks to Public Health. National Academy Press, Washington, DC
75. Air CHIEF. The air ClearingHouse for inventories and emission factors, CD-ROM. http://www.epa.gov/oppt/greenengineering/software.html. Accessed 2005
76. Mackay D, Shiu W, Ma K (1992) Illustrated hand book of physical-chemical properties and environmental fate for organic chemicals, vol 1–4. Lewis Publishers, Chelsea, MI
77. Shonnard DR, Hiew DS (2000) Comparative environmental assessments of VOC recovery and recycle design alternatives for a gaseous waste stream. Environ Sci Technol 34(24):5222–5228
78. WAR (WAste Reduction Algorithm). http://www.epa.gov/oppt/greenengineering/software.html
79. SACHE, Safety and Chemical Engineering Education. American Institute of Chemical Engineers. http://www.sache.org. Accessed 2005
80. Shonnard DR (2005) Tools and materials for green engineering and green chemistry education, green chemistry and engineering education—a workshop organized by the Chemical Sciences Roundtable of the National Research Council, 7–8 Nov 2005
81. Genco JM (1991) Pulp. In: Kroschwitz JI (ed) Kirk-Othmer encyclopedia of chemical technology, vol 20. Wiley, New York, p 493
82. Collins TJ, Horwitz C, Gordon-Wylie SW (1999) TAML™ Activators: general activation of hydrogen peroxide for green oxidation processes, provided by Mary Kirchhoff. Green Chemistry Institute, American Chemical Society
83. Miller K (2005) Comments at the panel discussion in the session "Building the business case for sustainability," AIChE spring meeting, Atlanta, 12 April 2005
84. Schrott W, Saling P (2000) Eco-efficiency analysis—testing products for their value to the customer. Melliand Textil 81 (3):190, 192–194
85. Landsiedel R, Saling P (2002) Assessment of toxicological risks for life cycle assessment and eco-efficiency analysis. Int J Life Cycle Assess 7(5):261–268
86. Azapagic A (1999) Life cycle assessment and its application to process selection, design, and optimization. Chem Eng J 73(1):1–21
87. Burgess AA, Brennan DJ (2001) Application of life cycle assessment to chemical processes. Chem Eng Sci 56, 8 (April) 2589, 2609
88. Royal Commission on Environmental Pollution (1988) Best practicable environmental option. Twelfth report, Cm130, London, England
89. Beaver ER (2004) Calculating metrics for acetic acid production. In: AIChE sustainability engineering conference proceedings, Austin, TX, Nov 2004, pp 7–15
90. International Organization of Standardization (ISO) (1997) Environmental management—life cycle assessment—principles and

framework. International Organization of Standardization, Geneva, Switzerland (International Standard ISO14040:1997(E))
91. International Organization of Standardization (ISO) (1998) Environmental management—life cycle assessment—goal and scope definition and inventory analysis. International Organization of Standardization, Geneva, Switzerland (International Standard ISO14041:1998(E))
92. International Organization of Standardization (ISO) (2000) Environmental management—life cycle assessment—life cycle impact assessment. International Organization of Standardization, Geneva, Switzerland (International Standard ISO14042:2000(E))
93. International Organization of Standardization (ISO) (2000) Environmental management—life cycle assessment—life cycle interpretation. International Organization of Standardization, Geneva, Switzerland (International Standard ISO14043:2000(E))
94. International Organization for Standardization (ISO) (2006) Environmental management—life-cycle assessment—principles and framework. International Organization for Standardization, Geneva, Switzerland (International Standard ISO 14040:2006)
95. International Organization for Standardization (ISO) (2006) Environmental management—life-cycle assessment—requirements and guidelines. International Organization for Standardization, Geneva, Switzerland (International Standard ISO 14044:2006)
96. Ukidwe NW, Bakshi BR (2004) Economic versus natural capital flows in industrial supply networks—implications to sustainability. In: AIChE sustainability engineering conference proceedings, Austin, TX, Nov 2004, pp 145–153
97. Graedel TE (1998) Streamlined life-cycle assessment. Prentice-Hall, Englewood Cliffs, NJ
98. Curran M (ed) (1997) Environmental life-cycle assessment. McGraw Hill, New York
99. Azapagic A, Perdan S, Clift R (2004) Sustainable development in practice: case studies for engineers and scientists. Wiley, New York
100. Chevalier J, Rousseaux P, Benoit V, Benadda B (2003) Environmental assessment of flue gas cleaning processes of municipal solid waste incinerators by means of the life cycle assessment approach. Chem Eng Sci 58(10):2053–2064
101. Ekvall T, Finnveden G (2001) Allocation in ISO 14041—a critical review. J Cleaner Prod 9(3):197–208
102. Bakshi BR, Hau JL (2004) A multiscale and multiobjective approach for environmentally conscious process retrofitting. In: AIChE sustainability engineering conference proceedings, Austin, TX, Nov 2004, pp 229–235
103. Ukidwe NW, Bakshi BR (2004) A multiscale Bayesian framework for designing efficient and sustainable industrial systems. In: AIChE sustainability engineering conference proceedings, Austin, TX, Nov 2004, pp 179–187
104. Suh S, Lenzen M, Treloar GJ, Hondom H, Harvath A, Huppes G, Jolliet O, Klann U, Krewitt W, Morguchi Y, Munksgaard J, Norris G (2004) System boundary selection in life-cycle inventories using hybrid approaches. Environ Sci Technol 38(3):657–663
105. Lei L, Zhifeng L, Fung R (2003) "The most of the most"—study of a new LCA method. In: IEEE proceedings, San Francisco, CA pp 177–182
106. Bakshi BR, Hau JL (2004) Using exergy analysis for improving life cycle inventory databases. In: AIChE sustainability engineering conference proceedings, Austin, TX, Nov 2004, pp 131–134
107. Cornelissen RL, Hirs GG (2002) The value of the exergetic life cycle assessment besides the LCA. Energy Convers Manage 43:1417–1424
108. Becalli G, Cellura M, Mistretta M (2003) New exergy criterion in the "multi-criteria" context: a life cycle assessment of two plaster products. Energy Conserv Manage 44:2831–2838
109. Jimenez-Gonzalez C, Overcash MR, Curzons A (2001) Waste treatment modules—a partial life cycle inventory. J Chem Technol Biotechnol 76:707–716
110. Jimenez-Gonzalez C, Kim S, Overcash MR (2000) Methodology for developing gate-to-gate life cycle inventory information. Int J Life Cycle Assess 5:153–159
111. Curzons AD, Jimenez-Gonzalez C, Duncan AL, Constable DJC, Cunningham VL (2007) Fast life cycle assessment of synthetic chemistry FLASC tool. Int J Life Cycle Assess 12:272–280
112. Grinter T (2010) The development of an environmentally sustainable process for radafaxine. In: Dunn P, Wells A, Williams T (eds) green chemistry in the pharmaceutical industry. Wiley-VCH, Weinheim, pp 197–219
113. Xun J, High KA (2004) A new conceptual hierarchy for identifying envirnomental sustainability merrics. Environmental progress 23(4):291–301
114. Shonnard DR, Kichere A, Saling P (2003) Industrial applications using BASF eco-efficiency analysis: perspectives on green engineering principles. Environ Sci Technol 37:5340–5348
115. Mueller J, Griese H, Schischke K, Stobbe I, Norris GA, Udo de Haes HA (2004) Life cycle thinking for green electronics: basics in ecodesign and the UNEP/SETAC life cycle initiative. In: International IEEE conference on Asian green electronics, Hong Kong and Shenzhen china, pp 193–199
116. Widiyanto A, Kato S, Maruyama N, Kojima Y (2003) Environmental impact of fossil fuel fired co-generation plants using a numerically standardized LCA scheme. J Energy Res Technol 125:9–16
117. Goralczyk M (2003) Life-cycle assessment in the renewable energy sector. Appl Energy 75:205–211
118. Schleisner L (2000) Life cycle assessment of a wind farm and related externalities. Renew Energy 20:279–288
119. Koroneos C, Dompros A, Roumbas G, Moussiopoulos N (2004) Life cycle assessment of hydrogen fuel production processes. Int J Hydrogen Energy 29:1443–1450
120. Furuholt E (1995) Life cycle assessment of gasoline and diesel. Resources Conserv Recycl 14:251–263
121. MacLean HL, Lave LB (2003) Life cycle assessment of automobile/fuel options. Environ Sci Technol 37:5445–5452
122. Narayanan D, Zhang Y, Mannan MS (2007) Engineering for sustainable development (ESD) in biodiesel production. Trans IChem E 85:349–359
123. Dinh LTT, Guo Y, Mannan MS (2009) Sustainability evaluation of biodiesel production using multicriteria decision making. Environ Prog Sust Energy 28:38–46
124. Amatayakul W, Ramnas O (2001) Life cycle assessment of a catalytic converter for passenger cars. J Cleaner Prod 9:395–403
125. Baratto F, Diwekar UM (2005) Life cycle of fuel cell-based APUs. J Power Sources 139:188–196
126. Eagan P, Weinberg L (1999) Application of analytic hierarchy process techniques to streamlined life-cycle analysis of two anodizing processes. Environ Sci Technol 33:1495–1500
127. Tan R, Khoo BH, Hsien H (2005) An LCA study of a primary aluminum supply chain. J Cleaner Prod 13(6):607–618
128. Jodicke G, Zenklusen O, Weidenhaupt A, Hungerbuhler K (1999) Developing environmentally sound processes in the chemical industry: a case study on pharmaceutical intermediates. J Cleaner Prod 7(2):159–166
129. Jimenez-Gonzalez C, Curzons AD, Constable DJC, Cunningham VL (2004) Cradle-to-gate life cycle inventory and assessment of pharmaceutical compounds. Int Life Cycle Assess 9:114–121
130. Jimenez-Gonzalez C, Overcash MR (2000) Energy optimization during early drug development and the relationship with environmental burdens. J Chem Technol Biotechnol 75:983–990
131. Wall-Markowski CA, Kicherer A, Saling P (2004) Using eco-efficiency analysis to assess renewable-resource-based technologies. Environ Progress 23(4):329–333
132. Slater CS, Savelski MJ, Carole WA, Constable DJC (2010) Solvent use and waste issues. In: Dunn P, Wells A, Williams T (eds) Green chemistry in the pharmaceutical industry. Wiley-VCH, Weinheim, pp 49–82

133. Slater CS, Savelski MJ (2009) Towards a greener pharmaceutical manufacturing environment. Innov Pharm Tech 29:78–83
134. Raymond MJ, Slater CS, Savelski MJ (2010) LCA approach to the analysis of solvent waste issues in the pharmaceutical industry. Green Chem 12:1826–1834
135. Slater CS, Savelski MJ, Taylor S, Kiang S, LaPorte T, Spangler L (2007) Pervaporation as a green drying process for solvent recovery in pharmaceutical production. Paper 223f, American Institute of Chemical Engineers, 2007 annual meeting, Salt Lake City, UT, Nov 2007
136. Slater CS, Savelski MJ, Hounsell G, Pilipauskas D, Urbanski F (2008) Analysis of separation methods for isopropanol recovery in the celecoxib process, Paper 290b. In: Proceedings 2008 meeting, American Institute of Chemical Engineers, Philadelphia, PA, Nov 2008
137. Raluy RG, Serra L, Uche J, Valero A (2004) Life-cycle of desalination technologies integrated with energy production systems. Desalination 167:445–458
138. Pre Consultants, Amersfoort, The Netherlands http://www.pre-sustainability.com
139. Papasavva S, Kia S, Claya J, Gunther R (2001) Characterization of automotive paints: an environmental impact analysis. Prog Org Coat 43:193–206
140. Dobson ID (2001) Life cycle assessment for painting processes, putting the VOC issue in perspective. Prog Org Coat 27:55–58
141. Lopes E, Dias A, Arroja L, Capela I, Pereira F (2003) Application of life cycle assessment to the Portuguese pulp and paper industry. J Cleaner Prod 11:51–59
142. Rios P, Stuart JA, Grant E (2003) Plastics disassembly versus bulk recycling: engineering design for end-of-life electronics resource recovery. Environ Sci Technol 37:5463–5470
143. Song H-S, Hyun JC (1999) A study on the comparison of the various waste management scenarios for pet bottles using life-cycle assessment (LCA) methodology. Resour Conserv Recycl 27:267–284
144. Ekvall T (1999) Key methodological issues for life cycle inventory analysis of paper recycling. J Cleaner Prod 7:281–294
145. Shiojiri K, Yanagisawa Y, Fujii M, Kiyono F, Yamasaki A (2004) A life cycle impact assessment study on sulfur hexafluoride as a gas insulator. In: AIChE sustainability engineering conference proceedings, Austin, TX, Nov 2004, pp 135–143
146. Vlasopoulos N, Memom FA, Butler D, Murphy R (2006) Life cycle assessment of wastewater treatment technologies treating petroleum process waters. Sci Total Environ 367:58–70
147. Sauers L, Mitra S (2009) Sustainability innovation in the consumer products industry. Chem Eng Progress 105:36–40
148. Cederberg C, Mattson B (2000) Life cycle assessment of milk production—a comparison of conventional and organic farming. J Cleaner Prod 8:49–60
149. Zabaniotou A, Kassidi E (2003) Life cycle assessment applied to egg packaging made from polystyrene and recycled paper. J Cleaner Prod 11:549–559
150. Bohlmann GM (2004) Biodegradable packaging life-cycle assessment. Environ Progress 23(4):342–346
151. Anderson K, Ohlsson T, Olsson P (1998) Screening life cycle assessment of tomato ketchup: a case study. J Cleaner Prod 6:277–288

Additional Suggested Reading: Introduction to Green Chemistry and Green Engineering

Allen D, Rosselot K (1997) Pollution prevention for chemical processes. Wiley, New York

Allen DT, Shonnard DR (2001) Green engineering: environmentally conscious design of chemical processes and products. AIChE J 47 (9):1906–1910

Anastas PA, Zimmerman JB (2003) Design through the twelve principles of green engineering. Environ Sci Technol 37(5):94A–101A

Boethling R, Mackay D (2000) Handbook of property estimation methods for chemicals. Lewis Publishers, Roca Raton, FL

Byrd D, Cothern R (2000) Introduction to risk analysis. Government Institutes, Rockville, MD

Daugherty J (1998) Assessment of chemical exposures. Lewis Publishers, Roca Ratan, FL

El-Halwagi M (1997) Pollution prevention through process integration. Academic, San Aeogo, CA

EPA Exposure Assessment Web site. www.epa.gov/oppt/exposure

EPA Pollution Prevention Framework Web site. www.epa.gov/oppt/p2framework/

Hesketh RP, Slater CS, Savelski MJ, Hollar K, Farrell S (2004) A program to help in designing courses to integrate green engineering subjects. Int J Eng Educ 20(1):113–128

Graedel TE, Allenby BR (1995) Industrial ecology. Prentice-Hall, Englewood Cliffs, NJ

Martin A, Nguyen N (2003) Green engineering: defining the principles—results from the Sandestin conference. Environ Progress 22(4):233–236

Ritter S (2003) A green agenda for engineering. Chem Eng News 81 (29):30–32

Shonnard DR, Allen DT, Nguyen N, Austin SW, Hesketh R (2003) Green engineering education through a US EPA/academia collaboration. Environ Sci Technol 37(23):5453–5462

Slater CS, Hesketh RP (2004) Incorporating green engineering into a material and energy balance course. Chem Eng Educ 38(1):48–53

Socolow R, Andrews F, Berkhout F, Thomas V (1994) Industrial ecology and global change. Cambridge University Press, New York

Pollution Prevention Heuristics for Chemical Processes

EPA Green Chemistry Web site, Green chemistry expert system: analysis of existing processes, building new green processes, and design. www.epa.gov/greenchemistry/tools.htm

Freeman H (ed) (1994) Industrial pollution prevention handbook. McGraw Hill, New York

Allen DT, Shonnard DR (2002) Green engineering: environmentally conscious design of chemical processes. Prentice-Hall, Upper Saddle River, NJ

Dyer JA, Mulholland KL (1998) Prevent pollution via better reactor design and operation. Chem Eng Process 94(2):61–66

Wynn C (2001) Pervaporation comes of age. Chem Eng Process October:66–72

Understanding and Prediction of the Environmental Fate of Chemicals

Crosby DG, Wong AS (1977) Environmental degradation of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). Science 195:1337–1338

Hansch C, Leo A (1995) Exploring QSAR: fundamental and applications in chemistry and biology. American Chemical Society, Washington, DC

Jeffers PM, Wolfe NL (1998) Green plants: a terrestrial sink for atmospheric methyl bromide. Geophys Res Lett 25:43–46

Mackay D (1991) Multimedia environmental models: the fugacity approach. Lewis Publishers, Boca Raton, FL

Meylen WM, Howard PH (1995) Atom/fragment contribution method for estimating octanol-water partition coefficients. J Pharm Sci 84:83–92

Syracuse Research Corporation [SRC] (2005) Syracuse, NY. Internet address: http://www.syrres.com

Environmental Performance Assessment for Chemical Process Design

Cano-Ruiz JA, McRae GJ (1998) Environmentally conscious chemical process design. Annu Rev Energy Environ 23:499

Life-Cycle Assessment

Carnegie Mellon University. http://www.eiolca.net/index.html, developed by Green Design Initiative, Carnegie Mellon University. Accessed 21 March 2005

Bauman H, Tillman A-M (2004) The Hitch Hiker's guide to LCA: an orientation in life cycle assessment methodology and applications. Studentlitteratur, AB, Lund

Graedel TE (1998) Streamlined life-cycle assessment. Prentice-Hall, Englewood Cliffs, NJ

Industrial Catalysis: A Practical Guide

6

Robert J. Farrauto

The Importance of Catalysis

Every student of chemistry, material science, and chemical engineering should be schooled in catalysis and catalytic reactions. The reason is quite simple; most products produced in the chemical and petroleum industry utilize catalysts to enhance the rate of reaction and selectivity to desired products. Catalysts are also extensively used to minimize harmful byproduct pollutants in environmental applications. Enhanced reaction rates translate to higher production volumes at lower temperatures with smaller and less exotic materials of construction necessary. When a highly selective catalyst is used, large volumes of desired products are produced with virtually no undesirable byproducts. Gasoline, diesel, home heating oil, and aviation fuels owe their performance quality to catalytic processing used to upgrade crude oil.

Margarine, cakes, chocolate, salad oils, and other everyday edible products are produced from natural oils via catalytic hydrogenation. Polyethylene and polypropylene plastics, commonly used in packaging of foods, films, fibers, liquid containers, etc., require catalysts for cost-effective high volume production. Because of highly active and selective catalysts, polyester fibers used in clothing can be produced at reasonable prices for the mass market. Catalysts enhance the production of ammonia-based fertilizers that enrich the earth's nutrient deficient soils for efficient agriculture. Catalytically produced ethylene oxide is a precursor to antifreeze. Formaldehyde is produced catalytically and used as a preservative and as a component in some polymer resins.

It is good to keep in mind the importance of catalysts in protecting the environment. They are frequently installed in the exhaust ducts from chemical operations to convert volatile organic compounds (VOC) generated during manufacturing operations, into harmless products. Catalysts also provide the environmental benefit of clean air by abating pollutants generated from stationary and mobile combustion sources. In many locations around the industrialized world, coal- and gas-fired power plants have special catalysts installed in the ducts to eliminate pollutants dangerous to our health. Many gas-fired compressors that pump natural gas through millions of miles of pipelines are also equipped with exhaust catalysts to clean emissions at moderate conditions. Even fastfood restaurants are being equipped with catalysts to eliminate odors from the cooking process. The most widely used treatment of exhaust pollutants is that of the catalytic converter present in the exhaust manifold that cleans emissions from the internal combustion engines of gasoline- and diesel-fueled automobiles and trucks. As modern commercial passenger jets fly above 30,000 feet, there is a need to destroy the few ppm ozone that enters the airplane with make-up air to ensure passenger and crew comfort and safety. Radiators on select vehicles have a catalytic coating deposited on their surface that decomposes harmful groundlevel ozone as the vehicle is driven.

All of this gives the consumer the benefits of readily available high-quality products at reasonable prices. From food to clothing to medicines to clean energy, catalysts play a major role in products people use in everyday life.

The forthcoming description of catalysts and catalytic processes should only serve as a primer towards understanding the basic principles with some examples of applications in the field of petroleum processing, alternative fuels, chemical production, and environmental air purification. Table 6.1 gives a list of some of the many commercial catalytic applications.

How Does a Catalyst Work?

A catalyst increases the reaction rate or activity relative to an uncatalyzed process by providing a less energetic pathway for conversion of reactants to products. In this regard the

---

R.J. Farrauto (✉)
Department of Earth and Environmental Engineering, Columbia University, 418 S.W. Mudd, NY 10027, USA
e-mail: RF2182@Columbia.edu

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_6, © Springer Science+Business Media New York 2012

Table 6.1 Some commercial catalytic reactions

Reaction name	Example of chemical reaction	Major catalyst components	Commercial applications
Petroleum processing			
Cracking	$C_{16}H_{34}=C_8H_{18} + C_8H_{16}$	Faujasite Zeolite	Naphtha, heating oil
Hydrodesulfurization	$RS + H_2=H_2S + RH$	Co, Mo/Al_2O_3 (S)	Sulfur-free fuels
Naphtha reforming	$C_6H_{12}=C_6H_6 + 3\ H_2$	Pt, Re/Al_2O(Cl)	High-octane gasoline
Alkylation	$C_3H_6 + C_4H_{10}=C_7H_{16}$	Liquid acids	Gasoline
Hydrogenation of functional groups			
Double bonds	$H_2R=R'–R''H_3 + H_2=H_3R–R'HR''H_3$	Ni or Pd or Ru on carriers	Air stable compounds, edible/non-edible oils
Aldehydes	$RCHO + H_2=RCH_2OH$	Pd or Ru or Ni on carriers	Alcohols
Selective oxidations			
Nitrogen dioxide for nitric acid	$4NH_3 + 5O_2=4NO_2 + 6H_2O$	Pt, Rh gauze	Fertilizer, explosives
Sulfuric acid	$SO_2 + 1/2O_2=SO_3$	V_2O_5 on TiO_2	Dissolving minerals
Ethylene oxide	$CH_2=CH_2 + 1/2O_2=C_2H_4O$	Ag on alpha Al_3O_3	Antifreeze, polyester, fibers, bottles
Formaldehyde	$CH_3OH + 1/2O_2=CH_2O + H_2O$	Bi, Mo or Ag	Monomers, preservatives
Synthesis gas/H_2			
Synthesis gas	$CH_4 + H_2O=3\ H_2 + CO$	Ni/Al_2O_3	Production of chemicals
Water gas shift	$CO + H_2O=H_2 + CO_2$	Fe, Cr and Cu, Zn, Al	High purity H_2
Pollution control			
Automotive (TWC)	CO, HC + $O_2=CO_2 + H_2O$ CO + $NO=1/2N_2 + CO_2$	Pt, Rh or Pd on $Al_2\ O_2$ on a Monolith	Clean emissions
Stationary reduction of NO_x	$2NO_2 + 4NH_3 + O_2=3N_2 + 6H_2O$	V_2O_5/TiO_2 on a monolith and metal exhanged zeolite	Reduction of NO_x from power plants
Polymerization	$nCH_2=CH_2=(–CH_2–)_{n+1}$	Cr/SiO_2 or $TiCl_4$ + alkyl aluminum halide	Plastics

catalyst provides a chemical and energetic shortcut by lowering the energy barrier (i.e., activation energy) of reactants going to products. If no catalyst were present, higher temperature would be required to initiate the reaction. Higher temperatures often lead to undesirable byproducts and sometimes decomposition of one of the reactants. Therefore, by initiating the reaction at a lower temperature, the process is more controlled and the desired product can be produced. This is the most important advantage for catalytic processes that is exploited in many product applications.

The catalyst is not consumed in the process; it accelerates but does undergo various chemical changes during the process by interacting with the reactants and products. Mechanistically some or all of the reactants adsorb onto active sites of the catalyst where bonds are rapidly made or broken. For a heterogeneous solid catalyst processing a liquid and/or gas, the adsorption of reactants is called chemisorption that has the kinetics and reaction energies equivalent to a chemical reaction. Frequently chemisorbed species decompose to an intermediate that is rapidly converted to other intermediates or the final product. After the reaction is complete, the catalyst returns to its original state. In this regard there is no net change of the catalyst. Therefore, a very small amount of catalyst can process many molecules.

What Are the Catalytic Metals and Metal Oxides?

Most catalytic metals and metal oxides are derived from Group VIII of the periodic table. Of special importance are Fe, Co, Ni, Rh, Pd, and Pt, but also of importance are Cu and Ag in Group 1b, V in Group Vb, and Cr and Mo in Group V1b. Three of the precious metals Rh, Pd, and Pt are extensively used in many industries due to their extremely high activity and selectivity. They are rare in nature and very expensive, and thus spent catalysts are routinely recycled, purified, and reused. However, the so-called base metals Fe, Co, Ni, Cu, V, Cr, and Mn but especially Ni and Cu are used for specialty chemical applications. Base metal catalysts usually have modest activities, but are much less expensive and in certain cases more selective than the precious metals. Therefore, it is always desirable to search for less expensive base-metal catalysts whenever possible. This has been especially the case for replacing precious metal-containing automotive emission control catalysts, but because of lower activity and stability in the severe environment of an automobile exhaust they are only used as promoters.

More examples of the efficient use of catalytic metals and metal oxides will be given in the applications section of this brief review.

The Structure of Heterogeneous Catalysts

The process of chemisorption of reactants requires adsorption on the surface of the catalyst. Therefore to maximize the rate, the catalytic surface area should also be maximized. This is achieved by dispersing the catalytic species onto a high surface area inorganic carrier. An ideal dispersion of Ni on Al_2O_3 is shown in Fig. 6.1.

Ideally every Ni atom should be accessible to the reactants for maximum efficiency in the conversion process. Although this is possible when the catalyst is first prepared, the dynamics of the catalytic reactions lead to some agglomeration. Catalyst scientists, however, have developed procedures and stabilizers to minimize the extent of agglomeration and therefore dispersed catalysts can be classified as nanomaterials with sizes only slightly greater than 1 nm or 10 Å.

The carrier can be thought of as a sponge with pores from 1 to 100 nm (10–1,000 Å) in diameter. If one were to measure the internal surface area of just 20 g with an internal surface area of 200–300 m^2/g, it would be equivalent to about 1 football field. Carriers such as Al_2O_3, SiO_2, TiO_2, CeO_2, ZrO_2, C, and combinations of these materials are commonly used. All have different surface properties and are used in applications dependent on the requirement for acidity, inertness to solubility, interactions with reactants, affinity for catalytically active components, and resistance to components in the gas phase. High surface area Al_2O_3 is not well suited for combustion reactions in which SO_2/SO_3 are present due to the formation of $Al_2(SO_4)_3$. In such cases, high area TiO_2 and/or ZrO_2 are used because of their inertness. Carbons are mostly used for supporting precious metals in hydrogenation reactions. In addition to their chemical role, precious metal recovery is achieved simply by burning the carbon.

The most common carrier is gamma alumina (γ-Al_2O_3). It has an internal area of >200–300 m^2/g. Its surface is highly hydroxylated (i.e., Al-O^-H^+). The H^+ sites provide acidity required for many reactions and exchange sites for catalytic metal cations.

Zeolites are combinations of Al_2O_3 and SiO_2 that are crystalline structures with precisely defined pore structures in the molecular size range (0.4–1.5 nm or 4–15 Å). A related group of materials known as mesoporous silica–alumina has extended the range of pore sizes attainable in ordered SiO_2–Al_2O_3 supports to 4 nm (40 Å). They are commonly used in the chemical and petroleum industry due to their surface acidity and ability to exclude molecules larger than the pore diameter. For this reason, they are often referred to as molecular sieves. Their surfaces contain Al–OH groups with acidic and exchangeable H^+. In the application section some of these materials will be more thoroughly discussed.

Rate-Limiting Steps for a Supported Catalyst

Supporting a catalytic component introduces a physical size constraint dictated by the pore size of the carrier. Thus, a key consideration is the accessibility of the reactants to the active

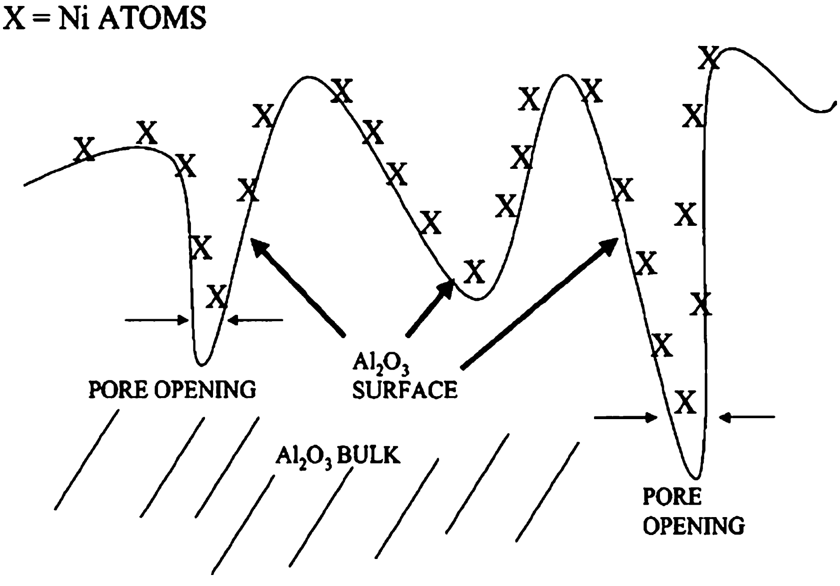


Fig. 6.1 Ideal dispersion of an active catalyst species on a porous carrier. Example given is Ni on Al_2O_3

catalytic sites within the high surface area carrier. Consider a hydrogenation reaction in which Ni is located in extremely small pores (i.e., 1 nm or 10 Å). The H_2 molecule has easy access, but a large molecule, having a size comparable to the diameter of the pore, would experience great resistance moving towards the active sites. If large amounts of Ni are present in pores and are not accessible to the molecules to be hydrogenated, the reaction rate will not be enhanced to its fullest potential. Thus, the carrier with its geometric sizes and its pore size distribution must be carefully designed to permit the reagents and products to move in the pores with minimum resistance.

Following are the seven fundamental steps in converting a reagent molecule(s) to its product(s) using a supported heterogeneous catalyst.

1. Bulk diffusion of all reactants to the outer surface of the catalyzed carrier from the external reaction media
2. Diffusion through the porous network to the active sites
3. Chemisorption onto the catalytic sites (or adjacent sites) by one or more of the reactants
4. Conversion and formation of the chemisorbed product
5. Desorption of the product from the active site
6. Diffusion of the products through the porous network to the outer surface of the catalyzed carrier
7. Bulk diffusion of the products to the external fluid

Steps 1, 2, 6, and 7 depend on the physical properties of the catalyzed carrier and are not activated processes (no intermediate chemical complex is formed). For this reason, we use the term apparent activation energy which is a term useful for comparing temperature dependence as will be described later. Steps 3–5 are chemically activated (with intermediate complexes formed during conversion to products) and depend on the chemical nature of the interaction of the reactants and products with the active sites [1–3]. Step 1 is referred to as bulk mass transfer. It describes the transfer of reactants from the bulk fluid to the surface of the catalyzed carrier. When this step is rate limiting, reactant molecules arriving at the external surface of the catalyst are converted instantaneously resulting in zero concentration of reactants at the surface. Thus, the internal surface of the catalyst is not used. Such as mass transfer controlled process is nonactivated and we assign an apparent activation energy of less than 2 kcal/mol. Rates vary only slightly with temperature ($T^{3/2}$) which, as will be shown below, allows it to be distinguished from other rate-limiting steps. Step 7 is similar to Step 1 except that the products diffuse from the external surface of the catalyst particle into the bulk fluid. The temperature dependence of this phenomena is relatively weak and has an apparent activation energy similar to that observed in Step 1 when it is rate limiting. When only the external surface of the catalyst particle is participating in the catalysis, it is said to have a low effectiveness factor. The effectiveness factor is defined as the actual rate divided by the maximum rate achievable when all catalytically active sites participate in the reaction. In the case of bulk mass transfer, the effectiveness factor approaches zero.

Steps 2 and 6 are both pore diffusion processes with apparent activation energies between 2 and 10 kcal/mol. This apparent activation energy is stated to be about 1/2 that of the chemical rate activation energy. The concentration of reactants decreases from the outer perimeter towards the center of the catalyst particle for Step 2. In this case some of the interior of the catalyst is being utilized but not fully. Therefore, the effectiveness factor is greater than zero but considerably less than one. These reactions are moderately influenced by temperature, but to a greater extent than bulk mass transfer.

Steps 3, chemisorption of the reactant(s), 4, chemical reaction forming the adsorbed product, and 5, desorption of the product(s) from the active site(s) are dependent on the chemical nature of the molecule(s) and the nature of their interaction with the active site(s). Activation energies are typically greater than 10 kcal/mol for kinetically or chemically controlled reactions. Chemical kinetic phenomena are controlling when all transport processes are fast relative to the reactions occurring at the surface of the active species so the effectiveness factor is one. All available sites are being utilized and the concentration of reactants and products is uniform throughout the particle. These reaction processes are affected by temperature more than either transport mechanisms. Figure 6.2 shows the conversion vs. temperature and concentration profiles of a reagent (R) for the three regimes of rate control.

Because of the significant differences in temperature dependence the kinetically limited reactions can be distinguished from pore diffusion that in turn can be differentiated from bulk mass transfer. This is shown in Fig. 6.2 in which conversion of reactants is measured against temperature. The first evidence of conversion is the sharply increasing slope that indicates kinetic control, whereas pore diffusion shows a lower change in slope as the temperature increases. The bulk mass transfer process shows little change in conversion with increasing temperature. Thus, at low temperature the reaction is controlled by chemical reactions (3, 4, or 5) and pore diffusion limited reactions exist when the supply of reactants to the active sites within the particle is limiting the rate (2 or 6). Finally, at the highest temperature, chemical reactions at the external surface are faster than bulk mass transfer (1 or 7) and the reaction is considered limited by bulk mass transfer.

The corresponding concentration of reactant R is also shown for each regime. The concentration of R is constant within the catalyst for kinetically limited reactions. The concentration of reactant gradually decreases within the catalyst particle for the pore diffusion limited case because the rate is limited by transport through the porous network.

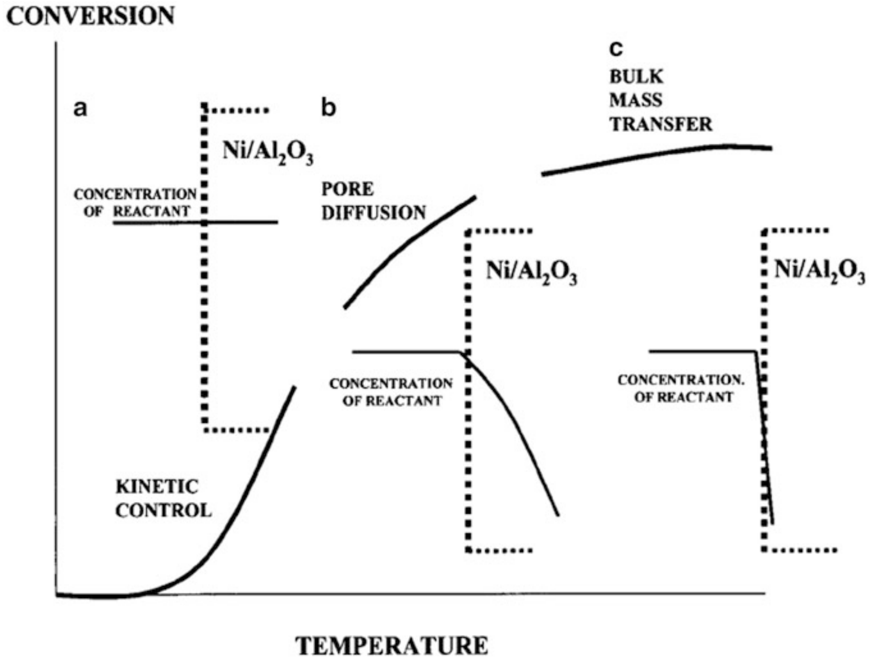


Fig. 6.2 Conversion of a reactant vs. temperature. The concentration of reactants [*R*] within the porous catalyst structure. Concentration of *R* is (**a**) uniform for kinetic control, (**b**) decreasing within the catalyst for pore diffusion control, and (**c**) zero immediately at the surface of the catalyst for bulk mass transfer

For bulk mass transfer limited cases, the concentration of R is zero at the gas/solid interface.

Activation Energies and Rate-Limiting Steps. The heterogeneous catalyzed NH_3 synthesis from N_2 and H_2 will be used to illustrate the relationship between rate and activation energy. There are a series of steps in the Fe catalyzed process.

The process steps within the pore structure of the catalyst are

1. Diffusion of N_2 and H_2 to the active Fe site within the catalyst pore structure
2. Chemisorption of H_2 on the active Fe surface
3. Dissociation of chemisorbed H_2 to H atoms on the Fe site
4. Chemisorption of N_2 on the Fe site
5. Dissociation of N_2 to N atoms on the Fe surface
6. Surface reaction between adsorbed N and H atoms forming chemisorbed NH_3
7. Desorption of NH_3 from the surface
8. Diffusion of the NH_3 into the bulk gas

Dissociation of chemisorbed N_2 (Step 5) is the slowest and thus is rate limiting.

$$N_2\text{—Fe(active site)} \rightarrow 2\text{N—Fe}$$

The overall rate of reaction is determined by the slowest of these steps. In other words, the overall reaction cannot be faster than the slowest step. The slow step and hence the overall reaction rate is characterized by the apparent activation energy.

An important detail is that an individual rate-limiting step may be endothermic, whereas the overall reaction is exothermic as in this case. This is illustrated in Fig. 6.3. The chemisorption of N_2 is exothermic and its dissociation is endothermic (1A). However, the overall reaction of N_2 + H_2 to NH_3 is exothermic (1B). The overall activation energy and kinetics are dictated by the slow step. The reaction heat liberated (ΔH_{25}°C) $= -11$ kcal/mole is the thermodynamic value associated with the overall reaction.

It is very important to understand that the catalyst only promotes the rate of a reaction and cannot change the equilibrium concentrations of reactants and products. It cannot make a thermodynamically unfavorable reaction occur. It increases the rate at which equilibrium is achieved while always respecting the thermodynamics of the equilibrium constant and the enthalpy (ΔH) and free energy (ΔG) of the overall reaction. Process conditions (T&P) are changed to give more favorable thermodynamics and rates.

Consider an everyday example of how we are all influenced by rate-limiting steps. If you are driving on a one-lane road behind a slow-moving truck, your speed and that of those behind you is no greater than that of the truck although you certainly have the potential to increase your rate. Thus, the time required to arrive at your destination is controlled by the speed of the truck. Taking an analogy, we can liken the truck to a slow chemical reaction step where the overall reaction rate, and the time required to achieve products, is no faster than the speed of the conversion at the surface of the catalyst. Returning to our highway story if you take a bypass road you can increase your rate of speed and

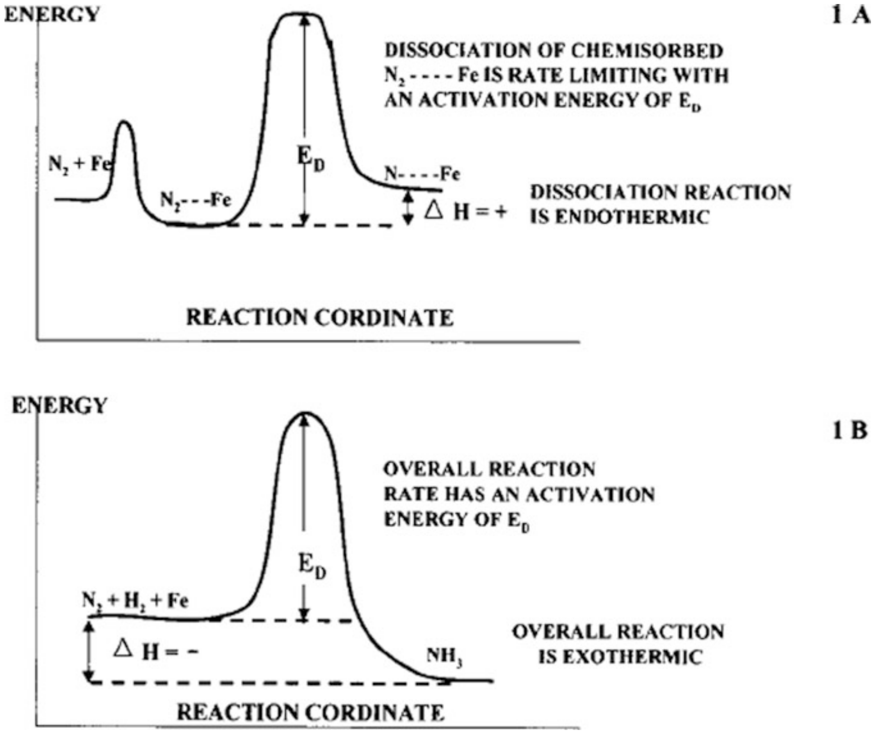


Fig. 6.3 Activation energy diagram for NH_3 synthesis: (1A) chemisorbed N_2 dissociation (rate-limiting step); (1B) overall reaction for NH_3 synthesis

decrease the time needed for you to arrive at your destination. The new road is analogous to a catalyst that provides a different pathway to the final product. It is likely, however, that you will again be limited by another obstacle (narrowing of the new road due to construction) that will slow you and the others behind you, as you maneuver through it. This may be likened to pore diffusion where you are limited by the width of the passage. Mass transfer control can be thought of as reaching the maximum speed your vehicle can safely achieve within the local speed limits. The activation energy reflects the slow step and the kinetics of the overall reaction rate.

Selectivity

In many processes, multiple reaction pathways are possible and it is the role of the catalyst and reaction conditions to maximize the rate along the desired path. Selectivity is defined as the amount of desired product divided by reactant converted. A practical example of the catalyst directing reactants to a selective product is shown by the oxidation of ammonia to nitric oxide, which is used in the production of fertilizers.

$$4NH_3 + 5O_2 \rightarrow 4NO + 6H_2O;$$
$$\Delta H_{25} = -54.2 \text{ kcal/mole } NH_3$$

The operating temperature of the process is 900°C and both the standard free energy of $\Delta G_{25} = -57.2$ kcal/mol of NH_3 and the equilibrium constant of $K_{NO} = 10^{10}$ are very favorable.

However, the decomposition pathway to N_2 is even more thermodynamically favorable with $\Delta G_{25} = -77.9$ kcal/mol of NH_3 and an equilibrium constant of $K_{N2} = 10^{15}$ at 900°C.

$$4NH_3 + 3O_2 \rightarrow 2N_2 + 6H_2O;$$
$$\Delta H_{25} = -75.5 \text{ kcal/mol of } NH_3$$

The presence of a PtRh gauze catalyst catalyzes the reactants along the NO pathway with a selectivity of 98%. Therefore, although the free energy is more favorable and the equilibrium constant for the N_2 reaction is 10^5 times greater, the highly selective PtRh catalyst promotes the NH_3 oxidation reaction to NO. In contrast, the presence of Pd favors the N_2 product. In each case the catalyst respects the equilibrium constant, but directs the reactants to specific products.

A second reaction that is currently receiving a great amount of attention because of low-temperature fuel cells is the purification of traces of CO present in a H_2 stream. The fuel cell directly converts chemical energy (H_2 and O_2) to electricity bypassing the mechanical (pistons, turbines, etc.) and combustion steps associated with conventional power generation. The mechanical step limits efficiency and combustion generates pollutants (CO, HC, and NO_x). The heat and power generated from the fuel cell hold promise for powering vehicles and for providing heat and electricity to residential and commercial buildings with the only product being H_2O. H_2 and CO are produced by catalytic steam reforming of a hydrocarbon (e.g., natural gas). The subsequent water gas shift reaction generates more H_2 from the $CO + H_2O$ reaction. Traces of CO exiting the shift reactor must be removed from the H_2 because it poisons the anode of the low-temperature fuel cell. The H_2 content of the gas is about 75%, and the CO is about 0.1% (balance is H_2O and CO_2). Although both standard state free energies are similar, a highly selective Pt containing catalyst promotes the oxidation of the CO with minimum oxidation of the H_2 purifying the latter for a low-temperature fuel cell.

$$CO + \frac{1}{2}O_2 \rightarrow CO_2 \ \Delta G_{25} = -61 \text{ kcal/mole}$$
$$\Delta H_{25} = -68 \text{ kcal/mole}$$
$$H_2 + \frac{1}{2}O_2 \rightarrow H_2O \ \Delta G_{25} = -57 \text{ kcal/mole}$$
$$\Delta H_{25} = -55 \text{ kcal/mole}$$

A small amount of air is injected into the reactor. The inlet H_2/CO ratio is 750, whereas the exit ratio must be 75,000. Thus, the free energy for CO oxidation is becoming less favorable (more positive) as CO is reduced below 10 ppm. An effective catalyst [4] currently in use commercially is Pt on Al_2O_3 promoted with a small amount of Fe. It operates at an inlet of 90°C and reduces the CO to less than 0.001% with a selectivity of well over 50% depending on management of the exothermic heat of reaction. This is quite remarkable given the increasingly large excess of H_2 as the reaction approaches completion. The same catalyst, but without the Fe, requires 170°C to achieve the same conversion of CO but with a selectivity less than 25%.

Catalyst Preparation

In the example given above, a small amount of Fe is added to a Pt on Al_2O_3 catalyst. The catalyst is prepared by a very unique procedure that must be strictly adhered to in order to achieve the desired results. The Pt and Fe must be in such close proximity that the CO chemisorbs on the Pt and the O_2 on the Fe after which they react to form CO_2 [4]. Simply reproducing the composition will not give acceptable performance. The specific details of catalyst preparation may be confidential and are most often covered by patents and trade secrets.

Some general guidelines for supported catalyst preparations are presented below; however, the reader should consult the many references and patents available on the subject [5]. Even by doing so the precise details used by industry to optimize the catalyst will often not be found.

Known amounts of salt(s) of catalytic metals are dissolved in aqueous solutions and impregnated into carrier materials. The wet mass is dried at 110°C and calcined in air at 300–500°C, releasing the decomposable salt components and depositing the metal oxide on the surface within the depths of the porous carrier. For many oxidation reactions, the catalyst is now ready for use, but for hydrogenation it is necessary to reduce the impregnated metal oxide or salt chemically. Usually this is accomplished by flowing H_2, under conditions consistent with the maximum temperature of use for the reaction of interest.

The carrier can be in the form of a powder used for slurry reactions or a particulate such as a sphere, cylinder, or tablet (typically a few mm in diameter) used in fixed bed reactors. The size and shape depend very much on what is anticipated to be the rate-limiting step. For example, for a reaction limited by pore diffusion, it is customary to use a smaller particle in the shape of a star, trilobe, or wagon wheel to decrease the diffusion path while increasing the external geometric surface area.

Mechanical strength and solubility under reaction conditions must be considered in the selection. Although it is often stated that the carrier is inert, there are many cases where this is not the case. Some carriers provide acid or basic sites that act as cocatalysts with the metal or metal oxides performing other functions. Petroleum reforming (discussed later) requires a hydrogenation function, provided by the metal, to dissociate H_2 and the carrier provides the acid site to enhance isomerization reactions.

Multi-channel ceramic monoliths (Fig. 6.4) are now the primary choice as support structures to carry the active catalytic species for cleaning emissions from various sources of pollution [6]. Figure 6.4 shows the shapes used for both automotive and stationary pollution abatement applications.

The largest application is the automotive catalytic converter that converts carbon monoxide (CO), hydrocarbons (HC), and nitric oxides (NO_x) to nontoxic compounds. The monolith structure offers high cell densities and thus geometric surface area upon which the catalyst is deposited permitting smaller reactor sizes, high mechanical strength, excellent thermal shock resistance, and low-pressure drop [6]. A powdered carrier, usually stabilized Al_2O_3, is impregnated with catalyst precursor salts. A slurry of the impregnated carrier is prepared and milled to some desirable particle size. The monolith is dipped into the slurry and "washcoated" onto the walls of all of the channel surfaces (see Fig. 6.4). Air is blown through the channels to remove excess slurry. It is then dried and calcined in air at about 500°C. The finished structure now contains the proper amount of catalyst uniformly deposited throughout the channel length. The washcoat thickness is greatest at the corners or fillets of the cell due to its sharp angle. The reactants flow through the channels and catalysis occurs on the washcoated walls. There are many other variations of

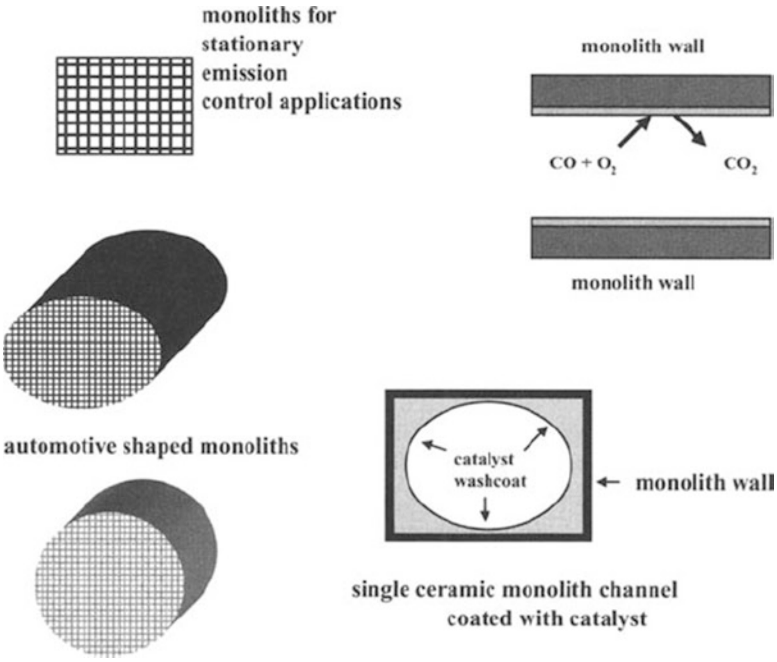
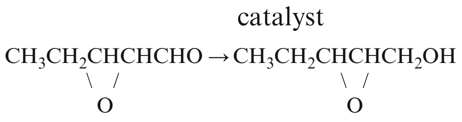


Fig. 6.4 Typical ceramic monolith geometries used for environmental emission control for vehicle and stationary applications

preparing monolith catalysts with different carriers and compositions. There are monoliths made of metal, some of which have parallel channels and others with nonparallel channels designed for tortuous flow to enhance mass transfer.

A Heterogeneous Catalytic Reaction: An Example

There is a great desire to use naturally occurring and renewable biomass for producing useful products. Furfural is extracted from cornhusks and contains an aldehyde functional group. If this group is selectively hydrogenated to the corresponding alcohol, the product furfuryl alcohol can be used to make corrosion-resistant resins for preparing polymers to make molds for shaping products. This reaction is selectively carried out with a Cu, Cr_2O_3 catalyst (copper chromite powder) in a slurry phase stirred batch reactor (see “Reactor Types” below) at 3,000 psig and 150°C.



$$\underset{\backslash\ /\ \ \mathrm{O}}{CH_3CH_2CHCHCHO} \xrightarrow{\text{catalyst}} \underset{\backslash\ /\ \ \mathrm{O}}{CH_3CH_2CHCHCH_2OH}$$

Hydrogen gas is dissociatively chemisorbed onto the surface of the Cu-containing catalyst producing highly active hydrogen atoms. The high pressure is needed to ensure adequate solubility of the H_2 in the furfural liquid. The aldehyde functional group forms a weak bond with these active adsorbed atoms and is hydrogenated to the finished product. In the absence of the catalyst, the diatomic hydrogen molecule would have to be dissociated in the gas phase at a much higher temperature leading to the decomposition of the aldehyde group.

Active Catalytic Sites

Not all catalytic sites are equal. Ideally each catalytic site is an atom having equal activity. This is never the case for a supported heterogeneous catalyst. One of the great mysteries in catalysis is the exact nature of the active site. Some catalytic species may be so well dispersed that they have no defined structure or are amorphous (no long-range structural order), whereas others may be highly crystalline. Amorphous catalytic components have greater catalytic surface area because fewer atoms are buried within a large crystal. However, the nature of the carrier and the catalytic species and the method used to deposit it on the carrier gives rise to a very heterogeneous surface with different sites having different surface energies and different activities. For example, it is believed that defects in the crystal structure produce highly energetic and active sites for catalytic reactions. This may be true but the more crystalline the catalytic site, the lower is the number of surface atoms and the lower is its catalytic surface area. All this being said, there are reactions that favor certain catalyst crystalline sizes and are said to be structure sensitive. The above discussion points to the mystery of catalysis. The goal of finding a universal model describing the nature of the active catalytic site still eludes us today and will undoubtedly be the subject of fundamental research for years to come.

Reactor Types

There are many different reactor designs, but the two most commonly used are fixed bed and batch slurry phase. For a fixed bed reactor, a given volume of solid particulate or monolith supported catalyst is fixed in a heated tube located within a furnace and liquid and/or gaseous reactants flow through the bed. This type of process is commonly used for large continuous-volume production where the reactor is dedicated to making only one product such as a bulk chemical or petroleum product.

Monolithic supports are commonly used for environmental applications and will be discussed in more detail later [6]. Batch reactors are used mostly for small-scale production such as the hydrogenation of intermediates in the production of medicines in the pharmaceutical industry. The catalyst powder is mixed in a precise amount of reactant in a pressurized-stirred autoclave. A gaseous reactant, usually H_2, is introduced at elevated pressures and the reaction proceeds with continuous monitoring of the H_2 consumed. The catalyst is separated from the product via filtration and is often used again depending on its retained activity and selectivity.

For the production of gasoline and other fuels by catalytic cracking of oils, a fluid bed reactor is used. This is a hybrid of a fixed bed and slurry phase reactor. The catalyst is fluidized as it interacts with the feed to be processed. This application is so important that it will be highlighted in the application section of this review.

Kinetics

The overall kinetics of a heterogeneous catalytic reaction can be controlled by any of the seven steps listed above [7–10]. We can distinguish which is rate controlling by determining the temperature dependence of the reaction. Once we know this we can design the catalyst to enhance the rate of the slowest step.

For example, bulk mass transfer (Steps 1 and 7) can be enhanced by increasing the turbulence of the reactants by increased stirring for a batch process or by increasing the linear velocity (see below) in the case of a fixed bed reactor. Increasing the geometric surface area of the catalyst also favors a reaction limited by bulk mass transfer. This is accomplished by decreasing the particle size of a particulate or by increasing the number of channels in a monolithic structure. Turbulence can be introduced in a monolith channel by modifying the surface from smooth to rough. Because kinetically controlled reactions have a stronger temperature dependence than transport controlled reactions, they are affected the most by increasing temperature. Pore diffusion resistance is decreased by increasing the pore size of the carrier or by using a smaller diameter carrier. One may also deposit the active catalytic species nearer the surface of the carrier to decrease the diffusion path length. The rate of a reaction limited by pore diffusion is moderately enhanced with temperature.

For chemically controlled reactions, one must modify the catalyst itself by increasing the number of active sites (increasing the catalytic surface area) or finding a crystal size that is more active for a given reaction. Often the activity is increased by the addition of promoters to the catalyst (i.e., Fe addition to Pt described under "Selectivity") that enhance the activity. Having the highest activation energy, it is affected more than the transport regimes by increasing the temperature. Many examples of this will be given in the example section of this chapter.

General Kinetic Rate Equations. The rate of a bi-molecular reaction is given by

$$\text{Rate} = d[A]/d(t) = -[k_F(A)^a(B)^b - k_{REV}(D)^d(E)^e]$$

Rate is the disappearance of reactants with time expressed as the derivative—$d[A]/d(t)$. $[A]$ and $[B]$ are the concentrations of reactants and $[C]$ and $[D]$ are the concentrations of the products. The exponents $a, b, c,$ and d are the reaction orders for each compound. The rate constants are k_F for the forward and k_{REV} for the reverse reaction. For those cases where the reaction is far away from equilibrium, the reverse rate is negligible and this term is dropped from the rate expression.

To determine the rate constant and the reaction order at a specific temperature, it is often convenient to increase the concentration of one reactant at least 20 times that of the other to maintain it relatively constant during the reaction. Thus, with a high concentration of reactant B one my write $k_F\ [B]^b = k_F{}^*$

$$d[A]/\mathrm{d}t = -k_F{}^*(A)^a$$

If the reaction order is to be determined, one may take the natural log of the rate equation and obtain

$$-\ln(d[A]/\mathrm{d}t) = \ln k_F{}^* + a\ln[A]$$

A plot of the $-\ln(d[A]/\mathrm{d}t)$ vs. $\ln[A]$ will produce a straight line with a slope equal to a and intercept $\ln k_F{}^*$

If one assumes $a = 1$ and integrates the rate expression

$$\int d([A]/[A]) = -k_F{}^* \int d(t)$$

Integration from the initial concentration A_o to A at anytime and from $t = 0$ to t

$$\ln([A_o]/[A]) = k_F{}^*t$$

$$[A] = [A_o - xA_o] \quad \text{where } x \text{ is the fraction converted}$$

Plotting $A_o \ln[1/(1-x)]$ vs. t will give a straight line with a slope equal to $k_F{}^*$ *Kinetics for Fixed Bed Continuous Reactions*. For continuous flow reactors, we use the term space velocity (*SV*) defined as the volumetric flow rate at *STP* divided by the volume of catalyst. That ratio yields the reciprocal of the residence or space-time (t)

$$\mathrm{SV}(\mathrm{h}^{-1}) = \frac{\text{flow rate cm}^3/\text{h(STP)}}{\text{catalyst volume}(\text{cm}^3)}$$

$$\frac{1}{\mathrm{SV}} = \text{residence time or space-time}$$

Thus, the rate equation for continuous flow reactions is

$$\ln([A_o]/[A]) = k_F{}^*t = k_F{}^*/\mathrm{SV}$$

The linear velocity (*LV*) or superficial velocity is an important engineering term because it relates to pressure drop and turbulence. This parameter is often increased in fixed bed reactors to enhance bulk mass transfer and heat transfer.

$$\mathrm{LV(cm/h)} = \frac{\text{flow rate cm}^3/\text{h(STP)}}{\text{frontal area of catalyst}(\text{cm}^2)}$$

Kinetics of a Slurry Phase Reaction in a Batch Process. This example is for the liquid phase hydrogenation of nitrobenzene to aniline with a powdered catalyst. These reactions typically are controlled by the supply of H_2 to the active sites.

$$3H_2 + C_6H_5NO_2 \rightarrow C_6H_5NH_2 + 2H_2O$$

H_2 must be

1. Transported from the bulk gas phase and dissolved in the liquid nitrobenzene.
2. Diffuse to the outside of the catalyst particle and into the pore structure.
3. H_2 and nitrobenzene react at the catalytic site.
4. Products diffuse through the pores and into the bulk liquid.

Steps 1, 2, and 4 are mass transfer phenomena while step 3 is kinetic.

At steady state the rate of mass transfer of reactants $(Rate)_M$ is equal to the kinetic rate $(\mathrm{Rate})_R$. This assumes Step 4 is fast and not rate limiting.

$$(\mathrm{Rate})_{\mathrm{net}} = (\mathrm{Rate})_{\mathrm{MT}} = (\mathrm{Rate})_{\mathrm{R}}$$

$$(\mathrm{Rate})_{\mathrm{MT}} = k_m(\mathrm{H}_{2g} - \mathrm{H}_{2s})/\mathrm{H}_{2g}$$

$$= 1 - (\mathrm{H}_{2s}/\mathrm{H}_{2g})$$

where H_{2g} = H_2 concentration in the gas

H_{2s} = H_2 concentration at catalyst surface

k_m = Mass transfer rate constant

$$(\mathrm{Rate})_{\mathrm{R}} = k_R(\mathrm{H}_{2s}Q)/H_{2g}$$

where k_R = kinetic rate constant

Q = the amount of catalyst

Equating $(\mathrm{Rate})_{\mathrm{MT}}$ and $(\mathrm{Rate})_{\mathrm{R}}$ and rearranging one obtains

$$(\mathrm{Rate})_{\mathrm{net}} = \frac{k_R k_m Q}{k_R Q + k_m}$$

Taking the inverse for the general rate equation and dividing both sides by $k_R\ k_m Q$, one obtains

$$(\mathrm{Rate})_{\mathrm{net}}{}^{-1} = k_M{}^{-1} + (k_R Q)^{-1}$$

A plot of inverse $(\mathrm{Rate})_{\mathrm{net}}$ vs. the inverse of Q yields a straight line with the slope equal to the inverse of k_R and the intercept the inverse of k_m. This is shown in Fig. 6.5.

When the amount of catalyst Q is large $k_R Q$ ≫ k_m

$$(\mathrm{Rate})_{\mathrm{net}} = k_m$$

The rate is limited by mass transfer because the reactants are consumed immediately at the outer surface of the catalyst. For small amounts of catalysts k_m ≫ $k_R Q$

$$(\mathrm{Rate})_{\mathrm{net}} = k_R Q$$

The reaction is kinetically controlled limited by the amount of catalyst.

Arrhenius Equation. The general rate constant (k) is an exponential function of temperature as described by the Arrhenius equation

$$k = k_o \exp^{(-E/RT)}$$

where

E = Activation energy for chemical control ("apparent" for diffusion limited processes)

R = Universal gas constant

T = Absolute temperature

k_o = Absolute rate constant

Taking the natural log of the equation gives

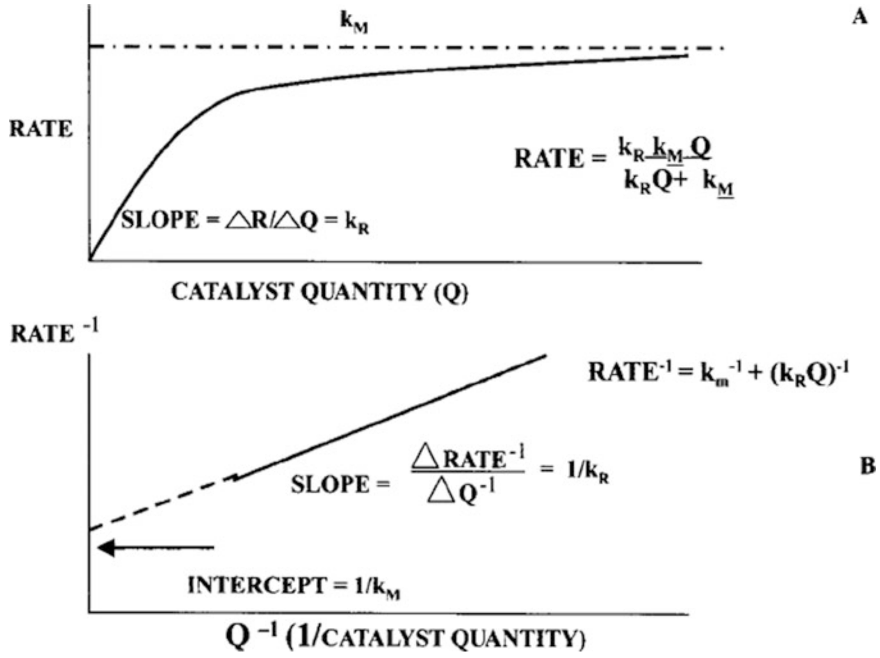


Fig. 6.5 Kinetics for slurry phase/batch reaction: (**a**) rate vs. quantity of catalyst; (**b**) line plot of the rate expression

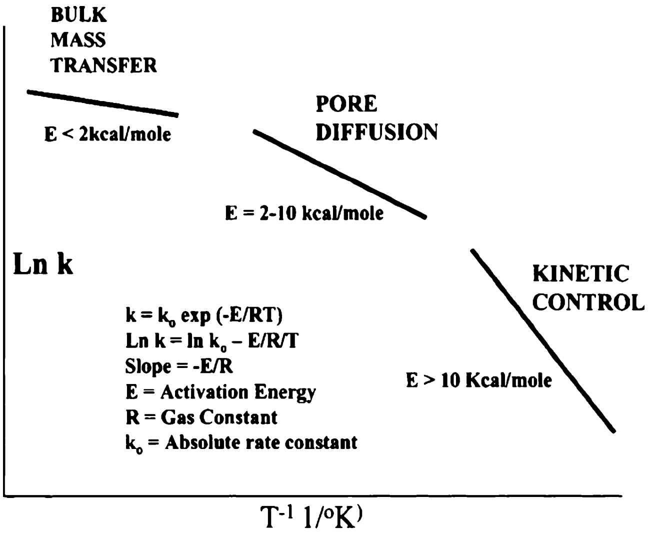


Fig. 6.6 Arrhenius profile for three regimes of rate control with activation energies (E)

$$\ln(k) = \ln(k_o) - E/RT$$

The plot of ln (k) vs. T^{-1} gives a straight line with a slope equal to $-E/R$ and intercept the absolute rate constant k_o as shown in Fig. 6.6. The lowest slope represents reactions controlled by bulk mass transfer, and the largest is for chemical or kinetic control. This method allows for the comparison of different rate-limiting steps, but it must be clearly understood that diffusion processes are not activated and thus we use the term apparent activation energy for them only to allow comparisons to activated processes such as chemically controlled processes.

Rate Models

The Langmuir-Hinshelwood kinetic model describes a reaction in which the rate-limiting step is reaction between two adsorbed species such as chemisorbed CO and O reacting to form CO_2 over a Pt catalyst. The Mars-van Krevelen model describes a mechanism in which the catalytic metal oxide is reduced by one of the reactants and rapidly reoxidizd by another reactant. The dehydrogenation of ethyl benzene to styrene over Fe_2O_3 is another example of this model. Ethyl benzene reduces the Fe^{+3} to Fe^{+2}, whereas the steam present reoxidizes it, completing the oxidation-reduction (redox) cycle. This mechanism is prevalent for many reducible base metal oxide catalysts. There are also mechanisms where the chemisorbed species reacts with a gas phase molecule and the combination rapidly converts to the final product. There are many kinetic models that describe different mechanisms and the reader is directed to some outside references [7–10].

Catalyst Deactivation

The first indication of catalyst deactivation is a significant change in the activity/selectivity of the process. Catalyst deactivation occurs in all processes, but it often can be controlled if its causes are understood. This subject is very extensive and the reader is encouraged to seek additional information in references given here [11, 12]. In the following we will present some of the most common deactivation modes especially for heterogeneous catalysts. These are pictorially shown in cartoon form in Fig. 6.7.

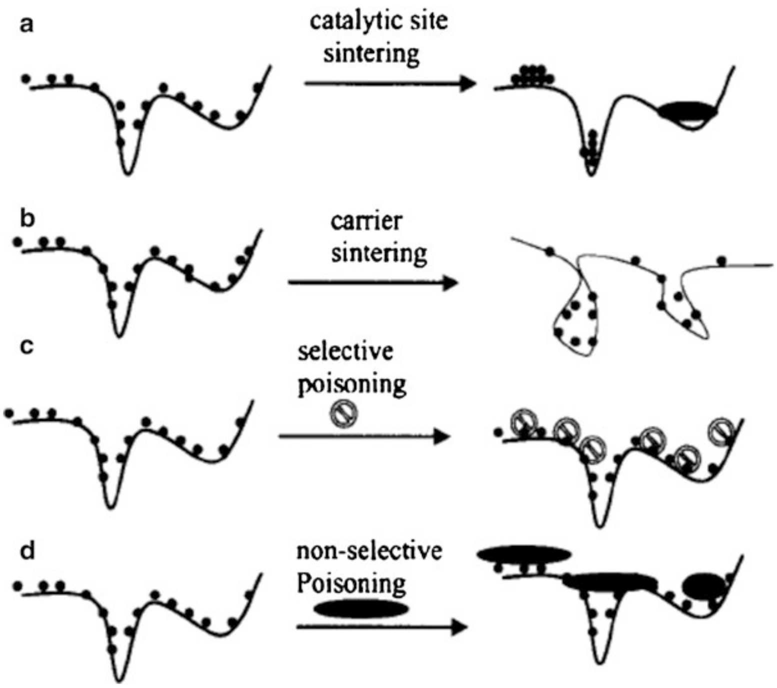


Fig. 6.7 Idealized poisoning mechanism: (**a**) sintering of the catalytic components, (**b**) sintering of the carrier, (**c**) selective poisoning, and (**d**) nonselective poisoning

Sintering of the Active Components. Catalytic scientists go to great lengths to disperse the active catalytic species over the surface of a carrier to maximize the number of sites available to the reactants. Small particles or crystallites have a high surface-to-volume ratio that is a highly unstable thermodynamic condition. The simple principle of Ostwald ripening indicates that small crystallites tend to grow to larger ones to bring the surface to volume condition to a favorable low free energy state. Thermal sintering occurs when small particles of active catalyst migrate over the surface of the carrier and agglomerate to form larger particles. There are other mechanisms of sintering, but conceptually this is the easiest to understand. The net effect is the loss of catalytic surface area that leads to loss of activity. The most frequently encountered cause is high temperature. This condition is encountered in Pt-, Pd-, and Rh-containing catalytic converters present in automobile exhausts where temperatures close to 1,000°C are commonly experienced. An oxidizing environment promotes the sintering of Pt by the formation of highly mobile or volatile Pt oxides. PdO, on the other hand, tends to form a stronger bond with the Al_2O_3 surface and thus sintering is not significant at modest temperatures. In contrast, it does sinter more readily in reducing environments. A catalytic species strongly bound to the surface is less likely to sinter. For this reason, a carrier such as SiO_2, which contains few OH groups on the surface relative to Al_2O_3, leads to sintering of the supported metal or metal oxide more readily. Catalyst companies have incorporated "rare earth stabilizers" into the formulations to minimize the rate of growth of the metal and metal oxide components. Stabilizers slow the rate of sintering, but do not completely prevent it due to the thermodynamic nature of the phenomenon. The goal is to minimize the rate to acceptable levels to ensure acceptable life of the catalyst.

Carrier Sintering. The purpose of the carrier is to provide a high surface area upon which the catalytic components can be dispersed. The high surface area leads to sintering by collapse of the pore structure that subsequently blocks (or occludes) the active sites by preventing access of the reactant. For some carriers such as Al_2O_3, there are changes to the crystal structure that occur as the temperature is increased. The most common is the conversion of high surface area (gamma) γ-Al_2O_3 (200 m^2/g) to low area (alpha) α-Al_2O_3 (1–5 m^2/g) at temperatures greater than about 800°C. This process occludes the catalytic components within the carrier and prevents the reactants from having access. The easiest analogy to understand is the truck that breaks down at the tunnel entrance; it prevents other vehicles from entering. High temperatures and steam are two of the most significant contributors to carrier sintering. Catalyst companies have incorporated metal oxides, such as Ba and La in precise percentages, into the carrier to minimize the sintering rate.

Poisoning. Specific components present in the reactant feed can adsorb selectively onto active catalytic sites rendering them inactive, in much the same way as CO can react with Fe-hemoglobin in the blood. For heterogeneous catalysts, sulfur compounds are the most universal poisons for both base metal catalysts and to a lesser extent precious metals.

Sulfur compounds present in petroleum, chemical, and environmental streams adsorb on the surface of Ni, Cu, Co, etc. forming metal sulfides that have little or no activity. In general, poisoning by sulfur compounds is irreversible. For this reason, upstream processes are used to reduce the sulfur to acceptable levels.

Sulfur oxides (SO_2 and SO_3) present in flue gases from upstream combustion operations adsorb onto the catalyst surface and in many cases form inactive metal sulfates. It is the presence of sulfur compounds in petroleum-based fuels that prevent the super-sensitive base metal catalysts (i.e., Cu, Ni, Co, etc.) from being used as the primary catalytic components for many environmental applications. Precious metals are inhibited by sulfur and lose some activity, but usually reach a lower but steady state activity. Furthermore, the precious metals are reversibly poisoned by sulfur compounds and can be regenerated simply by removing the poison from the gas stream. Heavy metals such as Pb, Hg, As, etc. alloy with precious metals and permanently deactivate them. Basic compounds such as NH_3 can deactivate an acidic catalyst such as a zeolite by adsorbing and neutralizing the acid sites.

Water is a reversible poison in that it will weakly adsorb (physically adsorb) on sites at low temperature, but readily desorbs as the temperature is increased.

One interesting example of different selective poisoning mechanisms is that of SO_3 deactivation of Pt on Al_2O_3 used for abating emissions from combustion reactions. The Pt oxidizes the SO_2 to SO_3 and the latter adsorbs onto the Al_2O_3 forming a sulfate. Slowly the carrier surface becomes so sulfated that it occludes the Pt within the pores and the catalyst slowly deactivates. By using a nonsulfating carrier such as TiO_2 or ZrO_2, deactivation can be prevented. In contrast, SO_3 directly adsorbs on Pd sites and deactivation occurs rapidly.

Poisoning is not always bad. There are situations where a catalyst is intentionally poisoned to decrease activity towards an undesirable reaction. In the hydrodesulfurization and -demetallization of a petroleum feedstock, the catalyst is presulfided prior to introducing the feed to decrease its activity and minimize cracking reactions that will produce unwanted gases. Another is the use of ammonia to slightly poison a Pt catalyst used in the hydrogenation of fats and oils to decrease undesirable oversaturation.

Nonselective poisoning or masking is caused by debris depositing on the surface of the catalyst physically blocking sites. Corrosion products from the reactor walls and contaminants such as dust, oil, etc. can be eliminated by careful filtration upstream, but this mechanism of deactivation is a constant problem in many applications. Regeneration is possible for precious metal oxidation catalysts designed to abate VOC from flue gases. The reactor is bypassed when the activity begins to decline to unacceptable levels. High-velocity air is passed through the catalyst bed and loosely held debris is dislodged. In some cases chelating solutions are used to solubilize the metal contaminants such as Fe without destroying the catalyst. Coking is a common phenomenon when petroleum and/or high molecular weight chemical compounds are processed. Hydrogen-deficient-hydrocarbons are formed from undesirable side reactions and block access to the catalytic sites deep within the pores of the catalyst. This deactivation mode has been positively integrated into the fluid catalytic cracking process for converting heavy oils to useful products. The coked catalyst is regenerated with air in a separate reactor and the heat liberated used to preheat the feed as it enters the cracker.

Catalyst Characterization

The goal of catalyst characterization is to relate the physical and/or chemical properties of the catalyst to performance. Some of the most important catalytic properties are physical surface area, pore size distribution, active catalytic surface area, the morphology or crystal structure of the carrier and active components, the location of the active components within the carrier, and the presence of surface contaminants or poisons on the surface. Fortunately, there are many instrumental tools readily available in modern laboratories to measure these properties for fresh and spent catalysts. There are many reference books and monographs available that describe the strengths and limitations of the instrumental methods used in characterizing catalysts [13, 14].

The chemical composition can be measured by traditional wet and instrumental methods of analysis. Physical surface area is measured using the N_2 adsorption method at liquid nitrogen temperature (BET method). Pore size is measured by Hg porosimetry for pores with diameters larger than about 3.0 nm (30 Å) or for smaller pores by N_2 adsorption/desorption. Active catalytic surface area is measured by selective chemisorption techniques or by X-ray diffraction (XRD) line broadening. The morphology of the carrier is viewed by electron microscopy or its crystal structure by XRD. The active component can also be measured by XRD, but there are certain limitations once its particle size is smaller than about 3.5 nm (35 Å). For small crystallites, transmission electron microscopy (TEM) is most often used. The location of active components or poisons within the catalyst is determined by electron microprobe. Surface contamination is observed directly by X-ray photoelectron spectroscopy (XPS).

Making the characterization measurements is of critical importance in the diagnostics of the catalysts, but interpreting those most responsible for changes in activity or selectivity requires experience and good comparative kinetics for fresh and aged materials. It should be standard practice to compare fresh and aged catalytic performance with the changes observed in your characterization diagnostics. Measuring rate-limiting steps and activation energies will provide invaluable insight into the major causes of deactivation.

Homogeneous Catalytic Reactions

In a homogenous catalytic reaction, the reactants and catalysts are in the same phase. The catalyst is a metal (Rh, Co, Ni, etc.) chelated with organic ligands (often phosphine-containing) soluble in the reaction media, and because no support is used, pore diffusion does not exist. However, bulk mass transfer is a concern especially when the reaction is a hydrogenation because H_2 must be dissolved in the liquid and make contact with the catalyst. This is accomplished by using high pressure and vigorous stirring. Homogeneous catalysis is most often used in the pharmaceutical industry where the desired selectivity can only be achieved with active complexes. A significant issue is separation of the catalyst from the final product to achieve the required purity. Furthermore, recovery of the catalyst is most often necessary especially for expensive precious metal containing complexes such as Rh. Distillation is sometimes used, provided there is a significant difference in vapor pressure of the product from the catalyst. The catalyst is also recovered by ion exchange with a suitable sequestering agent such as an amine compound. The efficiency of the separation allows for catalyst reuse and is essential for an economic process.

An example will be given in "Commercial Applications."

Commercial Applications

There are literally hundreds of commercial catalytic processes carried out for high and low volume premium products. Only a few have been selected below as examples of everyday products essential for a high-quality life. Table 6.1 also presents listings of some of the major catalytic processes, but the reader is directed to references given in this review for a more complete listing [10].

Petroleum Processing

Hydro-Demetallization (HDM) and -Desulfurization (HDS) of Heavy Oils. The hydrocarbon petroleum fractions separated by distillation contain varying amounts of inorganic impurities such as nickel, vanadium, and sulfur-containing compounds, all of which must be removed to make high-quality products both functionally and environmentally. The high-boiling fractions contain the highest concentration of metals and sulfur [15, 16]. Metals, if present in gasoline or diesel, will create significant engine wear and the sulfur would produce sulfur oxides during combustion and ultimately sulfuric acid in the atmosphere. Furthermore, they will deactivate the catalysts used in the petroleum upgrading processes and in their ultimate application as a fuel will damage the performance of the abatement catalyst.

Crude oil contains about 0.01% metals and up to 5% sulfur present in large aromatic structures. These levels are highly dependent on the origin of the crude. For example, California crude is relatively low in sulfur but higher in metals than crude from Kuwait. Any process to remove them must be economical with little destruction of the hydrocarbons and minimum consumption of H_2. The catalyst is Co, Mo/Al_2O_3 with particles a few mm in diameter. Although sulfur is usually a poison for catalytic reactions, it is used here in a positive function to control selectivity. It is presulfided to decrease activity towards excessive consumption of H_2 that leads to unwanted saturation of aromatic molecules.

$$\begin{aligned}
&\mathrm{SCo,\ Mo/Al_2O_3}\\
&\mathrm{R-M+H_2 \rightarrow M+R-H}\\
&\mathrm{R'S+H_2 \rightarrow R-M+H_2S}\\
&\mathrm{R\ and\ R'} = \text{organic host of metals and sulfur}\\
&\mathrm{M} = \text{metal (Ni or V)}
\end{aligned}$$

R and R′ = organic host of metals and sulfur M = metal (Ni or V)

The hydrogenation process is carried out at 500°C and pressures in excess of 30 atm in fixed bed reactors containing catalysts with varying physical properties to accommodate the metal deposition that occurs during the reaction. In some cases moving bed reactors are used where spent catalyst is continuously removed and fresh catalyst added.

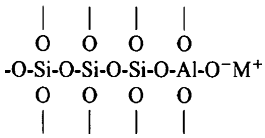
The first reactor contains the Co, Mo deposited on a low surface area Al_2O_3 with large pores to allow deep penetration of the metals into the particle. The second bed will treat a feed with less metal so its pore size is smaller and surface area slightly larger. The metal penetration here is less deep than in the first bed and allows for some hydrodesulfurization. The final bed contains the highest surface area and smallest pores and is designed to perform most of the desulfurization.

The catalyst is regenerated frequently during its useful life, but once spent it is leached and the metals recovered.

Catalytic Cracking for the Production of Useful Fuels. Gasoline and diesel fuel, home and commercial heating oil, kerosene, jet fuel, etc. are all produced by catalytically cracking fractions of distilled crude oil. Crude oil is distilled in large vertical towers where the various fractions present are separated according to their boiling ranges. The light gases (C_3 and C_4 propane and butane, respectively) are distilled first and the light/heavy naphtha fraction (C_5 to C_{10} pentanes to branched cyclopentanes), the precursors to gasoline, distill between roughly 70 and 200°C. Diesel fuel and heating oils (No. 1 and 2) are collected between 200 and 340°C. The remaining heavy hydrocarbons (called vacuum distillates) are used for lubricants and road construction.

The composition and molecular weight distribution of the crude oil depends on its origin, but generally less than 50% is within the molecular and boiling range for transportation and heating fuels. Thus, the role of the cracking process is to break or crack the higher molecular weight fractions into lower molecular weight compounds to be used for more useful products. Therefore, the catalyst is at the heart of the refining industry.

Cracking Catalysts. The catalysts used for cracking are called zeolites [17, 18]. They are SiO_2–Al_2O_3 materials in which Si, in its tetrahedral SiO_2 structure, is replaced with Al cations. They are produced by reacting sodium silicate with a water-soluble salt of Al followed by hydrothermal treatment in an autoclave. The zeolite is unique in that it has a well-defined crystal structure with a precise pore size (or aperture) ranging typically from 0.3–4 nm or 3–40 Å. This unique pore structure is responsible for separating molecules in accordance with their cross-sectional area. A molecule smaller than the aperture can enter the interior although a larger one cannot. Hence, the term molecular sieve is often used to describe zeolites. The composition and pore size can be varied giving rise to a large number of different zeolites with different pore sizes and crystal structures. They are usually identified by the Si/Al ratio, the crystal structure, and the size and shape of the pore. The Si^{+4} is bonded to 4 O^{-2} and each is bonded to another Si^{+4} establishing charge neutrality. Substituting Al^{+3} for Si cation upsets charge neutrality and requires another positive charge to satisfy the oxygen ions.



$$\begin{array}{c} | \quad\ \ | \quad\ \ | \quad\ \ | \\ \text{O} \quad \text{O} \quad \text{O} \quad \text{O} \\ \text{-O-Si-O-Si-O-Si-O-Al-O}^{-}\text{M}^{+} \\ \text{O} \quad \text{O} \quad \text{O} \quad \text{O} \\ | \quad\ \ | \quad\ \ | \quad\ \ | \end{array}$$

Neutrality is satisfied by a cation (e.g., M^+), which is usually Na^+ derived from the salts used in the synthesis. When the cation is exchanged with a proton, an acid site is created. This is the key active site for catalytic cracking reactions. The first exchange is with NH_4^+ which when heat-treated decomposes to NH_3 and the H^+ is retained on the zeolite. The acid zeolite is designated HZ.

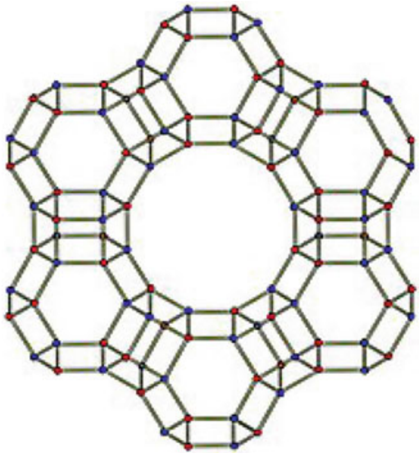
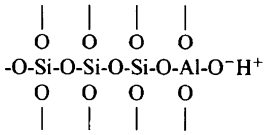


Fig. 6.8 Faujasite zeolite



$$\begin{array}{c} | \quad\ \ | \quad\ \ | \quad\ \ | \\ \text{O} \quad \text{O} \quad \text{O} \quad \text{O} \\ \text{-O-Si-O-Si-O-Si-O-Al-O}^{-}\text{H}^{+} \\ \text{O} \quad \text{O} \quad \text{O} \quad \text{O} \\ | \quad\ \ | \quad\ \ | \quad\ \ | \end{array}$$

The active zeolite for cracking reactions is called Faujasite and is classified as an X zeolite (HX). It has a Si/Al ratio of 1.0–1.5 with a pore or aperture size of 0.74 nm or 7.4 Å forming an aperture composed of 12 oxygen anions as shown in Fig. 6.8. The midpoint of each line represents an O^{-2} bonded to either Si^{+4} or Al^{+3}. It is the AlO^- site that requires a metal cation for charge balance. For cracking catalysts these sites are H^+. The higher the Al content (lower Si/Al), the greater the number of acid sites, but the lower the thermal stability. The zeolite is embedded within an amorphous SiO_2–Al_2O_3 structure that initiates the cracking of the large molecules, but also captures impurities such as organic compounds containing Ni and V that will severely deactivate the zeolite. Having its own acid sites, it also functions to break large molecules into smaller sizes where the zeolite can polish them to desired products. Catalyst particle sizes vary between 50 and 100 μm depending on the fluidization dynamics of the process.

During the fluidized catalytic cracking (FCC) process, a C–C bond is broken and a proton transferred from the catalyst to the molecule forming a positively charged carbocation. This ion can react with other hydrocarbons transferring its proton generating new carbo cations. Ultimately the large molecule is cracked to a smaller alkane and alkene with the regeneration of the protonated zeolite completing the catalytic cycle.

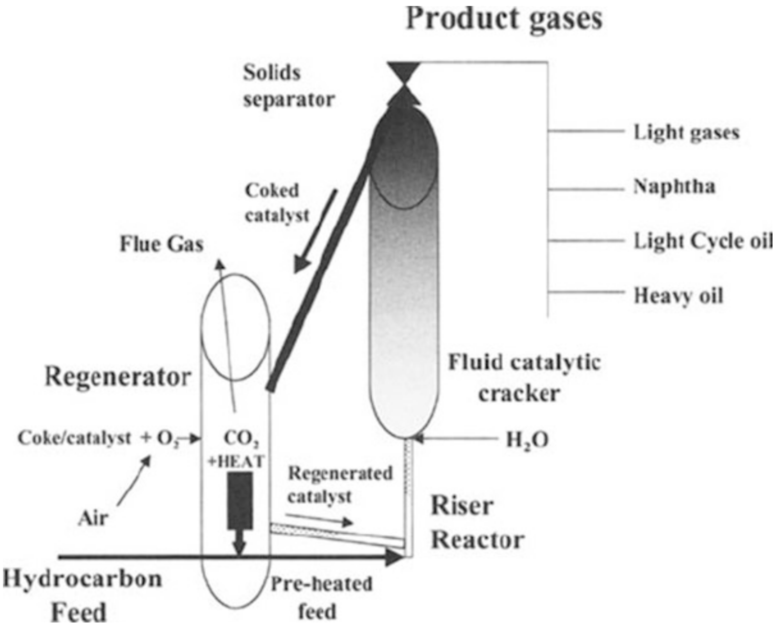


Fig. 6.9 Schematic of Fluid Catalytic Cracking (FCC) reactor with catalyst regenerator

Paraffin cracking: $C_{16}H_{34} \overset{HX}{\rightarrow} C_8H_{18} + C_8H_{16}$

Catalyst HX

Dealkylation: $C_6H_5 - CH_2CH_3 \overset{HX}{\rightarrow} C_6H_5H + CH_2 = CH_2$

Catalyst HX

Excessive extraction of H leads to the formation of hydrogen-deficient, high-boiling hydrocarbons called coke. Coking reactions are catalyzed by acid. The coke masks the surface and blocks the pore of the catalyst preventing access of the feed molecules leading to a loss in activity.

Cracking is carried out in a fluid bed process as shown in Fig. 6.9. Catalyst particles are mixed with feed and fluidized with steam upflow in a riser reactor where the reactions occur at around 500°C. The active life of the catalyst is only a few seconds because of deactivation caused by coke formation. The deactivated catalyst particles are separated from the product in a cyclone separator and injected into a separate reactor where they are regenerated with a limited amount of injected air. The regenerated catalyst is mixed with the incoming feed which is preheated by the heat of combustion of the coke.

Zeolites play a major role as catalysts and/or adsorbents in the petroleum, chemical, and lately in a growing number of environmental applications. The reader should consult references available [18].

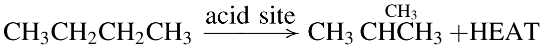
Naphtha Reforming for High-Octane Gasoline. Gasoline is volatilized and injected into the cylinders of the internal combustion engine where it is ignited under compression by a spark plug in the power stroke. Maximum power occurs when the cylinder reaches top dead center (maximum compression) and the mixture ignited by the spark plug. A high-octane gasoline is formulated not to preignite before reaching top dead center during compression to avoid the pinging or "knocking" sound that detracts from power. Before the mid-1970s, tetraethyl lead was added to quench preignition reactions, but because lead is no longer permitted the gasoline must be formulated to resist combustion until initiated by the spark. High-octane compounds such as aromatics and branched-paraffins are used in place of lead compounds. Today oxygenates are added to boost octane allowing decreases in carcinogenic aromatics.

Fuel-quality gasoline is made by a process called catalytic reforming [19, 20] in which molecules in the gasoline boiling range (called naphtha) are isomerized, dehydrogenated, and aromatized to make high-octane products. The most widely used reforming catalyst is Pt, Re on chlorinated Al_2O_3 particles (3–5 mm diameter). The Pt is the active component primarily for dehydrogenation and aromatization reactions and the Cl adds to the acidity of the carrier and is the active site for isomerization. The Re is believed to minimize coke formation. Dehydroisomerization requires both metal and acid functions. Some reactions are endothermic (dehydrogenation and dehydroisomerization) and others are exothermic (isomerization and dehydroaromatization). One can see below that the reactions lead to an increase in octane number

The dehydrogenation of cyclohexane to benzene and H_2 increases the octane number from 75 to 106.

$$C_6H_{12} + \text{HEAT} \xrightarrow{\text{Pt site}} C_6H_6 + 3H_2$$

Isomerization of *n*-butane to *i*-butane increases octane from 94 to 101.



$$CH_3CH_2CH_2CH_3 \xrightarrow{\text{acid site}} CH_3\,\overset{CH_3}{CH}CH_3 + HEAT$$

Heptane has a defined octane number of 0 and when dehydroaromatization occurs toluene is formed with an octane number of 116.

$$CH_3CH_2CH_2CH_2CH_2CH_2CH_3 \xrightarrow{\text{Pt + acid site}} C_6H_5CH_3 + 4H_2 + HEAT$$

The formation of benzene by the dehydroaromatization coupled with isomerization of methyl cyclopentane also increases octane from 76 to 106.

$$CH_3{-}C_5H_9 + HEAT \xrightarrow{\text{Pt + acid site}} C_6H_5H + 3H_2$$

The reforming process operates with three or four reactors in series. The feed is delivered to the first reactor at 500°C that is charged with the smallest amount of catalyst (5% of the total and the highest space velocity) to perform the easy but highly endothermic dehydrogenation reactions. To minimize coke formation, a small amount of H_2 is recycled from the product. The products and unreacted feed are then reheated to 500°C and fed to a second bed containing about 15% of the total catalyst charge where isomerization reactions occur. The unreacted feed and product are then reheated to 500°C where the more difficult dehydroisomerization reactions take place with 20% of the total catalyst charge. The final reactor contains 60% of the total catalyst charge and performs dehydrocyclization. Swing reactors are in place to allow the process to continue as each bed is being regenerated by coke burn-off. After regeneration, the catalyst must be rejuvenated by the addition of chloride. The final step is reduction of the metal to its active state.

Alternative Fuels

A secure energy supply and the need for reduction of greenhouse gas emissions will continue to be a technological goal for the world throughout the next few decades. It has become apparent that we need a fossil free energy supply. Solar, wind geothermal, tidal, etc. are natural sources of energy that are beginning to be utilized for stationary power generation. The promise of a hydrogen economy and fuel cell vehicles is also on the horizon. These technologies are slowly being introduced; however, we need a transitional approach for liquid fuels for vehicular applications. Biofuels, derived from plants, are already making an impact partially replacing fossil fuels in both gasoline and diesel applications [21]. In the US, ethanol is commonly added (10%) to gasoline, while bio diesel additions to fossil derived diesel can range from 2 to 100% depending on the locations and country of use.

Ethanol is derived primarily for corn, and other starchy plants, by the well known fermentation process where enzymes (nature's catalysts) accelerate the conversion of starch to ethanol. Bio diesel is synthesized by the homogeneous alkali catalyzed transesterification of triglyceride oils derived from plants such as soy and canola. Both alternative fuels are derived from edible plants which conflict with the food chain.

For this reason, there is a strong interest in utilizing non-edible portions of the plant, called lignin-cellulose as a source of fuel, especially ethanol. One of the main issues is the efficient penetration of the lignin (un-reactive aromatic polymer) portion of the plant which exists as a protective fiber surrounding the hemi cellulose and cellulose (sugar polymers of glucose), both of which are somewhat reactive towards enzymatic catalyzed fermentation to ethanol. This problem has stimulated the chemical and biochemical approach of developing new pretreatment techniques, including new enzymes, which can breakdown the lignin without destroying the ability to ultimately ferment the cellulose to ethanol. Alternatively, the thermal-chemical approach of gasification or pyrolysis of lignin-cellulose to produce gases and bio oils is being explored as a source of fuel. The gaseous and liquid products will require more traditional heterogeneous catalysts for upgrading to useable fuels.

Biodiesel can be derived also from non-edible plants such as jathropa and pennycress, both of which are rich in triglyceride oils relative to edible plants. These plants grow on arid lands during summer and winter seasons with little need for fertilizers. Research in harvesting and utilizing these plants is in progress with catalysts being used to upgrade the products to useful fuels.

Catalysts for Controlling Automotive Emissions

Oxidation Catalysts to Abate Unburned Hydrocarbon and CO Emissions

Catalytic converters were first installed in U.S. cars in 1976 [22, 23]. They were passive devices in that they were simply placed in the exhaust with no communication with the engine or its control strategy. It catalyzed the oxidation of the unburned hydrocarbons (C_yH_n) and carbon monoxide (CO) emitted during the incomplete combustion of the fuel. In some vehicles, excess air was pumped into the exhaust to ensure sufficient oxygen to complete the catalytic oxidation. This resulted in about a 90% reduction of these two pollutants relative to the uncontrolled uncatalyzed vehicle.

$$CO + \frac{1}{2}O_2 \rightarrow CO_2 + HEAT$$
$$C_yH_n + (1 + n/4)O_2 \rightarrow yCO_2 + n/2H_2O + HEAT$$

The presence of the catalyst provides a lower-energy chemical path than that offered by a thermal reaction. A catalyst accelerates oxidation of hydrocarbon/carbon monoxide/air mixtures that lie outside the flammability range required for thermal reactions. In the exhaust of the automobile, the composition of the pollutants is far below the flammability range, yet the oxidation reactions occur by the catalyst providing a lower-energy chemical path to that offered by the thermal reaction. An excellent example is the oxidation of CO with and without a catalyst. Without a catalyst, the rate-limiting step is O_2 dissociation at 700°C followed by reaction with gas phase CO. In the presence of the Pt catalyst, O_2 dissociation is rapid and the rate-limiting step becomes the surface reaction between adsorbed O atoms and CO that occurs below 100°C.

Two approaches were used in the design of the converters, both of which were positioned in the exhaust physically under the driver's seat. Both used precious metals (Pt and Pd) as the active catalytic components dispersed on Al_2O_3 (stabilized against carrier sintering with 1–2% CeO_2 and sometimes alkaline earth metal oxides).

One major automobile company used catalyzed Al_2O_3 beads (4 mm in diameter) and a spring-loaded pancakelike vessel to decrease the linear velocity and thus pressure drop. This design decreases back pressure that detracts from the power by offering less resistance to flow. Another used a new ceramic monolithic structure with hundreds of parallel channels (see Fig. 6.4). Upon the walls was deposited a coating of stabilized Al_2O_3 containing the active precious metals. The cordierite structure ($2MgO–5SiO_2–2Al_2O_3$) has a melting point over 1,300°C sufficiently high to withstand the expected temperatures in the exhaust. It was extruded and had excellent resistance to breakage due to thermal shock experienced during the transient operation of the vehicle. The cellular structure had between 200 and 400 cells per square inch (cpsi) parallel to the flow. With channel diameters of 0.059 in. (200 cpsi) and 0.044 in. (400 cpsi), they had open frontal areas of about 70%, offering little resistance to flow and thus low back pressure. It was incorporated into the exhaust system with retainer rings and surrounded by layers of insulation to minimize breakage due to vibration and heat. The regulations required that the converter have a life of 50,000 miles. To ensure this life, it was necessary to remove the tetraethyl lead, used to boost octane, in gasoline because the Pb poisoned the Pt and Pd by alloy formation.

Oxidation catalysts were used until 1979 in both the particulate (bead) form and monolith structure. Road experience demonstrated that the particulate beds were not mechanically stable and were breaking apart. In contrast, the washcoated monoliths were found to be highly reliable so they became the structure of choice.

Three-Way Catalytic Conversion

In 1980, additional regulations imposed by the U.S. Environmental Protection Agency (EPA) required control of NO_x (NO, NO_2, N_2O) emissions. Its removal coupled with the continuing need to remove CO and C_yH_n proved to be quite challenging because the latter had to be oxidized and the former reduced. Thus, it appeared two separate environments were needed. This problem was solved by the development of the three-way catalyst or TWC capable of catalyzing the conversion of all three pollutants simultaneously, provided the exhaust environment could be held within a narrow air-to-fuel range. This is shown in Fig. 6.10.

This range was defined between the fuellean and fuel-rich sides of the stoichiometric point, where the amount of O_2 is precisely sufficient for oxidizing both the CO and C_yH_n. This control required an O_2 sensor that is discussed below.

In the TWC, the Pt functions primarily as the catalyst for the oxidation reactions and the Rh catalyzes the NO_x reduction.

$$CO + NO(\text{or } NO_2) \rightarrow \frac{1}{2}N_2 + CO_2$$

$$H_2 + NO(\text{or } NO_2) \rightarrow \frac{1}{2}N_2 + H_2O$$

$$C_yH_n + (2 + n/2)NO(\text{or } NO_2) \\ \rightarrow (1 + n/4)N_2 + yCO_2 + n/2H_2O$$

The second reaction requires H_2 that is produced catalytically by the steam reforming reaction that occurs when excess C_yH_n is present.

$$C_yH_n + (n - 2)H_2O \rightarrow (n + 1)H_2 + yCO_2$$

O_2 Sensor. The control of the exhaust composition was essential to maintain the air-to-fuel ratio close to stoichiometric for simultaneous conversion of all three pollutants. This control came about with the invention of the O_2 sensor [23, 24]. The sensor head of this device was installed in the exhaust immediately at the inlet to the catalyst and was able to measure the O_2 content instantly and precisely.

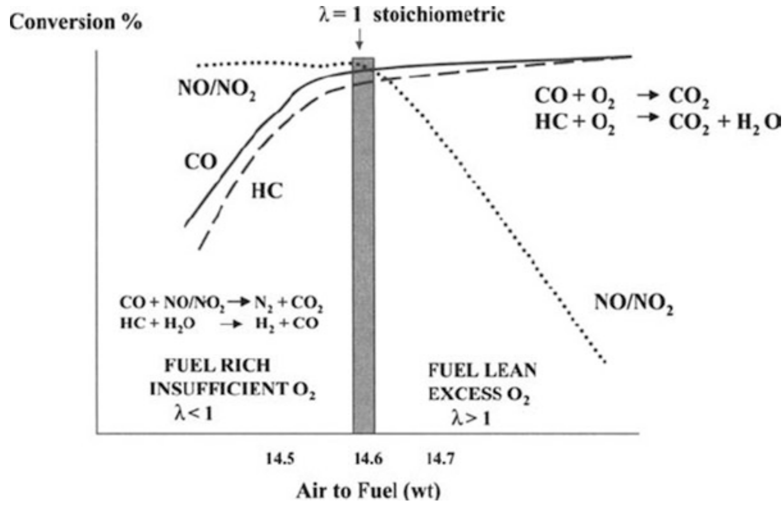


Fig. 6.10 Three-way catalytic (TWC) converter profile for conversion vs. air-to-fuel ratio

It generates a voltage consistent with the Nernst equation in which the partial pressure of O_2 $(PO_2)_{exhaust}$ in the exhaust develops a voltage (E) relative to a reference. The exhaust electrode was Pt deposited on a solid oxygen ion conductor of yttrium-stabilized zirconia (ZrO_2). The reference electrode, also Pt, was deposited on the opposite side of the electrolyte, but was physically mounted outside the exhaust and sensed the partial pressure $(PO_2)_{ref}$ in the atmosphere. E_o is the standard state or thermodynamic voltage. R is the universal gas constant, T the absolute temperature, n the number of electrons transferred in the process, and F the Faraday constant.

$$E = E_o + RT/nF[\ln (PO_2)_{ref}/(PO_2)_{exhaust}]$$

The CO and C_yH_n catalytically react with the O_2 at the surface of the Pt electrocatalyst. When the O_2 content is below stoichiometric, the electrode surface is depleted of O_2 causing an increase in the $(PO_2)_{ref}/(PO_2)_{exhaust}$ generating a large voltage. When the O_2 is higher than stoichiometric, the voltage is decreased. Thus, the electrodes must also function as catalysts. The voltage signal generated continuously fluctuates as the O_2 content is adjusted from sub to excess stoichiometric. Naturally, the exhaust electrode had to be resistant to exhaust poisons and temperature variations, so it was engineered with great care.

The voltage signal is fed back to the air/fuel intake system of the engine through an electronic control unit that controls the ratio necessary to maintain the proper window in the exhaust. Given the finite time necessary for the feedback system to function, it creates a perturbation of the O_2 content in the exhaust. The TWC catalyst had to be engineered to respond to these changes. The catalyst was composed of Pt, Rh on stabilized Al_2O_3 on a ceramic monolith, but an oxygen storage component (OSC) capable of storing and releasing O_2 was added. When the engine momentarily delivers less O_2 than the stoichiometric amount, the hydrocarbons present reduce the OSC. During higher O_2 spikes, the excess is stored on the OSC according to the fuel lean reaction below.

$$\text{Fuel rich:}\quad 2CeO_2 \rightarrow Ce_2O_3 + \frac{1}{2}O_2$$

$$\text{Fuel lean:}\quad Ce_2O_3 + \frac{1}{2}O_2 \rightarrow 2CeO_2$$

The current OSC material is CeO_2–ZrO_2 (proprietary promoters are added to stabilize it against sintering) where the oxidation state of the cerium is sufficiently labile to respond to the requirements for the OSC. The ZrO_2 is added to enhance thermal stability of the OSC.

U.S. federal regulations require that the driver be alerted to a malfunctioning catalyst through a signal on the dashboard. Currently, there is no instrumentation commercially available to sense the effectiveness of the catalyst to meet the onboard diagnostic requirement. An indirect solution is to place a second O_2 sensor at the exit of the catalytic converter. If the OSC in the catalyst is working properly, its voltage signal would have virtually no fluctuations because the O_2 content would be always zero. If the OSC is not functioning properly, O_2 will break through at the exit and the sensor would undergo similar fluctuations as the inlet sensor. Comparing these two signals generates the diagnostic that informs the operator of a malfunction.

A modern converter with dual O_2 sensors (one at the inlet and one at the outlet) is shown as Fig. 6.11.

Modern Catalytic Converter Systems

Modern TWC-equipped vehicles are required to meet minimum emission standards for 150,000 miles [23, 24]. It should be understood that after this time period the catalyst

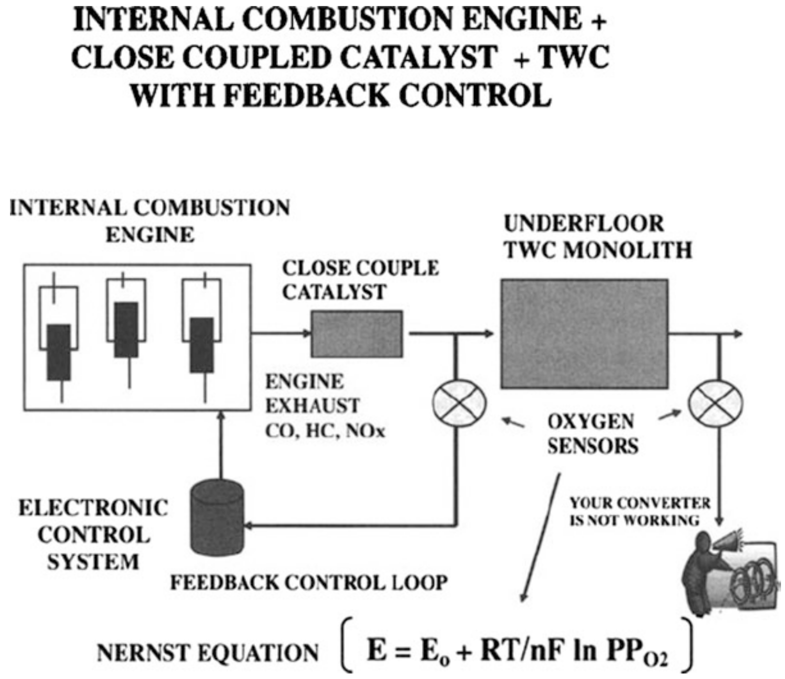


Fig. 6.11 Schematic of a exhaust system with a close couple and TWC catalyst. The system is equipped with dual O_2 sensors, one at the inlet and one at the outlet of the TWC converter

is still extremely active, but has lost sufficient activity that it no longer meets the stringent EPA standards. The source of deactivation is sintering of the catalytic metals, especially the OSC and the carrier due to the extremely high temperatures (900–1,000°C) experienced in the exhaust. The steam produced from combustion enhances the degree of sintering. There are proprietary stabilizers added to the formulations that minimize the extent of sintering. Poisoning effects by sulfur and oil components (Zn, P, Ca, S, etc.) have been minimized by reductions in fuel sulfur and careful design of the washcoat to prevent contact of the poisons with the catalytic components. These catalyst improvements, coupled with enhanced engine control, have resulted in lifetimes of at least 150,000 miles.

At start-up, the catalyst is cold and there is a substantial emission of hydrocarbons. It is during the first 1 or 2 min of cold-start operation that the vehicle can fail the federal test procedure. Here the reaction is kinetically controlled. Once it gets sufficiently warm, the reaction exotherm quickly raises the temperature and the reaction becomes limited by bulk mass transfer. The space velocity varies between 5,000 (idle) and 75,000 h^{-1} at high speed. So manufacturers had to design the catalyst for kinetic control and bulk mass transfer conditions. The cold-start issue was addressed by positioning a small oxidation catalyst (close coupled) up against the exhaust ports of the engine to ensure rapid heat up and light-off. This is shown in Fig. 6.11. The newest cordierite monoliths have lower weights for faster light-off and high geometric areas (600–900 cpsi) to ensure adequate bulk mass transfer area and lower pressure drop to meet modern driving demands and ever-increasing regulations.

It is truly remarkable that catalysts can function so well in the exhaust of the modern high-speed vehicle. This fact has raised confidence in industry to use different monolithic (ceramic and metal) structures as supports for catalysts for other environmental applications such as diesel exhausts, power and chemical plants, restaurants, and even on widebody aircraft.

Controlling Emissions from Diesel Engines

The diesel engine was invented by Rudolph Diesel in the latter part of the nineteenth century. Due to significant benefits in fuel economy, they have enjoyed a surge in popularity in recent years, particularly in Europe where they represent approximately 50% of new cars sold. Although improved fuel efficiency is the key driver, improvements in the drivability of diesel vehicles as well as a reduction in their tailpipe emissions have also helped improve their image and spurred further growth. Unlike gasoline emissions that are mainly gaseous in nature, diesel emissions from passenger cars, buses, and trucks contain solid, liquid, and gaseous components. Both diesel fuel and operation of the engine differ significantly from gasoline spark ignited engines and therefore their emission profiles are much more complicated. Diesel fuel has a boiling range from 200 to 350°C. The four-stroke engine compresses air and at maximum compression (top dead center) injects liquid fuel into the mixture where combustion occurs spontaneously driving the piston downward in the work stroke. Diesel engines operate with a large excess of air ($\lambda \gg 1$)

and therefore three-way catalysts, which operates at ($\lambda \sim 1$), will not catalyze the reduction of NO_x. Furthermore, the reduction of total particulate matter (TPM) must be addressed with new technology [23].

Based on the combustion characteristics of lean burn compression ignition diesel engines, the focus of emission regulations is on two key tailpipe pollutants—TPM and NO_x. Both of these are produced in high quantity during combustion. Since diesel fuel is less volatile than gasoline and is injected directly into the cylinder as a liquid, combustion initiates before the fuel has had sufficient time to vaporize and mix completely with the air. As a result, combustion begins at the gas–liquid interface of the fuel spray and then progresses towards the center. This "diffusion combustion" phenomenon results in the production of significant amounts of NO_x at the fuel liquid–gas interface where the local temperature and oxygen concentrations are high. In contrast, high amounts of soot (dry carbon) are generated in the interior of the spray where the temperature and oxygen concentrations are lower. As a result, both soot and NO_x are produced simultaneously in large amounts from diesel engines. This relationship between combustion temperature-particulates and NO_x is called the NO_x-particulate trade-off. When NO_x emissions are high (e.g., at high combustion temperatures), particulate emissions are low. In contrast, NO_x emissions are low (e.g., at lower combustion temperature), but the particulate emissions are high.

Also shown on this profile are the emission standards to be met for different countries and years. Clearly, regulations are requiring that both particulate and NO_x emissions be reduced to close to zero by 2010 and beyond in the US with other countries following closely.

Due to the adverse health effects associated with diesel soot and the ozone forming potential of NO_x, both are the major focus for emissions regulations. Although CO and HC are also regulated, the amounts produced by diesel engines are generally low enough not to be a major obstacle for meeting emission regulations. This is particularly true for heavy duty diesel applications.

The solids emitted from diesel engines are essentially dry soot (carbon rich particles). The liquids are primarily unburned diesel fuel and lubricating oils (commonly referred to as soluble organic fraction or SOF) and to some extent sulfates originating from the combustion of the sulfur compounds present in the diesel fuel. The combination of solid and liquid pollutants is referred to as particulates or TPM. Note that H_2SO_4 derived from the combustion of sulfur compounds in diesel fuel is included since it is a liquid at the collection conditions for TPM. The gaseous pollutants are CO, HC, and NO_x.

Total Particulate Matter
- Dry soot
- Liquids (oil, fuel) called SOF
- H_2SO_4

Gases
- CO, HC, NO_x

Diesel Oxidation Catalysts

The diesel oxidation catalyst (DOC) has two primary functions (1) to oxidize the CO and HC's emitted from the engine and (2) oxidize injected diesel fuel to generate heat for regenerating the particulate filter (to be discussed next). This catalyst is primarily a mixture of Pt and Pd on stabilized Al_2O_3 deposited on a ceramic monolith.

Regulations for dry soot reduction required an additional solution. This led to the introduction of the wall flow or diesel particulate filter (DPF). A DPF is primarily a cordierite honeycomb structure with alternating adjacent channels plugged at opposite ends. Exhaust enters the open channels, but only the gaseous components can pass through the porous wall exiting via the adjacent channel. Soot that is entrained in the exhaust stream is trapped on the wall while the gaseous components pass unrestricted. Periodically (e.g., every 1,000 km of driving), the filter is heated to a temperature high enough (ca. 550°C) to combust the soot and regenerate the filter. This heat is generally provided by oxidizing the diesel fuel injected into the DOC generating at least 550°C to initiate combustion of the soot. Alternatively, for vehicles without a DOC, fuel is injected into the cylinders during the exhaust stroke to promote combustion in the exhaust manifold, thereby raising the exhaust and DPF temperature. For the light duty market, DPF are usually silicon carbide or aluminum titanate.

A DPF may also contain a Pt containing catalyst (i.e., a catalyzed soot filter or CSF) to assist with the combustion of soot and to oxidize CO generated during the soot regeneration process.

Controlling NO_x in Diesel Engines

The US and European standards require the reduction of all three phases of diesel emissions (i.e., solids, liquids, and gases). In particular, reduction of NO_x will offer considerable challenges due to the lean nature (large excess air) of the exhaust. Two of the most promising technologies for controlling NO_x at the tailpipe are Selective Catalytic Reduction (SCR) and lean NO_x Traps (LNT). Both utilize catalytic processes to eliminate NO_x by chemically reducing it to N_2.

SCR relies on the reduction of NO_x by ammonia (NH_3) over either a vanadia (V_2O_5) supported on titania or metal-exchanged zeolite-based catalyst. The major desired reactions are below

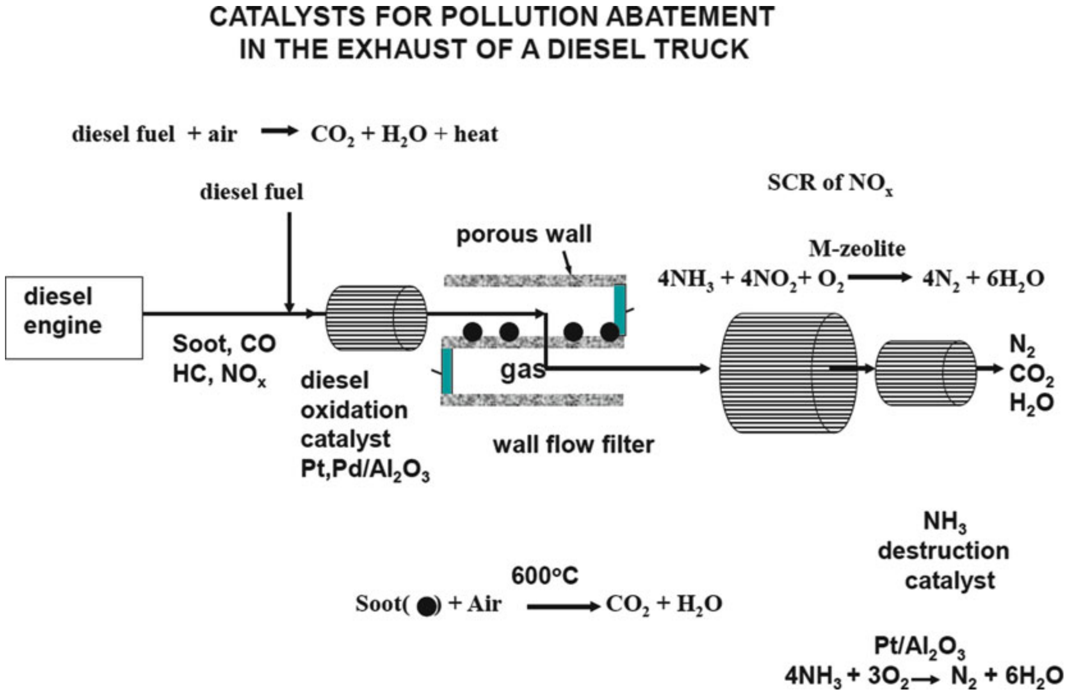


Fig. 6.12 Simplified diesel exhaust after-treatment system for heavy duty trucks. A diesel oxidation catalyst, wall flow filter, selective catalytic reduction, and ammonia decomposition catalyst

$$4NH_3 + 4NO + O_2 \xrightarrow{\text{catalyst}} 4N_2 + 6H_2O$$

$$4NH_3 + 2NO_2 + O_2 \xrightarrow{\text{catalyst}} 3N_2 + 6H_2O$$

Although commonly used in stationary source applications, SCR is relatively new for vehicle applications. Since handling gaseous ammonia is not practical in automobiles or trucks, an ammonia surrogate such as urea is utilized to generate ammonia in-situ in the vehicle exhaust. Typically, a solution of urea and water is injected into the exhaust stream before the SCR catalyst, and in the presence of water vapor, is hydrolyzed to ammonia which can participate in the SCR/NO_x reduction reactions. In addition, an ammonia "cleanup" catalyst (often referred to as an AMOX catalyst) may be required to remove any ammonia "slip" that may pass through the SCR catalyst unconverted. Clearly, integration of the SCR catalyst within the exhaust system is a challenge. Depending on the specific application, an oxidation catalyst (DOC), a CSF, an SCR catalyst, and an ammonia destruction catalyst (AMOX) may all be required to meet the combined CO, HC, and NOx regulations especially for heavy duty trucks. Sophisticated engine controls are required to ensure proper operation of all components within the system. The schematic in Fig. 6.12 shows the catalytic unit operations for meeting 2010 diesel emission standards using an SCR system.

The second promising technology for controlling NO_x at the tailpipe is LNT. The technology utilizes a Pt and Rh-based TWC catalyst in combination with an NO_2 trapping agent (e.g., an alkaline earth compound such as BaO). During the normal lean operation mode of the diesel engine, NO is oxidized to NO_2 over the Pt catalyst and the NO_2 is adsorbed by the BaO within the catalyst washcoat. Periodically (e.g., every 60–120 s), the trap is regenerated by introducing a "rich pulse" of reductant (e.g., diesel fuel) into the exhaust stream or by switching the engine operating mode to stoichiometric or rich for 1–2 s. This rich pulse provides the necessary chemical reductant to convert the adsorbed nitrate to nitrogen over the Rh catalyst.

Pt

$$\text{Lean mode:}\quad NO + \frac{1}{2}O_2 \rightarrow NO_2$$
$$NO_2 + BaO \rightarrow BaO\text{----}NO_2$$

Rh

$$\text{Rich mode:}\quad BaO\text{---}NO_2 + CO(H_2) \rightarrow N_2 + H_2O + CO_2$$

LNT technology has been successfully demonstrated in light duty vehicle applications; however, its primary disadvantages are high Pt levels are required to maintain sufficient catalyst durability and a fuel penalty (ca. 3–5%) results from the periodic trap regeneration. In addition, the

SO_x derived from the fuel-borne sulfur forms $BaSO_4$ that is much more stable than the corresponding nitrates and are not removed during the stoichiometric or rich operation mode. Therefore, the trap becomes progressively poisoned by sulfates. Complicated engine control strategies are being developed to desulfate the poisoned trap by operating the engine at a high temperature (>550°C) and rich of the stoichiometric air/fuel ratio for a short period of time. In addition, the air-to-fuel ratio must be carefully controlled to avoid the formation of H_2S during excessive rich conditions. LNT technology has the capability of removing up to 90% of the NO_x in the exhaust. Having lower sulfur fuels available will favor high NO_x conversion levels and also reduce the requirements for desulfation. Improvements in catalyst and engine control technologies hold promise for the future.

Catalytic Hydrogenation of Vegetable Oils for Edible Food Products

Triglycerides

Plant-derived oils such as soy, cottonseed, peanut, canola, corn, etc. are natural sources of edible products such as baking dough for cakes, cooking oils, salad dressing, chocolates, margarine, etc. Non-edible products such as lubricants, creams, lotions, etc. can also be produced depending on the processing of the oils. Natural oils are composed of long chains of fatty acid esters called triglycerides as shown in Fig. 6.13. The triglyceride chains are polyunsaturated, the degree of which influences their stability against oxidation in air.

Catalytic hydrogenation of the double bonds improves the stability against air and raises the melting point such that solids can be produced. Thus, the precursor to chocolate candy, margarine, or a cake mix is liquid oil that upon hydrogenation becomes an edible solid at room temperature. The more double bonds hydrogenated (the more saturated) the higher istowards air but often more injurious to our health by deposition of cholesterol in our blood the melting point, the lower is the reactivity vessels. The goal of a good catalyst coupled with the proper process conditions is to produce a reasonably healthy product with the desired melting point range with sufficient air stability to permit good shelf life.

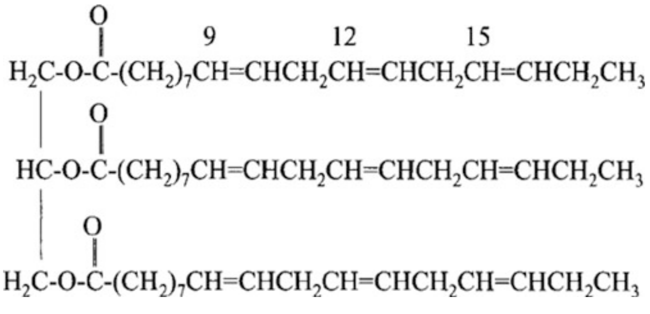


Fig. 6.13 Unsaturated fat molecule. Example shown is linolenic oil

Oils are classified by the length of the glyceride chain and degree of polyunsaturation. Typically nature produces oils with each chain length between 12 and 22 carbons with up to three unsaturated bonds usually all in the *cis*-form. Triglycerides with 18 carbons per length and three double bonds at positions 9, 12, and 15 counting from the first carbon in the ester group are called linolenic and designated C18:3. This structure is shown in Fig. 6.13. The outermost double bond is so reactive towards air that oils with three double bonds in the alkyl chain are rare. Therefore, the most prevalent in nature have double bonds at positions 9 and 12 and are referred to as linoleic (C18:2). Its reactivity is about half that of linolenic. The least reactive is oleic (1/20 that of linolenic) with only one double bond per length at position 9 (C18:1). Stearic is the term used for glycerides with all bonds saturated (C18:0). Not surprisingly, this form has virtually no reactivity towards air, has a high melting point, and is unhealthy.

The source of the oils plays a major role in producing a product with the desired melting point, stability, and health consequences. Cotton, sunflower, corn, and soy bean oils are a mixture of the four basic triglycerides with 50–70% C18:2 being the most dominant followed by 20–30% C18:1 with less than 1% C18:3. Less than 10% are other saturated oils such as C16:0. Palm kernel and coconut oils have almost 80% saturated triglycerides (C12:0, C14:0, and C16:0), have high melting points, and are stable against air but not healthy. They are used for protecting the skin against excessive sun exposure. Olive oil has up to 80% C18:1 and is therefore relatively healthy.

The most common oil hydrogenation catalysts are 20–25% Ni on Al_2O_3 and SiO_2. The nickel salts are either impregnated or co-gelled with a carrier precursor such as a soluble Al or Si salt. The catalytically active state of Ni is the reduced (metallic) form. The activation step is performed during manufacture at which time the catalyst is coated with a fatty gel to protect it from air oxidation during shipment.

Less than 1% by weight of catalyst is added to the batch reactor where the fatty protective gel slowly dissolves and the hydrogenation reaction commences. Temperatures of 100°C and H_2 pressures of 3–5 atmospheres are used to ensure adequate dissolution of the H_2 into the feed stream. Reactions are carried out between 100 and 160°C. Stirring is vigorous to maximize bulk gas diffusion to the catalyst surface. The catalyst particles are small to minimize pore diffusion resistance and increase liquid–solid mass transfer area.

Cu supported on Al_2O_3 is less active than Ni and this property is used to “brush” hydrogenate. Only minimum hydrogenation occurs maintaining the melting point but sufficient to improve stability against air. This is sometimes used for producing salad oils.

The reaction profile is generally sequential with hydrogenation first occurring on the most active double bonds followed by those less active [24–26]. Time distribution shows the linoleic form decreasing as the oleic form increases. After extended reaction time, the stearic begins to form as the oleic is slowly hydrogenated. Thus, one can design the process to control the product distribution in a predictable manner.

Hydrogenation of the linoleic form with a melting point of −13°C will produce a oleic product with a melting point of 5.5°C very suitable for consumption. During the partial hydrogenation process, it is most desirable to minimize isomerization to the *trans* isomer (the hydrocarbon groups are *trans* to each other across the remaining unsaturated bonds) because this structure raises the "bad" cholesterol or LDL (low-density lipids). A partially hydrogenated *cis*-structure may have a melting point of 6°C, whereas its *trans* isomer melts at 40°C. The *trans* isomer is more readily formed at high reaction temperatures, high Ni catalyst loadings, and at low hydrogen pressures (low concentration of H_2 at the catalyst surface). Pt containing oil hydrogenation catalysts produce considerably less *trans* than Ni; however, the high activity causes excessive hydrogenation of the double bonds. To minimize this effect, NH_3 is intentionally added to the feed or catalyst to poison its activity towards hydrogenation. By so doing, a low *trans* oil is produced without excessive saturation of the double bonds. In 2006, the U.S. Food and Drug Administration required labels that report the amount of *trans* components present in edible products.

Catalyst deactivation is mainly caused by mechanical attrition due to the rigorous stirring. In most cases, adsorption guard beds are used upstream to remove most of the impurities such as sulfur and phosphorous often found in the feed. Recognizing that some poisons may break through, the catalyst has an average pore size sufficiently large to admit the triglycerides but smaller than the average size of the organic compounds containing P and S. The spent catalyst is separated from the product by filtration. Given the increasing cost of Ni, it is recovered, refined, and used to make fresh catalyst.

Fertilizers and Hydrogen Generation

General Reactions

Ammonium nitrate (NH_4NO_3) and urea ($CO(NH_2)_2$) are two major sources of the world's fertilizers. The nitrate is produced by reaction of ammonia and nitric acid.

$$NH_3 + HNO_3 \rightarrow NH_4NO_3$$

Urea is produced by the reaction of NH_3 with CO_2 and its subsequent decomposition.

$$2NH_3 + CO_2 \rightarrow NH_2CO_2NH_4$$

Heart

$$\rightarrow CO(NH_2)_2 + H_2O$$

Ammonia is produced by the catalytic hydrogenation of N_2 with H_2

$$3H_2 + N_2 \rightarrow 2NH_3$$

Hydrogen is produced by a series of catalytic reactions, the first of which is hydrocarbon reforming, and the second is water gas shift. Considering natural gas (CH_4) as the starting hydrocarbon

$$\begin{gathered} CH_4 + H_2O \rightarrow 3H_2 + CO \\ CO + H_2O \rightarrow H_2 + CO_2 \end{gathered}$$

Nitric acid is produced by the selective catalytic oxidation of NH_3 and its subsequent hydration

$$\begin{gathered} 2NH_3 + \frac{7}{2}O_2 \rightarrow 2NO_2 + 3H_2O \\ \frac{3}{2}NO_2 + \frac{1}{2}H_2O \rightarrow \frac{3}{2}HNO_3 + \frac{1}{2}NO \end{gathered}$$

Each catalytic step will be discussed in this section.

Hydrogen Generation for the Production of NH_3

Producing H_2 from hydrocarbons such as natural gas is currently practiced in the chemical industry [27–30] under steady state conditions with carefully controlled catalytic unit operations. The overall process is as shown in Fig. 6.14.

Traces of organic sulfur compounds such as mercaptans, thiosulfides, and alkyl sulfides are added to natural gas to impart odor for safety detection of leaks. Because sulfur compounds are poisons to the downstream catalysts, they must be removed. The technology of choice is hydrodesulfurization or HDS.

$$RH - S + H_2 \rightarrow H_2S + RH_2$$

The reaction is carried out at about 300–400 psig and 300–400°C, but these conditions vary with the hydrocarbon feed. The catalyst is 3% Co, 15% Mo deposited on spheres

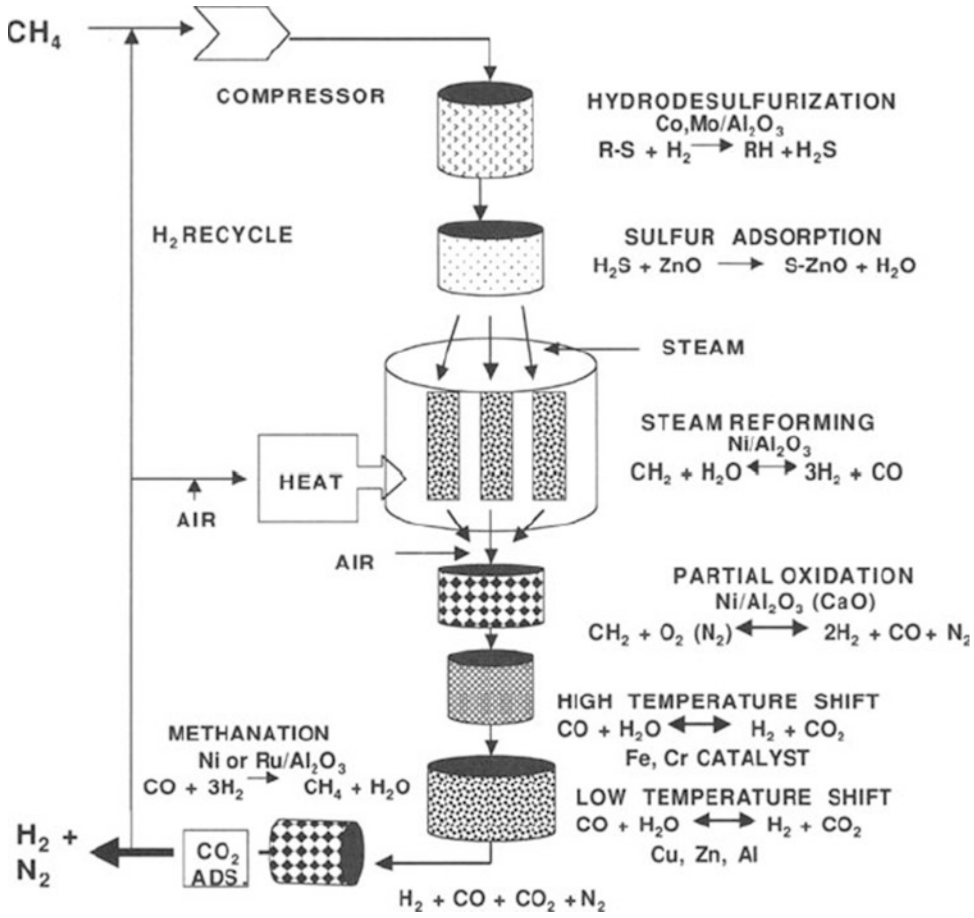


Fig. 6.14 The catalytic synthesis of hydrogen from natural gas for ammonia synthesis using hydrodesulfurization, steam reforming, partial oxidation, water gas shift, methanation, and CO_2 scrubbing

(25–75 m^2/g) Al_2O_3 with diameters of 2–3 mm. The catalyst is presulfided to decrease its activity towards undesirable side reactions such as coking.

The H_2S produced is removed downstream from the HDS reactor by adsorption on ZnO particles at about 400–500°C.

$$H_2S + ZnO \rightarrow ZnS + H_2O$$

Primary reforming of the sulfur-free natural gas (i.e., CH_4) is the first step to produce a H_2-rich gas.

$$CH_4 + H_2O \rightarrow CO + 3H_2$$

The reaction is highly endothermic and thus is favored at high temperatures. The maximum temperature achievable is limited by metallurgy of the reactor. Given the increase in gas volume, the reaction is favored by low pressures.

The catalyst is approximately 30% Ni with about 14% CaO on highly densified alpha alumina (α-Al_2O_3) with a surface area about 2–5 m^2/g. The CaO reacts with the Al_2O_3 forming $CaAl_2O_4$ for added mechanical strength under the severe operating conditions of 800°C and 300–400 psig and a steam environment up to 75%. The reaction rate is limited by a combination of heat transfer and pore diffusion, the latter due to the inadequate accessibility of the reactants to the catalyst interior. To counter pore diffusion limitations, the catalyst is manufactured as a donut with two to three holes to increase the external contact area and decrease the diffusion path. The space velocity is between 1,000 and 2,000 h^{-1}. The active catalyst is Ni metal, so it must be reduced carefully with H_2 prior to introducing the feed. The reaction is carried out in a series of tubular parallel reactors located in a large fired box furnace to provide the necessary heat.

Deactivation is due mostly to the slow accumulation of sulfur that breaks through the upstream HDS/ZnO guard beds. Sulfur irreversibly decreases the activity of the Ni that allows the methane decomposition rate to become significant leading to the accumulation of hydrogen-deficient carbon or "coke." This builds up within and between catalyst particles, leading to its fracture and an increase in pressure drop. During process shut down, the catalyst must be "passivated" to protect against air oxidation of the Ni and a subsequent fire due to its strong exotherm creating a safety hazard at the plant site. This is accomplished by periodically injecting small amounts of air and carefully oxidizing the surface of the Ni while monitoring the exotherm.

Equilibrium limits conversion of the CH_4, so partial oxidation or secondary reforming of the unreacted CH_4 is used to generate more heat and H_2 in a secondary reforming step. The addition of air also serves the purpose of providing the required N_2 for the subsequent ammonia synthesis reaction. Secondary steam reforming also uses a high-temperature resistant Ni containing catalyst that must retain its strength after prolonged exposure to close to 1,200°C due to the oxidation in the front end of the bed. The catalyst used is α-Al_2O_3 impregnated with about 18–20% Ni and 15% CaO.

$$2CH_4 + N_2 + 2O_2 \rightarrow CO + 3H_2 + CO_2 + H_2O + N_2$$

The exit from the secondary reformer contains about 10–12% CO, is cooled to about 350°C, and fed to a high-temperature water gas shift (HTS) reactor.

$$CO + H_2O \rightarrow H_2 + CO_2$$

The particulate catalyst is composed of 90% Fe and 10% Cr. The Cr minimizes sintering of the active Fe phase. The catalytic reaction is limited by pore diffusion, so small particles are used. The exit process gas contains about 2% CO as governed by the thermodynamics and kinetics of the reaction. This reaction is slightly exothermic and thermodynamics favor low temperatures that decrease the reaction rate. It is therefore necessary to further cool the mix to about 200°C where it is fed to a low-temperature shift reactor (LTS) containing another particulate catalyst composed of 30–35% Cu, 45% ZnO, and 13–20% Al_2O_3. The catalyst is active in the reduced form, so it must be carefully activated with H_2 avoiding excessive overheating which will cause sintering. The Zn and Al_2O_3 are added to stabilize the Cu because it is sensitive to sintering. The CO is decreased to its thermodynamic limit as imposed by the temperature and gas compositions. Typically, the CO is reduced to less than about 0.5%. The catalyst deactivates by traces of sulfur and sintering of the active Cu phase. Because of the necessity to operate at low temperatures, the reaction rate is slow and large volumes and low space velocities (1,500–2,500 h^{-1}) are used.

The active Cu-containing catalyst is also very air sensitive (like the Ni reforming catalyst) and will spontaneously oxidize generating uncontrolled reaction heats. Thus, it must be passivated before discharged and exposed to air. A small amount of air is added to the reactor and the temperature monitored. This process is continued until the exotherm is small enough that the catalyst can be safely removed from the reactor.

The remaining CO, which poisons the downstream ammonia synthesis catalyst, is removed by methanation using either a Ni or Ru on Al_2O_3 catalyst at 300°C

$$CO + 3H_2 \rightarrow CH_4 + H_2O$$

The CO_2 is scrubbed in an amine solution.

Ammonia Synthesis

The Haber process for the synthesis of ammonia from H_2 and N_2 has been practiced since the beginning of the twentieth century always with a massive Fe catalyst [31].

$$N_2 + 3H_2 \rightarrow 2NH_3$$

It is mildly exothermic, so the reaction is thermodynamically favored at lower temperatures but at high pressures due to the contraction of gaseous product volume. To obtain reasonable rates, the process is operated at about 450°C and pressures approaching 5,000 psig at a space velocity of 10,000–15,000 h^{-1}. The process is operated in a recycle mode, so ammonia is continuously removed aiding the equilibrium.

The active catalyst is 75–80% Fe metal, 10% Fe_2O_3, 4% Al_2O_3, less than 5% alkali and alkaline earth (Li, Ca, and Mg), with 1% SiO_2 added to minimize sintering of the Fe. The promoters are added to a melt of magnetite (Fe_2O_3). The solid mass is then ground to 1 mm particles, charged to the fixed bed reactor, and slowly reduced with H_2 at 500°C. The reduction generates active Fe metal with some porosity due to liberation of oxygen forming H_2O. The surface area is increased from about 1 m^2/g to about 20 m^2/g. The small particle size of the finished catalyst is necessary to minimize pore diffusion limitations. Special precaution is necessary during discharge from the reactor because air exposure will spontaneously oxidize the Fe surface generating large quantities of heat.

The catalyst is poisoned by CO, CO_2, and H_2O, so they must be rigorously removed upstream in the hydrogen synthesis process. Oxygen-containing molecules are permanent poisons. Other poisons such as sulfur, arsenic, halides, and phosphorous must be carefully removed upstream in as much as they too are permanent poisons.

Nitric Acid Synthesis

Nitric acid is produced by the selective oxidation of NH_3 over a gauze catalyst composed of 90%Pt, 10%Rh (some gauze is 90% Pt, 5% Rh, and 5% Pd) [32]. This reaction was used in the “Selectivity” section to demonstrate the high efficiency with which PtRh leads to NO production as opposed to more thermodynamically favored N_2.

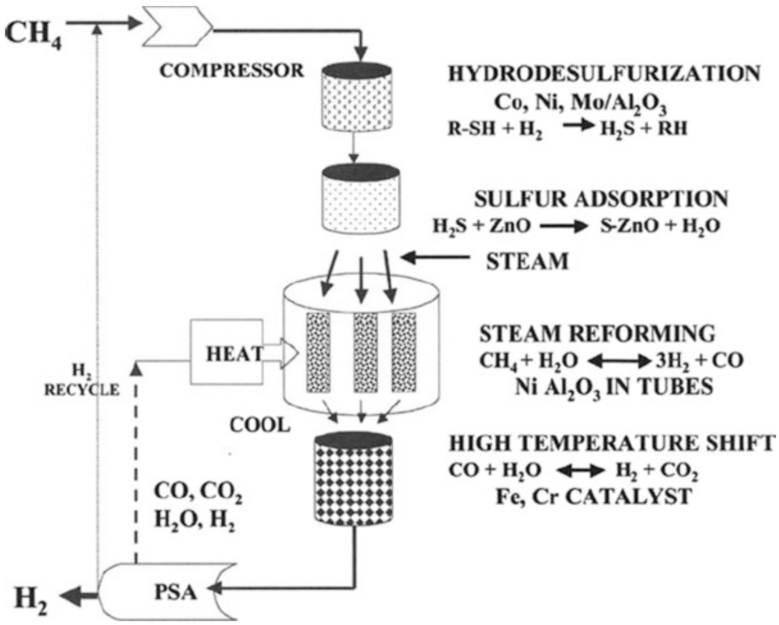


Fig. 6.15 The catalytic synthesis of hydrogen from natural gas using hydrodesulfurization, steam reforming, water gas shift, and pressure swing absorption (PSA)

$$4NH_3 + 5O_2 \xrightarrow{PtRh} 4NO + 6H_2O$$

$$4NH_3 + 3O_2 \rightarrow 2N_2 + 6H_2O$$

The low-pressure process (15–30 psig) produces NO with a selectivity of 98% and the high-pressure process (150 psig) has a selectivity of 94%. The high-pressure plant allows for a smaller reactor (and gauze) diameter (3 ft) compared to 12 ft for the low-pressure process.

The feed is composed of 10–12% NH_3 in air and is fed to the reactor at an inlet temperature of about 250°C and a space velocity approaching 50,000 h^{-1}. For a high-pressure plant, the exotherm generates an outlet temperature of 900°C. The NO product is cooled and noncatalytically converted to NO_2, its thermodynamically favored state at low temperatures, which reacts with H_2O forming HNO_3.

The finished alloy catalyst is manufactured by a knitting process to form a wire gauze that looks like a door screen. The Rh is added to impart mechanical strength during the wire drawing operations. During reaction, the catalyst undergoes an unusual morphology change. The Pt forms an oxy–nitrogen species and volatilizes from the gauze. The smooth wires become roughened and sprout resembling cauliflower. The surface area of the gauze increases by 20 times. The loss of Pt enriches the surface with Rh and the catalyst slowly loses activity. After approximately 90 days of operating a high-pressure plant, the Pt content of the gauze is reduced to 50%. The volatile Pt is captured downstream on a "getter gauze" made of Pd. The Pd surface catalytically decomposes the gaseous oxy–nitro Pt species and a Pt–Pd alloy forms that allows for easy recovery of the precious metals. The spent catalyst is returned to the supplier where the precious metal is recovered for future use.

Another source of deactivation is Fe contamination originating from the corrosion of upstream equipment depositing on the gauze resulting in decomposition of the NH_3 to N_2. Another source is from the Fe-containing ammonia synthesis catalyst.

Although the largest use for nitric acid is NH_4NO_3 fertilizers, it is also used for explosives and nylon polymers.

Pure Hydrogen Generation with Pressure Swing Adsorption Purification

For applications in which N_2 is not needed, such as H_2 or alcohol production, pressure swing adsorption (PSA) is used. The process flow diagram is shown in Fig. 6.15.

There is a renewed interest in hydrogen generation for the developing hydrogen economy with the anticipated use of fuel cells as a power source for vehicles [33]. The fuel cell generates electricity by electrochemically oxidizing H_2 and reducing O_2. Because it directly converts chemical to electrical energywithout using the traditional mechanical steps of piston-driven engines and turbines, it promises to be more efficient, cleaner, decrease our dependence on oil, and generate less greenhouse gas. The small-scale generation of H_2 for cost-effective refueling stations is a major issue that is aggressively being studied. Ultimately, it will be derived from water by electrolyzers using natural sources such as solar, wind, and geothermal energy. Until these technologies are available, natural gas reforming is a likely source because infrastructures exist in many cities in the world. The fuel cell powered vehicle will require an infrastructure similar to gasoline and diesel service stations. High-pressure H_2 will have to be available to refuel the vehicles.

Such demonstration stations are now being built in various parts of the world. PSA is used for final H_2 purification. The partial oxidation and the LTS are eliminated. There is therefore a loss of some hydrogen production, but the final H_2 is not diluted with N_2. The PSA purification unit replaces the methanator and CO_2 scrubber and produces pure H_2. Some of the H_2 is recycled for HDS and some combusted to provide the heat for steam reforming.

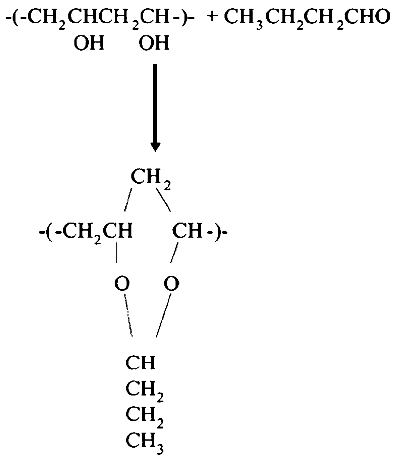
Given the need for smaller size reformers to be operated in local communities, safety will be an elevated concern. These reformers must contain nontoxic and air-insensitive catalysts, thus eliminating Ni, Cu, and Cr from consideration. A major research effort is underway to redesign the entire H_2 generation process system using modern materials such as monoliths and precious metal catalysts [33].

For fuel processors directly integrated to a residential fuel cell, H_2 must be maximized and no pressure is available for PSA. In these cases, a high- and LTS catalyst will be required. Hydrogen purification to reduce the CO to less than 10 ppm will be managed by preferential oxidation [34].

Production of Butyraldehyde: A Homogeneous Catalytic Reaction

Butyraldehyde

The incorporation of an inner layer of poly(vinyl butyral) or (PVB) in the glass of an automobile windshield protects against serious head injuries when a passenger strikes it during an accident. The strongly adherent coating is optically transparent and maintains the glass intact (anti-shattering agent) when a foreign object hits the surface. Thus, the glass does not shatter when a stone strikes its surface. PVB is produced by reaction of polyvinyl alcohol (PVA) with linear butylraldehyde ($CH_3CH_2CH_2CHO$).



Another important application of butyraldehyde is in the production of oxo-alcohols for use as plasticizers used to improve mixing of solid compounds that must be molded or extruded into specific shapes. The hydrogenation catalyst is Ni/Al_2O_3.

$$CH_3CH_2CH_2CHO + H_2 \xrightarrow{Ni/Al_2O_3} CH_3CH_2CH_2CH_2OH$$

Butylraldehyde is produced by a homogeneous catalytic process called hydroformulation in which CO and H_2 are added to liquid propylene using a soluble cobalt-containing complex catalyst $Co(CO)_6$. The reaction is carried out with a butanol solvent

$$\begin{aligned} &CH_3CH = CH_2 + H_2 + COCo(CO)_6 \\ &\quad \rightarrow 80\% CH_3CH_2CH_2CHO + 20\% CH_3CHCHO \\ &\quad + CH_3CHCHOCH_3 \end{aligned}$$

The conditions needed to catalyze the reaction are very severe; pressure = 3,000–4,500 psig and 150°C. The high pressure maintains the propylene in solution, ensures sufficient solubility of the H_2 and CO, and maintains the Co-carbonyl complex stable against decomposition. The product distribution is 4:1 linear to branch.

Much less severe conditions can be used with the Wilkinson homogeneous catalyst rhodium tricarbonyl triphenyl phosphate, $HRh(CO)_3(PC_6H_5)_3$. Pressures equal to 225 psig and temperatures of 100°C selectively produce the more useful linear form [35]. The milder conditions more than compensate for the more expensive Rh (1,000 times that of Co). The aldehyde product is distilled leaving the catalyst in the solvent ready for reuse.

Homogeneous catalysts are structurally well- defined complexes and, because they are soluble in the reaction mix, are not subject to pore diffusion limitations as are heterogeneous catalytic materials. They are almost always highly selective towards desired products. The main consideration is that the complex be stable and reactor conditions chosen such that all the gaseous reactants are adequately dissolved and mixed in the liquid phase. Homogeneous catalysts are easily characterized by standard instrumental methods for compound identification such as XRD or spectroscopy. Deactivation is associated with attack by traces of carboxylic acidic byproducts and impurities in the feed such as O_2 and chlorides that attack the ligand groups.

Polyethylene and Polypropylene for the Production of Plastics

Polyethylene

Specially prepared plastics are rapidly replacing traditional metal components because of their strength, transparency, resilence, lighter weight, and greater corrosion resistance.

The largest volume products are polyethylene and polypropylene. Each has its own contributions to the marketplace where the former is primarily used for low-strength applications such as milk and food containers. Polypropylene is used when enhanced strength, higher melting temperatures, and greater resistance to chemicals such as chemical holding tanks and automobile bumpers are required.

There are two prevalent methods of producing polyethylene, both of which involve heterogeneous catalysts. A slurry phase process utilizes chromium oxide deposited on SiO_2 dispersed in a solvent such as cyclohexane at 80–150°C and a pressure between 300 and 500 psig. The process operates in a recycle mode with a residence time of 2–3 h. The product containing the solvent and polymer is flashed leaving the polymer. The catalyst is usually left in the polymer because its concentration is extremely low. The operating conditions are adjusted to produce both high- and low-density polyethylene. The active site is Cr^{+2} produced by the reduction of Cr^{+6} by ethylene. The reaction mechanism proposed is that the polymer coordinates with one of the Cr^{+2} sites and the incoming ethylene coordinates with another site. Insertion of the ethylene into the double bond of polymer propagates its growth.

A second method of production utilizes the Ziegler–Natta $TiCl_4$ catalyst with liquid cocatalysts such as an alkyl aluminum halide. This is a reactive catalyst that must be prepared at the exclusion of air and water. The alkyl group of the cocatalyst coordinates with the Ti^{+3} site. The polymer grows by insertion of the ethylene into the double bond of the adsorbed polymer on another site.

Polypropylene

The most modern production route for polypropylene (PP) is also the Zeigler–Natta catalyst [36–38]. The catalyst is $TiCl_4$ supported on $MgCl_2$ along with aluminum alkyl halide cocatalyst such as diethyl aluminum fluoride $(CH_3CH_2)_2AlF$. The $MgCl_2$ is milled to a very disordered but active structure and the $TiCl_4$ is added. Production of the PP can be carried out in a fluidized gas phase reactor between 50 and 100°C and 100–600 psig. Ethyl benzoate is also used as part of the catalyst preparation and functions to reduce the $TiCl_4$ to active $TiCl_3$. The role of the alkyl component of the cocatalyst is to coordinate with the Ti^{+3} site where it inserts into the adsorbed polypropylene continuing the chain growth. The amount of catalyst used is so small it is retained in the final polymer product with no negative consequences. Unreacted gases are removed and recycled at the completion of the process. The most desirable product for the largest market is the isotactic form in which all CH_3 groups are on the same side of the polymer chain. Typically, it has a density of 0.9 g/cm^3, a melting point of 170°C, and an average molecular weight of 500,000. The polypropylene product is mixed in a separate reactor with ethylene to make a block polymer with enhanced mechanical properties.

Water, CO, and O_2 are the most significant poisons and are carefully removed upstream of the process.

The catalyst preparation and the process are far more complicated than presented here, so the reader is encouraged to refer to more detailed references [36–38].

Catalyst Challenges

Catalysts will have additional challenges as we move forward in the twenty-first century. In this author's mind, one of the most critical is the need to balance our rapidly expanding energy needs with the environment. Catalysts are already playing a dominant role in pollution abatement and in the production of specialty petroleum and chemical products. The main challenge will be to use bio-renewable energy sources (Chap. 33) as well as solar, wind, geothermal, etc. with the specific goal of freeing us from the use of fossil fuels. The hydrogen economy coupled with the fuel cell holds great promise as one road to meet this challenge [39]. Catalysts will play a key role in this pursuit, although the road map is not yet complete. This will be an exhilarating ride as we find our way to clean energy.

Another area of great importance is the application of biocatalysis using enzymes to produce a growing number of pharmaceutical agricultural products and fuels such as ethanol. This subject is outside the scope of this review so other sections of this Handbook (see Chap. 31) should be consulted.

References

1. Smith J (1981) Chemical reaction engineering. McGraw-Hill, New York
2. Schweich D (2003) Transport effects. In: Horvath I (ed) Encyclopedia of catalysis, vol 6. Wiley, New York, p 507
3. Fogler H (1992) Elements of chemical reaction engineering, 2nd edn. Prentice Hall, Englewood Cliffs, NJ
4. Liu X, Korotkikh O, Farrauto R (2002) Selective catalytic oxidation of Co in H_2: structural study of Fe-oxide promoted Pt/alumina catalyst. Appl Catal Gen 226:293
5. Bartholomew C, Farrauto R (2006) Fundamentals of industrial catalytic processes, 2nd edn. Wiley, New York, Chapter 2
6. Heck R, Farrauto R, Gulati S (2009) Catalytic air pollution control: commercial technology, 3rd edn. Wiley, New York, Chapters 2, 4, 7, and 9
7. Missen R, Mims C, Saville B (1999) Introduction to chemical reaction engineering and kinetics. Wiley, New York, Chapter 4
8. Broadbelt L (2003) Kinetics of catalyzed reactions. In: Horvath I (ed) Encyclopedia of catalysis, vol 4. Wiley, New York, p 472

9. Chorkendorff I, Niemantsverdrietm J (2003) Concepts of modern catalysis and kinetics. Wiley-VCH, Weinheim, Germany, Chapters 2 and 9
10. Bartholomew C, Farrauto R (2006) Fundamentals of industrial catalytic processes, 2nd edn. Wiley, New York, Chapter 1
11. Bartholomew C (2003) Catalyst deactivation/regeneration. In: Horvath I (ed) Encyclopedia of catalysis, vol 2. Wiley, New York, p 82
12. Bartholomew C, Farrauto R (2006) Fundamentals of industrial catalytic processes, 2nd edn. Wiley, New York, Chapter 5
13. Knozinger H, Maximlians L (2003) Catalyst characterization. In: Horvath I (ed) Encyclopedia of catalysis, vol 2. Wiley, New York
14. Bartholomew C, Farrauto R (2006) Fundamentals of industrial catalytic processes, 2nd edn. Wiley, New York, Chapter 3
15. Van Rysaselberghe V, Froment G (2003) Hydrodesulfurization. In: Horvath I (ed) Encyclopedia of catalysis, vol 3. Wiley, New York, p 667
16. Brunet S, Mey D, Perot G, Bauchy C, Diehl F (2005) On the hydrodesulfurization of FCC gasoline: A Rev Appl Catal Gen 278 (2):143
17. Gates B (2003) Catalysis by solid acid. In: Horvath I (ed) Encyclopedia of catalysis, vol 2. Wiley, New York, p 104
18. Csicsery S, Kiricsi I (2003) Shape selective catalysis. In: Horvath I (ed) Encyclopedia of catalysis, vol 6. Wiley, New York, p 307
19. Moser M (2003) Reforming. In: Horvath I (ed) Encyclopedia of catalysis, vol 6. Wiley, New York, p 1
20. Antos G, Aitani A, Parera J (eds) (1995) Catalytic naphtha reforming, science and technology. Marcel Dekker, New York
21. Soetaert W, Vandamme E (2008) Biofuels. Wiley, UK
22. Heck R, Farrauto R (2003) Automotive catalysis. In: Horvath I (ed) Encyclopedia of catalysis, vol 1. Wiley, New York, p 517
23. Heck R, Farrauto R, Gulati S (2009) Catalytic air pollution control: commercial technology, 3rd edn. Wiley, New York, Chapter 6
24. Albright L (1978) Hydrogenation (partial) of triglycerides. In: Peterson S, Johnson H (eds) Encyclopedia of food science. Avi Publishing, Westport, CT, p 398
25. Bartholomew C, Farrauto R (2006) Fundamentals of industrial catalytic processes, 2nd edn. Wiley, New York, Chapter 7
26. Albright L, Wisniak J (1962) Selectivity and isomerization during partial hydrogenation of cottonsead oil and methyl oleate: effect of operating variables. J Am Oil Chem Soc 39:14
27. Rostrup-Nielsen J, Nielsen TR (2002) Steam methane reforming. CATTECH 6(4):150
28. Osterkamp P (2003) Synthesis gas. In: Horvath I (ed) Encyclopedia of catalysis, vol 6. Wiley, New York, p 456
29. Nielsen J (2003) Hydrogen generation by catalysis. In: Horvath I (ed) Encyclopedia of catalysis, vol 4. Wiley, New York, p 1
30. Kondrantenko E, Baerns M (2003) Synthesis gas generation. In: Horvath I (ed) Encyclopedia of catalysis, vol 6. Wiley, New York, p 424
31. Pesce L, Jenks W (2006) Synthetic nitrogen products. In: Kent J (ed) Riegels handbook of industrial chemistry, 11th edn. Kluwer, New York
32. Bartholomew C, Farrauto R (2006) Fundamentals of industrial catalytic processes, 2nd edn. Wiley, New York, Chapter 8
33. Farrauto RJ, Hwang S, Shore L, Ruettinger W, Lampert J, Giroux T, Liu Y, Ilinich O (2003) New material needs for hydrocarbon reforming; Generating hydrogen for the prom fuel cell. Ann Rev Mater Res Soc 33:1
34. Shore L, Farrauto R (2003) Preferential oxidation of CO in H_2 streams. In: Vielstich W, Lamm A, Gasteiger H (eds) Handbook of fuel cells, 3, Part 2. Wiley, West Sussex, England, p 211
35. Kohlpaintner C (2003) Hydroformulation. In: Horvath I (ed) Encyclopedia of catalysis, vol 3. Wiley, New York, p 787
36. Vlad G (2003) Polymerization. In: Horvath I (ed) Encyclopedia of catalysis, vol 5. Wiley, New York, p 611
37. Rodriguez F (1996) Principles of polymer systems, 4th edn. Taylor & Francis, Washington, DC
38. Knapczyk J, Simon R (2006) Synthetic resins and plastics. In: Kent J (ed) Riegels handbook of industrial chemistry, 11th edn. Kluwer, New York
39. McNicol B, Williams K (2003) Catalysts for fuel cells. In: Horvath I (ed) Encyclopedia of catalysis, vol 2. Wiley, New York, p 387 and 42

Environmental Chemical Determinations

7

William L. Budde

Introduction

Environmental chemical determinations are identifications and measurements of the concentrations of elements, compounds, or ions in environmental media. In a chemical determination equal importance is given to the correct identification of the substance, and to its accurate and precise measurement. There has been a tendency in some environmental work to place more emphasis on making accurate and precise measurements, and to give less attention to ascertaining the correctness of the identification of the substance being measured.

Air, water, soil, and sediment are broad categories of environmental media and each of these can be divided into several subcategories, for example, ambient air, indoor air, industrial or workplace air, and vapor emissions from mobile or stationary combustion sources. Body fluids and tissue are also relevant environmental media because they are often analyzed to determine human, animal, and plant exposure to environmental chemicals.

Significance of Environmental Chemical Determinations

Accurate and precise identifications and measurements of specific chemical substances are fundamental to environmental studies and protection programs. Determinations are required to understand natural background concentrations of chemicals in the environment, the nature and extent of environmental pollution by anthropogenic chemicals, trends in concentrations of these substances, the transport and fate of chemical substances, and the causes of variations of concentrations in time and space. Accurate and precise determinations are also required to assess human health and ecological risks caused by exposure to natural and anthropogenic substances, establish air and water quality standards, develop pollution-control strategies, evaluate the effectiveness of pollution prevention and treatment technologies, and monitor compliance with and the effectiveness of discharge and other government regulations. These data are also needed to develop, calibrate, and verify mathematical models used to predict the impact of changes in concentrations of specific substances in the environment. Environmental chemical determinations are also required to set priorities and make cost estimates for the remediation of abandoned hazardous waste sites.

In addition to specific chemical substances, a variety of other determinations is very important and is often required for environmental studies and protection programs. These include identifications and measurements of bacteria, viruses, protozoa, and minerals such as asbestos fibers. Measurements of meteorological conditions, particulate matter in air by size, water turbidity, biological oxygen demand, chemical oxygen demand, and radioactivity are also very important. These and other similar determinations are beyond the scope of this chapter and the reader should consult other reference books for information about these topics.

Chemical Analysis Strategies

A wide variety of chemical analytical strategies is used in environmental analysis programs. In order to develop an appropriate strategy for a specific program, the objectives of the program must be thoroughly defined and understood by all participants. An analytical strategy for the program can then be developed by selecting and combining the most appropriate analytical methods and other key elements that will provide the results needed to fulfill the objectives of the

W.L. Budde (✉)
U.S. Environmental Protection Agency (retired), Cincinnati, OH, USA
e-mail: ohbudde@fuse.net

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_7, © Springer Science+Business Media New York 2012

program. Some of the major issues and strategies that must be considered in developing a specific strategy for an analysis program are described in this section.

Samples and Sampling Strategies

The goals of the environmental program will usually define the specific types of samples that must be acquired and analyzed, for example, ambient air, drinking water, ground water, soil, sediment, blood, sweat, urine, tissue, etc. The sampling strategy is a broad but detailed plan for the acquisition of the samples needed for the project. Sampling techniques and procedures are the actual physical processes used to acquire the samples. Sample preservation protects the integrity and validity of the samples before, during, and after the actual sampling process. Sampling strategies, techniques, and sample preservation are major factors contributing to the success or failure of environmental chemical determinations.

The sampling strategy defines in detail what, where, when, and how the environmental samples will be acquired. This plan should specify all details including the number of sampling stations for air, water, soil, sediment, or fish samples; the locations of the sampling stations; the time of the day of sampling; the frequency of sampling; the depth of sampling for water, soil, and sediment samples; the meteorological conditions and altitude for air sampling; and, the need for replicate samples taken at the same time and place. For body fluids and tissue samples myriad details must be considered including the number of persons contributing or the kinds of wildlife-related samples that must be collected. The plan should specify a series of discrete samples, a composite sample from contributions at various intervals, or a continuous flow of sample over a period of time. A sampling strategy that does not provide appropriate and representative samples seriously jeopardizes the value of the chemical determinations.

The sampling techniques and procedures define the type of sampling equipment, the containers used for the samples, the procedures used to clean the sample containers and sampling equipment, the calibration of sampling equipment, and sample compositing procedures if composite samples are required. Sample preservation must ensure that the chemical composition of the sample at the time of analysis is the same as it was at the time and place of sampling and is not the result of physical or chemical changes caused by the conditions of shipment and storage prior to chemical analysis. Significant research has been conducted to find sample containers that resist adsorption or degradation of sample components. Samples are often shipped and stored at low temperatures, and treated with various chemical reagents, to retard microbiological degradation of the analytes or to prevent chemical reactions that can change the composition of the sample. Appropriate sampling techniques and preservation procedures are required to provide valid samples and valid chemical determinations.

Determination of Total Elements or Total Related Substances

A standard analysis strategy is the determination of the total amount of an element in a sample where the element is present in several or more elemental forms, compounds, ions, oxidation states, or physical phases. This strategy was developed before chromatographic and other techniques were available to separate the individual substances containing the element of interest. It is widely used in environmental studies and government regulatory programs. However, the meaning of *total* is variable because some analytical methods do not include all the physical phases or chemical forms in the determination. The sample preparation procedures in individual analytical methods define exactly which elemental forms, compounds, ions, oxidation states, and physical phases are included in a total measurement. A similar strategy is the determination of the total amount or total concentration of a group of closely related compounds or ions without specifying the relative or absolute concentrations of the individual substances. Closely related substances usually have some common physical or chemical properties, elemental composition, or structure.

The broadest definition of, for example, total mercury (Hg) in a sample is the sum of the elemental Hg^0, the Hg in all inorganic compounds of Hg^1, the Hg in all inorganic compounds of Hg^2, and the Hg in all organic compounds in which Hg is bonded to C, O, N, S, or some other element. All phases are included because some species may be present in the vapor phase (Hg^0), some soluble in water, and some insoluble in water or present in the particulate phase of an air sample. More limited definitions of total are often used in analytical methods that separate physical phases or chemical forms that contain the element of interest.

If a water sample contains both soluble and insoluble manganese (Mn) compounds and ions, and it is filtered to separate the dissolved and insoluble fractions, and the filtrate and insoluble residue are analyzed separately, the results can be expressed as total dissolved Mn and total suspended or insoluble Mn. Phosphorus (P) can be determined colorimetrically as the ortho-phosphate ion, PO_4^{-3}, in aqueous samples after a reaction that forms an intensely blue-colored derivative. However polyphosphate ions and other ions and compounds containing P do not form this derivative. Total P in a sample can be determined with the same colorimetric procedure after acid hydrolysis and oxidation of all ions and compounds containing P to PO_4^{-3}. In some elemental analyses the sample is treated with reagents designed to make available for measurement some fraction of an element or elements but not the total amount. For example, a soil sample may be treated with water at pH 3 to simulate the leaching process of acid rain. A total elemental analysis of the filtrate provides information about just those elements solubilized by the mild acid treatment. This can be called the determination of total mild acid leachable elements.

The measurements of total organic carbon (TOC) and total organic halogen (TOX) in a sample are used to assess many types of environmental samples. The analytical methods for TOC employ procedures to physically separate the inorganic carbon, that is, carbonate, bicarbonate, cyanide, and other inorganic substances containing C, from the organic compounds and ions in the sample. The TOC is then measured by oxidation of the organic compounds and ions to carbon dioxide which is determined by one of several techniques. The analytical methods for TOX use procedures to separate inorganic halogen-containing ions from halogen-containing organic compounds and then determine the total halogens in the organic compounds.

Some analytical methods have been developed for the determination of groups of closely related compounds or ions. The classic example is the measurement of combined phenolic materials colorimetrically after a reaction that forms a red derivative with many phenols. However, different phenols form derivatives with somewhat different visible absorption spectra and various phenols have different reactivities with the derivatizing reagent. Therefore this method is calibrated with pure phenol and the measurement gives just an estimate of the total phenols in the sample. For this reason this method is probably no longer widely used and phenols are usually determined as individual compounds with other analytical methods. Subgroups of complex mixtures of congeners are sometimes measured together for convenience of interpretation or for government regulatory programs. These determinations give concentrations of, for example, total tetrachlorobiphenyls, total pentachlorobiphenyls, etc. Similar determinations of chlorinated dibenzo-*p*-dioxins and chlorinated dibenzofurans at each level of chlorination are specified in some analytical methods.

Determination of Specific Substances

Before the development of efficient chromatographic separation techniques and selective and sensitive detectors, analytical methods for the determination of specific analytes in environmental samples were very limited. Those methods depended on highly selective chemical reactions that are relatively rare and difficult to discover, or on very selective physical measurements such as atomic absorption or emission techniques for elemental analytes. Therefore only a relatively few analytical methods for the most common and amenable organic and inorganic compounds or multi-element ions were developed. It would have been essentially impossible, and enormously costly, to develop, test, document, and implement a large number of nonchromatographic analytical methods for a wide range of often similar organic and inorganic analytes.

Vapor phase or gas chromatography (GC) and high-performance liquid chromatography (HPLC) provided capabilities for the separation of microgram (10^6) and smaller quantities of often similar individual substances in complex mixtures. A variety of GC and HPLC detectors were developed and, depending on the detector, the separated substances could be determined with good to excellent sensitivity and with fair to excellent reliability. It was soon recognized that determinations of a broad range of specific chemical substances was not only practical but also essential to achieve the goals of many environmental studies and protection programs.

Gas chromatographic techniques were first applied during the 1960s to synthetic organic compounds and natural products which are often complex mixtures of organic compounds. The earliest applications of these techniques in environmental research were GC-based determinations of petroleum hydrocarbons in polluted air and chlorinated hydrocarbon pesticides in several types of samples. Analytical chemists and environmental scientists concerned about organic pesticides, industrial organic chemicals, and other organic compounds emphasized the development of analytical methods for the determination of specific organic compounds. In contrast, the emphasis in elemental and inorganic analyses during the 1960–1980s was, with a few exceptions, on determinations of the total amounts or concentrations of specific elements in a sample without regard to the specific compounds or ions containing those elements. With the recognition that toxicity and other environmentally significant properties varied widely with the specific compound or ion, inorganic analytical chemists later focused on chromatographic separations of individual species. The term *speciation* came into use, mostly by inorganic chemists, to distinguish this type of analysis from the conventional determinations of the total amounts or concentrations of the elements in a sample.

There are two general strategies for the determination of specific organic or inorganic compounds and ions in environmental samples [1]. The target analyte (TA) strategy dominated analytical chemistry before the development of chromatographic separation techniques and is by far the most commonly used with contemporary separation and detection techniques. The broad spectrum strategy became feasible with the development of high-resolution chromatographic separation techniques and spectroscopic detectors.

Target Analyte (TA) Strategy. Target analytes are known substances with known chemical, physical, and other properties. They are either known or thought to be in samples and they must be determined to meet the objectives of the environmental analysis program. The TA strategy is the analysis of the sample with an analytical method that is designed and optimized to determine the target analyte or a group of similar and separated target analytes. If the target analytes are a diverse group with sufficiently different chemical and physical properties, they are divided into subgroups according to their similar properties. Several optimized analytical methods and separate environmental samples are used to determine the analytes in the subgroups.

The TA strategy has many advantages that favor its widespread application. Sample preparation procedures can be designed to separate the target analytes from the sample matrix with maximum efficiency, and to concentrate them in a suitable solvent for further chromatographic separation. Some interferences can be separated from the analytes during sampling or sample processing by pH adjustments, chemical derivatization, evaporation of non-target substances, or other techniques. Chromatographic separation and detection techniques can be selected to give the best practical resolution of analytes, selective detection, and the lowest detection limits. The detector can be calibrated for quantitative analysis with standard solutions of the target analyte or analytes, and the analytical method tested, perfected, and validated using test sample matrices fortified with known concentrations of the target analyte or analytes.

The target analyte strategy is widely accepted, understood, and used in analytical chemistry, environmental research, environmental protection programs, and in many other fields of investigation. It is used in most environmental quality surveys and government regulatory compliance monitoring programs. Many analytical methods for a variety of target analytes have been developed, tested, documented, and implemented [1–4]. Cost estimates for analyses are readily made and analytical costs are not difficult to control. The target analyte strategy is used in the vast majority of chemical analyses reported in the scientific literature.

The Broad Spectrum (BS) Strategy. The objective of the BS strategy is to discover the substances present in the sample and to measure their concentrations without a predetermined list of target analytes [1]. As a starting point, the BS strategy may follow the general scheme of a target analytical method, but with minimum sample processing to allow a broad variety of generally similar substances to reach the chromatographic separation and the detector. Procedures to remove target analyte interferences are minimized or not used to avoid discarding interesting and potentially important components. If the sample is very complex, it is usually divided into fractions that are likely to contain components with similar properties. Several or more different types of chromatographic separations may be required for the various fractions depending on the components of the sample and the breadth of information desired. The chromatographic detectors are generally spectroscopic detectors that can provide information about the composition and structure of the sample components. This information is used to identify known substances, unexpected substances, and even unknown substances. The most important and widely used detector for the broad spectrum strategy is the mass spectrometer [1], but infrared, nuclear magnetic resonance, and other spectroscopic techniques are sometimes used.

The identification of all or most substances in one or more chromatograms, even with extensive spectroscopic data, can be a challenging and difficult process. The available data may not be sufficient to even tentatively identify all the components, especially if pure authentic samples of suspected substances are not available in the laboratory. Calibration of a broad spectrum method for quantitative analysis is delayed until the desired components are identified. For these reasons, and the general preoccupation with target analytes, the BS strategy is much less common than the target analyte strategy.

The BS strategy is obviously important for the discovery of unknown naturally occurring substances and anthropogenic chemicals in the environment. However, BS strategies are more difficult than TA strategies to develop, document, and implement. Cost estimates for analyses are difficult to make because the number of substances found, identified, and measured is not known until after the samples are analyzed. Therefore, analytical costs are difficult to control. Because of the potential costs, the instrumentation requirements, the technical skills needed, and time required, the BS strategy is not often employed especially by programs with limited objectives and budgets. A strategy sometimes used in environmental studies is to develop a target analyte method that can meet the objectives of the study and give some attention to other chromatographic peaks to identify potentially new or unexpected substances.

Single-Analyte and Multi-Analyte Methods

Traditional analytical methods, which were generally developed prior to the widespread application of GC and HPLC techniques, were nearly always designed for a single target analyte, for example, the colorimetric determination of PO_4^{-3}. With the development of GC and HPLC separation techniques, the determination of several or more similar compounds or ions in a sample was feasible and multi-analyte analytical methods were developed and documented. However, analytical chemists did not generally attempt to include more than about 10–20 target analytes in a method. This strategy was necessary because early GC and HPLC columns were not very efficient and most GC and HPLC detectors were either not selective or had limited selectivity. Sample preparation procedures designed to reduce or eliminate interferences continued to be very important. Analytical methods for chlorinated hydrocarbon pesticides with a GC separation and an electron capture detector require appropriate sample preparation. These multi-analyte pesticide methods are sometimes called multi-residue methods because the pesticides are residues in crops and other samples.

As more efficient high-resolution chromatographic separation techniques were developed, and spectroscopic detectors came into widespread use, the number of target analytes in multi-analyte methods was increased. The separation of most or all target analytes is often feasible and spectroscopic detectors usually provide sufficient information to make correct identifications of target analytes even when some are not fully separated. Well-tested and documented analytical methods for 80–100 or more analytes in some types of samples are presently available [1, 2]. Multi-analyte methods significantly reduce the cost of an analysis on a per analyte basis and add support to the already strong justification for high-resolution chromatographic separations and spectroscopic detectors.

Remote Laboratory Analyses and Field Analyses

Chemical analyses can be conducted in a laboratory remote from the locations where the samples are taken or in the field near the sampling sites (on-site). On-site analyses can be conducted in a field laboratory which may be a temporary building or a truck trailer, van, or recreational vehicle equipped with utilities services and analytical equipment. Another type of field laboratory is within a materials or fluids processing facility. Alternatively, field analyses can be conducted with mobile or portable instrumentation carried in a small van, sport utility vehicle, moved with a hand cart, or carried by a person. Each of these strategies has some advantages and some disadvantages.

The remote laboratory has the advantages of providing carefully controlled temperature, humidity, ventilation, and background conditions with adequate space and utilities to support a large array of major analytical instrumentation and a staff of skilled analytical chemists and technicians working in a convenient, comfortable, and safe environment. The major disadvantage is that environmental samples must be carefully preserved, shipped, and stored prior to analysis. Furthermore, the analytical results may not be available for several days or weeks because of the time required to transport the samples to the laboratory, incorporate the analyses into work schedules, and service the multiple clients of a remote laboratory.

Field determinations have the major advantage of greatly reducing the time between acquisition of the samples and the availability of the analytical results. This may allow utilization of the results quickly which can provide significant cost savings in the field operations that utilize the analytical information. For example, the rapid availability of results from field determinations can be used to fine-tune a sampling strategy to obtain the most significant samples for detailed remote laboratory analysis. Similarly, the results from field determinations can be used to direct the work of construction crews or well drillers using heavy and costly equipment. Mobile or portable instrumentation can provide rapid and low-cost results from a large number of samples taken over a broad geographic area in a short period of time.

The analytical methods that are feasible in the field may be significantly limited compared to what is feasible in a remote permanent laboratory. If a field laboratory is located in a temporary building or a large truck trailer, and adequate utilities and personnel are available, many of the kinds of analytical methods that are routinely implemented in a remote laboratory may be feasible in the field. However, because of space and power limitations, a broad variety of instrumentation is usually not available and the number of different analytical methods that can be implemented is smaller than in a remote permanent laboratory. Mobile or portable instrumentation is usually more limited and generally sample analyses are less complete and detailed than in a field or remote laboratory.

Discrete Samples and Continuous Monitoring

Field analyses can be conducted with discrete or composite samples similar to the samples used in a remote laboratory, or by continuous monitoring of substances in a flowing stream of gases or liquids. Continuous monitoring is required when the results are needed within a period of time that is shorter than the time required to acquire and to analyze conventional discrete samples. Continuous monitoring is often required to determine substances in a processing facility and provide rapid feedback of results that are used in process-control strategies. If continuous determinations are made with sufficient speed to permit changes in sampling or other strategies while the determinations are in progress, these are called real-time analyses.

Continuous environmental monitoring in the field is needed when the substances present and their concentrations are changing rapidly. Discrete samples taken at inappropriate times will give results that do not correctly assess the variable conditions in the atmosphere or a flowing stream. Sample compositing may provide a better assessment, but continuous monitoring and integration of the determinations over time gives the best assessment. The determination of the sources of fugitive emissions that are rapidly dispersed in the atmosphere or in a flowing water stream requires continuous and sometimes mobile continuous monitoring. The concentrations of air pollutants, such as CO, NO_x, SO_2, and O_3, that vary widely with sunlight, automobile traffic, wind speed, wind direction, and other meteorological conditions are monitored continuously.

Sampling and analytical equipment used in processing facilities and in the field have some of the same requirements. However, equipment used in a facility may have fewer constraints in regard to size, weight, and power requirements than mobile and portable field equipment. Standard chromatographic separation techniques are generally too slow for continuous measurements although GC separations, especially fast GC, are used for process control.

Analytical Quality Assurance and Control

Analytical quality assurance (QA) is a broad program of actions designed to ensure that the chemical determinations are of known and acceptable quality. The QA program encompasses all aspects of the chemical analysis from the design of the sampling strategy to the documentation of the results. Analytical quality control (QC) is the implementation of specific actions designed to control the quality of the determinations at some defined level of acceptance. All physical measurements have some degree of variability and uncertainty and the QA/QC program should define the limits of these. The quality of analytical determinations is judged by the attributes, or figures of merit, of an analytical method, which include selectivity, sensitivity, detection limits, signal/noise, recovery, accuracy, bias, precision, and validation. These attributes are determined using a variety of special measurements and fortified samples (QA/QC samples) which are described in the analytical method. The time relationship between the analyses of the environmental samples and the determination of the quality attributes is critical in estimating the quality of the environmental determinations.

Published environmental chemical determinations are often used in ways that are not anticipated by the original investigators. The determinations may be used by other investigators to estimate human or ecological exposure, assess the level of industrial discharges, or to develop environmental quality standards or discharge limitations. Therefore the results of the QA/QC measurements, and their time relationships to the analysis of environmental samples, should be permanently attached to the environmental sample results. These QA/QC data should demonstrate that the environmental data were obtained under controlled conditions that provide credibility to the results. Other users of the data, sometimes years in the future, should be able to assess the quality and applicability of the information and take into account the variability and uncertainty in the determinations.

Development and Documentation of Analytical Methods

The complete development and documentation of an analytical method usually occurs over a period of several years or more. The germ of most new methods is often a research project in which a determination is needed. Organizations that specialize in developing and manufacturing analytical instrumentation frequently play a major role in supporting the development of new analytical methods. But instrumentation alone does not constitute an analytical method. The general acceptance and widespread use of methods depends on many factors including the needs for research or environmental monitoring of various substances, the cost and complexity of instrumentation, the required laboratory or field skills, and especially government regulations. There is a high degree of variability in the completeness, documentation, testing, and validation of published analytical methods.

Research Methods

The basic concepts and fundamentals of most analytical methods are usually first published in scientific journals, for example, The American Chemical Society publications.

Analytical Chemistry, Environmental Science and Technology, and the *Journal of the American Chemical Society*. These descriptions are usually brief and often just summarize the techniques and procedures of the method. Research reported in scientific journals is often focused on a detailed investigation of a narrow subject area, and new analytical techniques and procedures developed for the research are rarely tested in a broader context. Analytical quality assurance and control are often minimal or not described. Techniques and procedures described in scientific research journals are the beginnings of analytical methods, however, they usually require considerable development, modifications, and testing before they become widely accepted analytical methods.

Methods in Development

Academic research groups, industrial laboratories, and government agencies interested in conducting environmental surveys or monitoring programs often adapt research techniques and procedures to the broader needs of the survey or program. During this stage considerable experimentation is underway to evaluate equipment, instrumentation, chemical reagents, and other materials used in the method. The incipient method may be tested with many potential analytes and sample matrices. Modifications are implemented to minimize or eliminate problems discovered during the development process. Techniques and procedures are developed for identifications of analytes, calibration of instruments for quantitative analysis, and analytical quality assurance and control. Methods in development may be described in a series of draft versions that are not usually formally published but may appear in bound or unbound technical reports. These methods are described in more detail than in scientific journal articles, and descriptions often contain detailed information about required equipment, supplies, reagents, instrumentation, and personnel skills. The technical reports and draft method descriptions are often distributed informally through personal contacts, scientific conferences, or the Internet.

Methods Published by Standard-Setting Organizations

Standard-setting organizations that publish analytical methods for environmental analyses include The American Society for Testing and Materials (ASTM) [4], The American Public Health Association (APHA) [3], The American Water Works Association (AWWA), The Water Environment Federation (WEF), The International Standards Organization (ISO), and the Association of Official Analytical Chemists (AOAC). These organizations assemble working committees of experts that consider developed, widely used, and generally accepted analytical methods for publication as standardized analytical methods. Requirements of individual organizations vary and these may include a description of the method in a specific editorial format and the availability of multi-laboratory validation data for the proposed analytical method. Some of these organizations also sponsor or participate in multi-laboratory validation studies of proposed analytical methods. When an analytical method is published by one of these organizations it usually has been thoroughly tested and used by many laboratories and analysts, and is widely known and accepted by specialists in that type of determination. Published methods are reviewed periodically, updated as needed, and may eventually be replaced by entirely new methods.

Methods Published or Referenced in Government Agency Regulations

The US Environmental Protection Agency (USEPA) and other federal, state, and local government agencies in the United States may require chemical analyses to determine compliance with air quality, water quality, liquid waste discharge, solid waste disposal, and other environmental regulations. The analytical methods specified for these purposes may be methods in development, methods published by standard-setting organizations, or new methods documented in the regulatory proposals [2]. The USEPA has promulgated two types of analytical methods. Some regulations, for example, some drinking water and waste water regulations, require either analytical methods designated in the regulations or approved alternative test methods for compliance monitoring. Other regulations, for example, some USEPA solid waste regulations, include or reference analytical methods that are suggested or optional but allow any other appropriate analytical methods.

Characteristics of Analytes, Samples, and Sampling Techniques

The physical and chemical properties of analytes and the nature of the sample have a major impact on, and often limit, the sampling and other procedures and techniques that can be employed in an analytical method. Major issues that must be considered when developing an analytical method are the volatilities, thermal stabilities, photochemical stabilities, polarities, water solubilities, and chemical reactivities of the sample components or target analytes; the physical state of the sample; and the nature of the sample matrix. Analytes, whether organic or inorganic, can be broadly divided into three categories based partly on vapor pressure (VP), or volatility, at ambient temperature and on some other physical and chemical properties. There are major differences in the procedures and techniques used to acquire and process condensed-phase and vapor-phase samples.

Sampling ground water requires expert selection of procedures and techniques to avoid significant analytical errors.

Volatile Analytes

Volatile analytes are usually defined as those having vapor pressures (VPs) greater than about 0.1 Torr at 25°C and an external pressure of 760 Torr (1 Torr 1 mm of Hg or 133 pascals (Pa)). Figure 7.1 shows the structures, molecular weights (MW), boiling points (BP), and VPs of three representative volatile compounds of environmental interest. Most volatile analytes have MWs below 200 but a low MW does not guarantee that an analyte will be volatile. Many substances have MWs below 200 but they are not volatile because they are ionic or have polar groups of atoms or engage in hydrogen bonding with other molecules. Some compounds have high VPs, for example, MTBE in Fig. 7.1, but reduced volatility in water because of their high water solubilities. Boiling points of volatile substances range

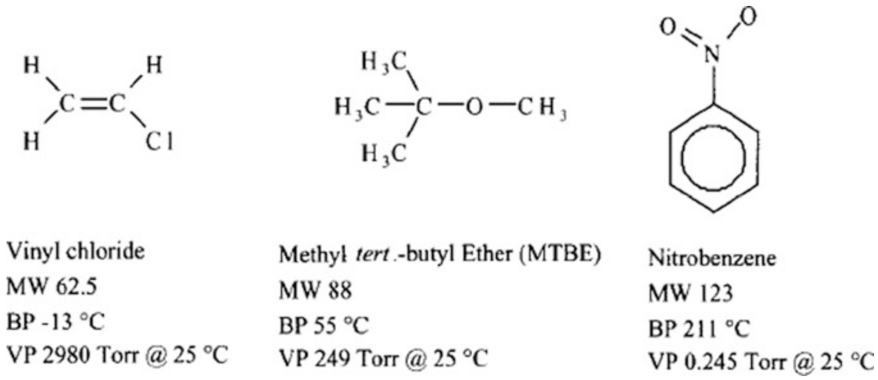


Fig. 7.1 The structures, molecular weights (MW), boiling points (BP), and vapor pressures (VP) of three representative volatile compounds of environmental interest

from below 0°C to above 200°C. Some compounds with BPs at the upper end of this range are surprisingly volatile. For example, nitrobenzene in Fig. 7.1 has a BP of 211°C but still has a VP of 0.245 Torr at 25°C. At normal room temperatures the distinctive odor of nitrobenzene vapor is readily detected by most people.

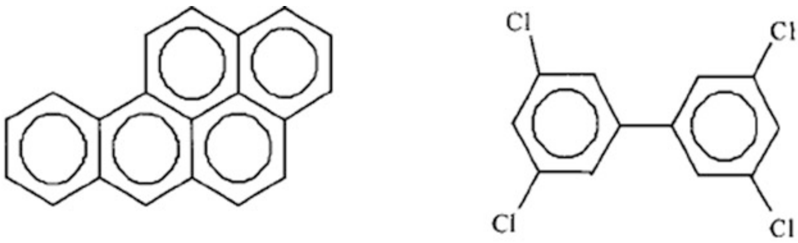
Semivolatile Analytes

Semivolatile analytes are usually defined as those having VPs in the range of 0.1 Torr to about 10^9 Torr, but this range is approximate and some substances with VPs in this range are considered volatiles or even nonvolatiles (next section). For example, Hg^0 has a VP of 2 10^3 Torr at 25°C, but is often considered a volatile analyte. Most semivolatile analytes have molecular weights in the range of 100–500, but a MW in this range does not guarantee that a compound is a semivolatile. However, because semivolatiles nearly always have higher MWs and contain more atoms than the volatiles, they have a significantly larger number of isomers, congeners, and chiral forms. Semivolatiles are usually devoid of structural groups that are susceptible to thermal decomposition below about 300°C, or cause high polarity, or impart high water solubility, or are very chemically reactive. Figure 7.2 shows the structures, MWs, some melting points (MPs), and VPs of four representative semivolatile compounds of environmental interest. Both benzo[a]pyrene and 2,3,7,8-tetrachlorodibenzo-*p*-dioxin have VPs at the low end of the semivolatile range, but they have little or no polarity and tend to behave as do other semivolatiles with higher VPs. The VPs of semivolatile compounds are generally insufficient to give vapor concentrations at ambient temperatures that can be detected by a distinctive odorb.

The large number of potential congeners and isomers of some semivolatile compounds is illustrated in Table 7.1, which shows the numbers of possible chlorinated biphenyl, chlorinated dibenzo-*p*-dioxin, and chlorinated dibenzofuran congeners and isomers. The total number of chlorinated congeners of each parent compound is the sum of the possible isomers at each level of chlorination. Dibenzofuran has more possible chlorinated congeners and isomers than chlorinated dibenzo-*p*-dioxin (Fig. 7.2) because it has just one-ring oxygen and a less symmetrical structure. If F, Br, I, CH_3, or any other uniform substituent replaces Cl, the same numbers of possible congeners and isomers would exist. If different atoms or groups of atoms were mixed as substituents on these parent compounds, a significantly larger number of different substances, congeners, and isomers are possible.

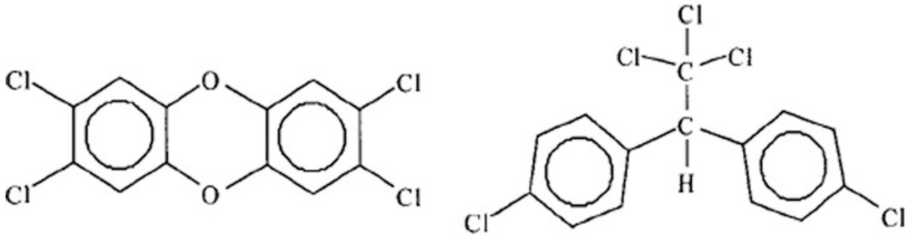
Nonvolatile Analytes

All other compounds and essentially all ions are classified as nonvolatile. Substances in this category have VPs lower than about 10^9 Torr or have structural groups that are susceptible to thermal decomposition below about 300°C, or that cause high polarity, or that impart high water solubility, or that are very chemically reactive. Figure 7.3 shows the structures, MWs, some MPs, and VPs of six representative compounds of environmental interest from this group. All these examples except the herbicide glyphosate have reported VPs in the semivolatile range, but they do not behave as typical semivolatile compounds. Nitroglycerine and *N*-nitrosodiphenylamine undergo thermal decomposition at temperatures well below 300°C. The pesticide carbofuran and the herbicide diuron are more thermally stable, but they also tend to decompose below 300°C and on hot surfaces. The polarities of the functional groups of this class of compounds cause adsorption on polar surfaces and dipole–dipole interactions with other molecules. Nearly all carboxylic acids, for example, the herbicide 2,4-D in Fig. 7.3, are susceptible to decarboxylation at elevated temperatures and they interact with basic substances or basic surfaces which reduces their volatility. Similarly, basic substances such as amines and some other nitrogen compounds interact with acids and acidic surfaces which reduce their volatility.



Benzo[α]pyrene
MW 252
MP 179 °C
VP 5.49×10^{-9} Torr @ 25 °C

3,3',5,5'-Tetrachlorobiphenyl
MW 292
VP 8.45×10^{-6} Torr @ 25 °C



2,3,7,8-Tetrachlorodibenzo-*p*-dioxin
MW 322
MP 305 °C
VP 1.5×10^{-9} Torr @ 25 °C

1,1-bis(4-Chlorophenyl)-2,2,2-trichloroethane (4,4'-DDT)
MW 354.5
MP 109 °C
VP 1.60×10^{-7} Torr @ 20 °C

Fig. 7.2 The structures, MWs, some melting points (MP), and VPs of four representative semivolatile compounds of environmental interest

Table 7.1 The number of possible chlorinated biphenyl, chlorinated dibenzo-*p*-dioxin, and chlorinated dibenzofuran congenors and isomers

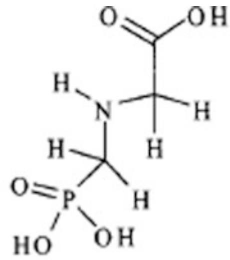
Parent compound	Total congeners	Cl_1	Cl_2	Cl_3	Cl_4	Cl_5	Cl_6	Cl_7	Cl_8	Cl_9	Cl_{10}
Biphenyl	209	3	12	24	42	46	42	24	12	3	1
Dibenzo-*p-dioxin*	75	2	10	14	22	14	10	2	1		
Dibenzofuran	135	4	16	28	38	28	16	4	1		

Condensed-Phase Samples

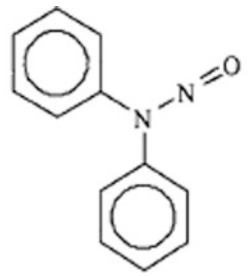
If only semivolatile and nonvolatile components are of interest, or the sample only contains these categories of analytes, sampling procedures for most condensed-phase samples are not difficult and require only a few precautions. These include the materials used for sample containers, the cleaning of sample containers, and the preservation of the sample. Water samples for elemental analyses are treated with acid at the time of sampling, or well before analysis, to reduce the pH to 2. This ensures the solubility of metal-containing ions which can precipitate or adsorb on container walls at a pH 2. Glass containers are not used for elemental analysis samples because aqueous acid solutions can leach trace elements from glass. Samples are taken in plastic containers which provide the added benefits of reduced weight compared to glass, little or no breakage during handling and shipping, and single-use containers that can be disposed of at a plastic recycling facility. However, samples for the determination of organic analytes are taken in glass containers to prevent background contamination of the samples by organic compounds that can leach from plastic materials. These samples are often preserved by addition of acid to make the sample pH 2 to retard microbiological degradation of some analytes. Acid leaching of glass containers generally has no effect on the concentrations of organic analytes.

Condensed-phase samples containing volatile analytes require special techniques because the volatile components are elusive and can be readily lost during sampling and

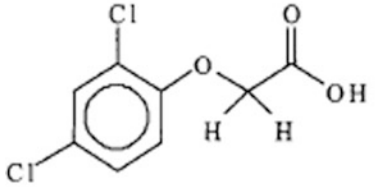
Fig. 7.3 The structures, MWs, some MPs, and VPs of six representative nonvolatile compounds of environmental interest



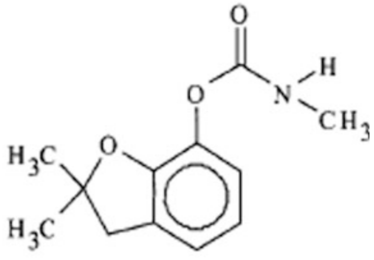
Glyphosate
MW 169
MP 200 °C
VP 2.89×10^{-10} Torr @ 25 °C



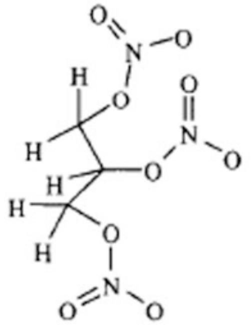
N-nitrosodiphenylamine
MW 198
VP 1.98×10^{-3} Torr @ 25 °C



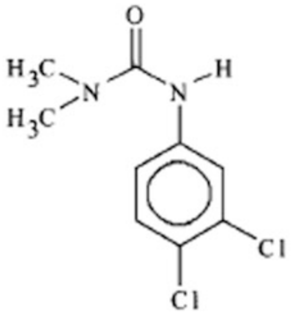
2,4-Dichlorophenoxyacetic acid (2,4-D)
MW 221
MP 141 °C
VP 6×10^{-7} Torr @ 25 °C



Carbofuran
MW 221
MP 151 °C
VP 4.85×10^{-6} @ 19 °C



Nitroglycerine
MW 227
Decomposes 50-60 °C
VP 2.6×10^{-4} Torr @ 20 °C



Diuron
MW 233
MP 158 °C
VP 2.7×10^{-6} Torr @ 30 °C

sample processing. Water samples must be taken with little or no agitation of the sample and poured into the container so there is no air space, often called a headspace, between the top of the sample and the air-tight seal of the container. This prevents vaporization of volatile components into the headspace during shipment and storage which reduces the concentration in the aqueous phase. Aliquots of water samples are removed from sample containers by inserting a syringe needle through the inert air-tight septum seal and drawing water into a gas-tight syringe. Solid samples, for example, soils and sediments, may also contain volatile components trapped in pores or dissolved in associated water and these components are easily lost during sample handling. Solid and semisolid samples are carefully and quickly placed in wide-mouth sample containers that are sealed in the field and not opened during any subsequent processing of the sample. Volatile components are usually partitioned into water, organic solvents, or the vapor phase and the extracts are analyzed using techniques developed for liquid- or vapor-phase samples. Water or an organic solvent are added to the sample containers through inert and air-tight septum seals with a syringe.

Ground water sampling is a significant challenge because water must be lifted to the surface through a bore hole from depths ranging from a few meters to 75 m or more. The lifting process can disturb the equilibrium between dissolved analytes and analytes associated with particulate matter, and significantly change the temperature and pressure of the water sample. These factors can cause changes in concentrations of analytes and raise questions about the representativeness of the sample. Volatile components of ground water samples are particularly susceptible to losses caused by temperature and pressure changes and degassing of the sample. Several types of down-hole water samplers are available and several types of pumps are used to purge a well and lift a stream of ground water to the surface. The materials of construction of the down-hole samplers, pumps, tubing used with pumps, and well casings can also have an impact on the concentrations of analytes in the sample. Considerable research has been conducted to evaluate the materials and techniques used to sample ground water, however, uncertainties remain because there is no satisfactory ground water standard of reference for comparison of various techniques.

Semivolatile analytes in vapor-phase samples are often associated with particulate matter or aerosols that are collected on glass or quartz fiber filters in a flowing air stream [1]. However some semivolatile analytes that have higher VPs, for example, 3,3′,5,5′-tetrachlorobiphenyl in Fig. 7.2, can slowly vaporize from a particle trapped on a filter in the flowing air stream, and will be lost unless captured by an in-stream sampling device. Small glass or metal tubes containing polyurethane foam are often used to capture vaporized semivolatile analytes. Other solid-phase adsorbents, which are described in the next section, are also used to trap semivolatile analytes vaporized from particulate filters.

Vapor-Phase Samples

There are two general approaches to sampling air, or vaporous emissions from stationary (stack) and mobile (automobile, truck, etc.) sources, for the laboratory determination of volatile analytes [1]. Bulk vapor-phase samples can be taken in the field in various containers and transported to a remote or field laboratory for analysis. Containers used for bulk vapor-phase samples include flexible polyvinyl fluoride (Tedlar™) bags, evacuated glass or metal reservoirs, and thermally insulated cryogenic collection vessels. Alternatively, the volatile analytes can be separated from the main components of air in the field and just the analytes and their collection devices transported to the laboratory. The principal techniques used to separate volatile analytes from air in the field are cryogenic traps, impingers, and solid-phase adsorbents.

Bulk Vapor-Phase Samples. Flexible plastic sample bags are generally limited to vapor-phase samples that can be analyzed within a short time after sample collection. This limitation is due to potential losses of analytes by surface adsorption and surface chemical reactions. Transportation of inflated bags over long distances to a remote laboratory is cumbersome and can result in total losses of samples due to punctures and other accidents. Plastic bags are used in some laboratory operations, for example, to collect automobile exhaust and vaporized fuel, and for samples that can be conveniently transported to a laboratory.

Evacuated stainless steel canisters are widely used collection devices for ambient air samples [5]. Sample canisters have smooth and inert internal surfaces and few or no active sites that adsorb volatile analytes or catalyze chemical reactions. The 1–6 L canisters are easily transported to a remote or a field laboratory. Canisters are leak tested and cleaned in the laboratory prior to use, evacuated to about 5 10^2 Torr or less, and transported to the sampling site where samples are taken by opening the sampling valve. Composite samples can be taken over time and/or space and an in-line pump can be used to pressurize the container with either additional sample air or pure air if sample dilution is required. Pressurized samples are useful when longer-term composite samples are taken or when larger samples are needed to lower detection limits.

Condensation of an entire air sample with liquid nitrogen or liquid helium has been used for many studies. However, this technique is expensive to implement and requires specialized portable equipment for handling cryogenic fluids in the field.

Separated Analytes. Cold trapping is used to separate volatile analytes from the main components of air in the field. Air is drawn by a pump through an inert, often nickel, metal tube immersed in a fluid at a very low temperature, for example, 150°C. The tube may be packed with some inert material such as Pyrex™ glass beads and the temperature is sufficient to condense most analytes but insufficient to condense oxygen or nitrogen.

Plugging of the condensation tube with ice or other solids can be a major problem when sampling large volumes of humid air. This problem is addressed by the employment of air dryers that trap moisture but allow the nonpolar volatile analytes to pass into the cold stage of the trapping system. However, the more polar and water-soluble analytes are also removed by efficient air drying systems. Another potential problem is that some analytes may react during subsequent processing with trapped ozone, nitrogen oxides, or other substances present in the air. A practical limitation of the cold trapping technique is the requirement for liquid nitrogen or liquid argon in the field during extended sampling periods.

Impingers are used to extract various substances from vapor-phase samples [1]. An impinger is a closed glass or metal vessel with an inlet tube that extends to near the bottom of a liquid and an outlet tube well above the surface of the liquid. The impinger may contain various aqueous or nonaqueous liquids including solutions of derivatizing agents. Vapor is drawn by a pump into the inlet tube and bubbled through the liquid which dissolves soluble analytes or the analytes react with reagents in solution to form soluble compounds. The liquid may also condense various substances, including water vapor, and collect fine particulate material that passes through a coarse filter or another separation device. A important advantage of an impinger compared to some bulk air sampling techniques is that hundreds of liters of air can be drawn through the device over a period of several hours. Sampling trains with multiple collection devices in series are used to collect different fractions of a vapor-phase sample. Components of sampling trains can include particulate filters of several types and sizes, particle-size separation devices, multiple impingers, and solid-phase adsorbents.

Solid-phase adsorbents are simple and inexpensive devices used to separate volatile analytes from the principal components of air in the field [1]. A porus solid-phase adsorbent is placed in a glass or metal tube which is taken to the field where air is drawn through the adsorbent to trap the analytes. The adsorbent tube is then sealed and returned to the laboratory for analysis. Many types of solid-phase adsorbents are used including alumina, activated carbons of various types, charcoal, graphitized carbon black (Carbopack™ B and C), carbon molecular sieves, GC packing materials such as Chromosorb™ 101 and 102, ethylvinylbenzene-divinylbenzene copolymer (Porapak Q™), styrene-ethylvinylbenzenedivinylbenzene terpolymer (Porapak P), silica gel, 2,6-diphenyl-*p*-phenyleneoxide (Tenax-GC™), styrene-divinylbenzene copolymers (XAD-1™, XAD-2™, and XAD-4™), and acrylic ester polymers (XAD-7™ and XAD8™). Adsorbent trapping is applicable to a wide variety of nonpolar and some polar volatile analytes.

Some adsorbents do not bind analytes strongly, which is an advantage for subsequent processing, but may result in the gradual release of some volatile analytes during long vapor sampling periods. Long sampling periods are often required, for example, with ambient air, because of generally lower analyte concentrations compared to vapor-phase samples from stationary or mobile sources of air pollutants. Breakthrough of the more volatile polar and nonpolar analytes from the adsorbent is controlled by placing additional adsorbents in the tube or additional sampling tubes in a sampling train. Analytes adsorbed on solid-phase materials may also react with substances in the air, for example, nitrogen oxides and ozone, to produce products not in the original sample. Long air sampling times also risk oxidation of some adsorbents and the production of background substances and other artifacts. These potential problems are usually evaluated during analytical method development and controlled by using appropriate quality control procedures.

Processing of Samples Before Determination of the Analytes

Some processing of samples before the determination of the analytes is often required to achieve optimum analytical method performance. Sample processing also defines exactly what elemental forms, compounds, ions, oxidation states, and physical phases are included in the determination of specific substances, or in the total amount of an element or a group of related substances in a sample. The degree of sample processing that is either appropriate or needed depends on a variety of considerations including:

The chemical analysis strategy, that is, a total target analyte, or broad spectrum determination
The physical phases of the sample that are included in the determination
Whether the samples will be analyzed in a remote or a field laboratory or with laboratory, mobile, or portable instrumentation
The complexity of the sample matrix
The number of target analytes
The diversity of properties of the target analytes
The nature and concentrations of analyte interferences
The detection limits required
The efficiency of the chromatographic separation used in the analytical method
The selectivity of the chromatographic detector or the nonchromatographic measurement technique
The willingness to risk severe contamination of a GC or HPLC column, or some other critical instrument component, by a concentrated or high background environmental sample

Some sample matrices, for example, ambient air or drinking water, are analyzed without sample processing with some analytical methods. Volatile compounds in ambient air collected in canisters or volatile compounds trapped from ambient air on solid-phase adsorbents are determined directly. Similarly, the total concentrations of some elements in low turbidity drinking or surface water samples preserved at a pH 2 can be determined without further sample processing. Common inorganic anions, for example, Cl, Br, NO_3, etc., are determined in surface, ground, drinking, and some other water samples without sample processing. However, many sample matrices and analytical methods require at least some, and sometimes considerable, sam- *CHROMATOGRAPHIC ANALYTICAL* ple processing prior to the determination of *METHODS* the analytes.

Semivolatile organic analytes trapped on air filters or on solid-phase adsorbents or in water, soil, and other solid samples are extracted with an organic solvent, or a solvent mixture, and the extracts are concentrated by evaporation of the solvent before the determination of the analytes. Solid samples are often extracted with the classical Soxhlet apparatus or with a variety of other techniques including several that use organic solvents at elevated temperatures and pressures. If a sample is highly concentrated, for example, an industrial wastewater or a soil from a hazardous waste land fill, fractionation of the solvent extract before determination of the analytes is usually required. Fractionation is often accomplished with open-column liquid chromatography or preparatory-scale HPLC. The determination of total elements in particulate matter collected on air filters, in most water samples, or in solids requires an acid digestion of the sample to ensure that the elements are in water-soluble chemical forms for the measurement. The vigor of the digestion also determines the recovery of the elements from sediment and other insoluble material.

Invariably the analytical method developer is required to make compromises between the amount and complexity of the sample processing and the separating power, selectivity, and other attributes of the chromatographic or nonchromatographic determination. These compromises are often strongly influenced by the projected cost and time required for various method options and by the desired quality, detail, and reliability of the results. Major issues usually are the availability of laboratory or field equipment and instrumentation, the experience and skill of the staff in using the equipment, and other laboratory or field infrastructure required to complete the analyses of the samples. Most research and standard analytical methods contain many compromises that may not be clearly defined in the method description, but should be understood by the user.

Nearly all contemporary analytical methods designed for specific compounds or ions in environmental samples employ some type of chromatographic separation. Gas chromatography is the dominant technique for the separation of volatile and semivolatile analytes (Figs. 7.1 and 7.2 and Table 7.1). Reverse-phase HPLC is used for the separation of nonionic analytes that are thermally unstable, nonvolatile, or reactive and not amenable to GC (Fig. 7.3). High-performance ion exchange chromatography and capillary electrophoresis (CE) are used for the separation of ionic substances. Analytes, either fully or partially separated, are sensed by in-line chromatographic detectors that produce electronic signals that are usually converted into digital form and stored in computer data systems. The computer systems, which are used with nearly all contemporary chemical analytical instrumentation, typically have software to control the operation of the instrument, acquire and store raw data, and reduce the data to more usable analytical information.

Chromatographic Separation Techniques

Gas Chromatography. The basic components of a gas chromatograph are a carrier gas system, a column, a column oven, a sample injector, and a detector. Very pure helium is the near-universal carrier gas for environmental and many other analyses. Open tubular GC columns are constructed of fused silica with low-bleed stationary phases of varying polarity chemically bonded to the silica surface. Columns are typically 30–75 m in length and have inside diameters (ID) in the range of about 0.25–0.75 mm. The column oven is capable of precise temperature control and temperature programming at variable rates for variable times.

Analytes are introduced into GC columns with several techniques. An aliquot of a relatively concentrated vapor or air sample, for example, from a plastic bag or a canister, can be introduced into a short section of tubing of known volume, called a sample loop, and subsequently purged with carrier gas into the GC column. Volatile analytes in ambient air samples in a canister or trapped on a solid-phase adsorbent are usually concentrated and focused in a cryogenic trap or a secondary adsorbent trap, then thermally vaporized into the GC carrier gas stream. However, in some analytical methods, volatiles trapped on an absorbent are thermally desorbed directly into the GC column. Aliquots of organic solvent extracts from various aqueous and solid samples are usually injected with a syringe into the carrier gas stream in a heated injection port. Both manual and automated syringe injection systems (autoinjectors) are used and the latter are generally very reliable, precise, and have the capacity to process many samples unattended.

Mixtures of analytes are separated by repeated equilibrations between the vapor state, where the analytes are entrained in the flowing carrier gas toward the detector, and the absorbed or dissolved state, where they are attracted to the stationary phase on the wall of the column by generally weak molecular forces. Analytes that tend to favor the absorbed state move more slowly through the column to the detector than do analytes that favor the vapor state. Complete separations of a few nanograms (10^9) or less of each of 40–50 or more analytes in 30 min or less is not uncommon.

HPLC. The basic components of a high-performance liquid chromatograph are a high-pressure mobile-phase delivery system, a metal column packed with fine particles containing the stationary phase, a sample injector, and a detector. A high-pressure pump is used to force the mobile-phase solvent or solvent mixture through the packed column. The term *high performance* is often used to distinguish this technique from open-column liquid chromatography conducted at atmospheric pressure with gravity flow of the mobile phase. Columns vary in diameter from 1 to 5 mm ID, or larger, and from a few cm in length to 30 cm or more. Column packings consist mainly of silica particles, usually 3–10 m in diameter, coated with low-bleed stationary phases of varying polarity chemically bonded to the silica surface. Organic polymers are also used as stationary phases in analytical separations. Automated or manual syringe injections of aliquots of liquid samples and known-volume sample loops are used for sample injection.

The dominant HPLC technique is the reverse-phase configuration in which a nonpolar or slightly polar stationary phase is used with a more polar mobile phase that is often water, methanol, acetonitrile, or mixtures of these solvents. The normal-phase configuration, which was developed before reverse phase, uses a more polar stationary phase and a nonpolar mobile phase, but it is not often used in contemporary HPLC. A vast array of organic compounds containing a variety of functional groups are retained on nonpolar or slightly polar stationary phases, and are sufficiently soluble in more polar mobile phases to give excellent separations. Analytes are retained on the column by an equilibrium process in which the dissolved molecule is alternately associated with the stationary phase, through weak noncovalent bonding interactions, and the mobile phase where it is transported toward the detector.

Isocratic elution is the use of a mobile phase that has a constant composition throughout the elution of analytes from the column. Gradient elution is the gradual changing of the mobile phase composition as the analytes elute from the column. Gradient elution is frequently used to enhance analyte resolution and shorten the time required for a separation. The pH of typical mobile phases can be adjusted over a reasonable range to improve resolution and selectivity. Various buffer substances, salts, and ion pair reagents can be used to control pH, ionic strength, and to facilitate the separation of some analytes. Reverse-phase HPLC is more complex than GC because of the number of operational parameters including column dimensions, column packings, mobile phases, gradient elution, and various mobile-phase additives.

High-Performance Ion Exchange Chromatography. Ion exchange is another form of HPLC that uses a stationary phase consisting of a cross-linked synthetic organic polymer, often called a resin, with SO_3H or NH_2 groups attached to phenyl or other aromatic rings on the polymer backbone. Mobile phases are usually water or water and a miscible organic solvent. In basic solution the SO_3H groups of a cation exchange resin are ionized and consist of $-SO_3$ groups and associated counter ions, for example, Na. Cationic analytes are retained on the cation exchange column by displacing the resin counter ions in an equilibrium process. The cationic analytes are subsequently eluted with a mobile phase containing a high concentration of counter ions or counter ions of a higher charge, for example, Ca. Isocratic and gradient elutions are used with gradients in both solvent composition and counter ion concentration. Mobile-phase pH is a very important operational parameter that can have a significant impact on the retention or elution of various ionizable analytes from a resin.

In an acid solution the amino groups of an anion exchange resin are protonated and exist as NH_3 groups and associated counter ions, for example, Cl. Analyte anions are retained on the column by displacing the resin counter ions in an equilibrium process. The anion analytes are subsequently eluted with a higher concentration of counter ions or some type of gradient elution. Anion exchange chromatography is widely used for the separation of inorganic anions and chelated metal anions in aqueous samples. With some natural or industrial water samples that contain high concentrations of ions such as Na, Ca, Mg, Cl, and SO_4, cation and anion analytes may not be retained on the column and this can result in poor analyte recoveries.

Other Chromatographic and Related Techniques. Supercritical fluid chromatography (SFC), capillary electrophoresis (CE), and several related separation techniques are occasionally used in environmental chemical determinations. The CE technique is very important in biochemistry and molecular biology because of the very high resolving power that can be achieved, the high speed of separations compared to HPLC, and its ability to separate charged species such as proteins, peptides, and deoxyribonucleic acid fragments. These techniques have considerable potential for environmental analyses and may emerge as very important techniques in the future.

Chromatography Detectors

Online detectors for GC, HPLC, and other chromatographic separation techniques are conveniently divided into three general classes:
Nonselective detectors
Selective detectors
Spectroscopic detectors

Nonselective detectors respond to most or all changes in the composition of the carrier gas or mobile phase and are capable of detecting nearly all entrained or dissolved analytes. The GC flame ionization and thermal conductivity detectors and the HPLC refractive index detector are examples of nonselective detectors. Nonselective detectors provide no direct information about the identity of the analytes except the time of arrival, the time of maximum concentration, and the time of return to pure carrier gas or mobile phase. Thus the peak shape of the analyte is usually well defined, the measurement of peak area or height can be precise, but the information needed to identify the analyte that caused the change is weak. Identifications are based on comparisons of measured retention times of separated analytes with retention times of expected analytes measured under the same chromatographic conditions. Although measured retention times can be very precise, especially with fused silica capillary GC columns and multiple internal standards, there is a high probability of coelution of two or more analytes in most environmental samples. This is because the peak capacity of a chromatogram, which is the number of analytes that can be fully separated, is limited. For most environmental samples the peak capacity is much smaller than the number of possible analytes that respond to a nonselective detector. Therefore, there is a significant risk of misidentifications or false positives.

Selective detectors respond to only certain classes of analytes and they are often used in environmental chemical determinations. The electron capture, photoionization, electrolytic conductivity, and flame photometric GC detectors are selective for limited groups of analytes. Selective HPLC detectors include electrochemical detectors and the single wavelength ultraviolet–visible (UV–VIS) absorption and fluorescence detectors. Preinjection or postcolumn online chemical reactions are used to convert nonresponsive analytes into derivatives that respond to a selective detector. The information produced by a selective detector about the nature of the analyte is improved, compared to the nonselective detector, but still is not strong because the most significant piece of information obtained is the retention time. Although the range of potential analytes that responds to a selective detector is much narrower than the range that responds to a nonselective detector, multiple potential analytes still can have the same retention time. The nature of the sample is a factor in considering whether multiple analytes could be present that may have the same retention time.

Spectroscopic detectors measure partial or complete energy absorption, energy emission, or mass spectra in real-time as analytes are separated on a chromatography column. Spectroscopic data provide the strongest evidence to support the identifications of analytes. However, depending on the spectroscopic technique, other method attributes such as sensitivity and peak area measurement accuracy may be reduced compared to some nonselective and selective detectors. The mass spectrometer and Fourier transform infrared spectrometer are examples of spectroscopic detectors used online with GC and HPLC. The diode array detector, which can measure the UV–VIS spectra of eluting analytes is a selective spectroscopic detector because only some analytes absorb in the UV–VIS region of the spectrum. The mass spectrometer, which can use a variety of ionization techniques, is probably the most widely used GC and HPLC detector for environmental chemical determinations [1]. All detectors respond, to some extent, to natural background substances in environmental samples and sample extracts. Background chemical noise is highest in the most contaminated environmental samples and at the highest instrument sensitivities, but spectroscopic detectors are best equipped to distinguish chemical noise from environmental analytes.

Nonchromatographic Analytical Methods

Analytical methods that do not employ a chromatographic separation of analytes are widely used for some environmental determinations. These methods generally depend on highly selective physical measurements or selective chemical reactions. Atomic absorption spectrometry (AAS), atomic emission spectrometry (AES), and mass spectrometry (MS) are most often used for determinations of the elements, but electrochemical and colorimetric techniques are sometimes used. Derivatives of analytes, which are sometimes called complexes, that strongly absorb in the UV–VIS spectrum are employed in colorimetric methods for the elements and some compounds and multi-element ions. Ozone, nitrogen oxides, and some other substances in air are determined by selective chemical reactions that produce measurable light emission (chemiluminescence). Fluorescence spectrometry is used in some methods, especially with fiber optics technology, for remote monitoring. Immunoassays have been developed for a few analytes and they are especially useful in field analyses.

Elemental Analysis

Atomic emission and atomic absorption techniques are highly selective because the number of elemental analytes is small and each has a unique atomic spectrum. Frequencies are selected for measurements to maximize selectivity and sensitivity, minimize interferences, and correct for interferences when necessary. Mass spectrometry is the separation and measurement of ions in the gas phase by their mass-to-charge ratios. Elemental analyses by mass spectrometry provides high selectivity, sensitivity, and the ability to correct for interferences when necessary [1]. Colorimetric and electrochemical techniques were used more frequently for elemental determinations before the development of the three major spectroscopic techniques. Their use has decreased significantly because of inherent limitations in selectivity, sensitivity, susceptibility to interferences, and the need for chemical processing to prepare energy-absorbing derivatives.

Atomic Absorption. During the 1960s, flame AAS became the dominant technique for determinations of the elements in low concentrations in water samples and aqueous extracts of other samples. Aqueous sample aerosols are injected directly into the flame and precise determinations of many elements can be made. However, instrument detection limits for some important elements, for example, As, Cr, Pb, Sb, Se, and Tl, in water and other matrices are 50 g/L, and often much greater, and AAS is limited to measuring one element at a time. The high-temperature graphite furnace sample introduction system provided instrument detection limits of 5 g/L for most elements, and allowed automation of sample processing, but was still limited to measuring one element at a time. Other AAS sample introduction techniques, especially chemical reduction and elemental Hg vaporization (cold-vapor) and conversion of As and Se to volatile hydrides, provided similarly low detection limits. The graphite furnace and Hg cold-vapor techniques are often referred to in the scientific literature as flameless AAS methods.

Atomic Emission. Elemental analysis with AES had been practiced since the mid-1930s using flames and arc or spark discharges to vaporize and atomize samples and excite the atoms for optical emission. The AES technique provided rapid simultaneous or sequential multi-element determinations, but the flames and arc or spark discharges had significant limitations. Interferences from electrode and other sample components were not uncommon, elemental measurements were often imprecise, detection limits were not sufficiently low, and liquid samples were difficult to analyze. The increasing demand for rapid, selective, and sensitive multi-element determinations led to the development of the inductively coupled plasma (ICP) sample vaporization, atomization, and excitation source which met the analytical requirements and gave new life to AES. The ICP allowed the direct injection of aqueous aerosols into a 5,500–8,000°K argon ion plasma and elemental measurements with minimal or no interferences from background components or other analytes. Analytical methods utilizing ICP/AES are widely used for the determination of multiple elements in environmental samples.

Mass Spectrometry. The ICP technique also produces gas-phase elemental ions and the ICP was quickly adapted as an ion source for mass spectrometry. This combination became one of the most useful and important techniques for rapid multi-element analyses of gases, liquids, and solids [1]. The argon ICP is a very efficient ion source that produces mainly singly charged ions. It is estimated that 54 elements, all metals, are ionized with 90% or greater efficiency. Only C, H, N, a few electronegative elements, and the noble gases are ionized with efficiencies less than 10%. A small number of elements, for example, As, B, Be, Hg, I, P, S, Se, and Te have estimated ionization efficiencies in the 10–90% range. Although Ar has a low ionization efficiency, it is present in the ion source in great quantity and gives a few significant ions at m/z 40 ($Ar^{\cdot}$) and m/z 80 ($Ar_2^{\cdot}$). Other background ions, for example, m/z 41 (ArH) and m/z 56 ($ArO^{\cdot}$), are formed from sample components, usually water, and mineral acids that are used to ensure dissolution of some analytes.

These ions obscure ions from some elements, but techniques are available to circumvent most of these interferences.

Organic and Inorganic Compounds and Ions

Colorimetric determinations are selective for specific substances, for example, the PO_4^{-3} ion, because of selective color-forming chemical reactions between reagents and target analytes. Several of these techniques continue to be important for the determination of a few inorganic anions. However, ion-exchange chromatographic techniques are increasingly used for determinations of multiple ions in environmental samples. Some electrochemical techniques, for example, ion-selective electrodes, are selective because the materials of construction, for example, ion-selective membranes, and operating parameters are carefully chosen so the devices respond only to specific analytes. Ion-selective electrodes are widely used for the determination of pH and a few other inorganic ions, for example, F^-. Immunoassay methods, which are very widely used and enormously important in clinical analyses, have been developed for a small number of organic compounds, for example, several triazine pesticides. However, many chemicals of environmental interest have MWs and shapes that are too small and similar for effective application of immunochemical techniques.

The further development of nonchromatographic analytical methods based on these and other techniques has been impeded by fundamental limitations in selectivity, and sometimes sensitivity, for a wide variety of similar substances, for example, those in Table 7.1 and Figs. 7.1–7.3. The cost of developing, testing, and documenting a large number of specialized methods, often for just one or a few analytes, is generally prohibitive compared to the cost of chromatographic methods that allow the determination of multiple generally similar analytes in an environmental sample.

Glossary

Accuracy The degree of agreement between the measured concentration of a substance in a sample and the true value of the concentration in the sample.

Analysis The process of investigation of a sample of the physical world to learn about its chemical components, composition, structure, or other physical or chemical characteristics. Generally only samples are analyzed and individual elements, compounds, and ions are separated from one another, identified, measured, or determined. A pure compound or multi-element ion is analyzed only when it is investigated to determine its components, composition, structure, or other physical or chemical characteristics.

Analyte A general term for any element, compound, or ion that is present in a sample or is targeted for determination in a sample.

Analytical method The complete process used to determine an analyte or analytes in a sample. The analytical method documents all the individual steps in the process from sampling to reporting the results.

Analytical method attributes Measures of the quality, reliability, and uncertainty of the determinations obtained with an analytical method. Typical analytical method attributes are selectivity, sensitivity, detection limits, signal/noise, recovery, accuracy, bias, precision, and validation. *Analytical method attributes* are sometimes called *figures of merit*.

Bias The systematic error in a measurement of the amount or concentration of an analyte in a sample.

Congeners Compounds or ions that are members of a series of related substances that differ only by the number of hydrogens that have been substituted by the same atom.

Derivatives Compounds or ions that are produced by chemical reactions of analytes. An analytically useful derivative has physical or chemical properties that are not possessed by the analyte but that can be employed to determine the analyte.

Detection limit The minimum quantity or concentration of an analyte that can be detected with an analytical method or technique. There are no generally accepted standard criteria for detection and detection limits often depend on the sample matrix. Therefore detection limits must include the criteria for detection and the nature of the sample matrix. A technique with a higher analyte sensitivity does not always provide a lower detection limit because interfering sample matrix components may also be observed with higher sensitivities.

Determination The identification and the measurement of the concentration of an analyte in a sample.

Isomers Two or more compounds or multi-element ions that have the same elemental composition but different structures.

Precision The degree of random variation in repetitive measurements of the concentration of an analyte in a sample. Precision is usually measured by the standard deviation or the relative standard deviation of the measurements.

Procedure A specific part of an analytical method that is concerned with one aspect of the method, for example, the liquid–liquid extraction of groups of similar analytes from a water sample.

Qualitative analysis The process of only identifying the analytes in a sample.

Quantitative analysis The process of both identifying and measuring the concentrations of the analytes in a sample.

Recovery The amount of analyte measured in a sample matrix as a fraction of the amount of the same analyte that was added to the sample. If the analyte is present in the sample before the addition, the native amount is subtracted from the measured quantity before calculating the recovery.

Sample matrix The general nature of the sample and its components that can have a significant impact on the performance of an analytical method. For example, sea water and fresh water sample matrices are significantly different and this difference can affect the performance of an analytical method.

Selectivity Is a qualitative estimate of how well the analyte identification procedure is able to distinguish an analyte in a sample from one or many similar analytes with similar, or even some of the same, physical or chemical properties.

Sensitivity Is the electronic or other measurable signal produced by the analytical method or measurement technique per unit amount of analyte.

Signal/Noise Is the ratio of analyte electronic or other measurable signal to the mean background matrix signal.

Technique Is a specific way of manipulating a sample or substance or measuring a substance. One or more techniques may be used within each procedure and several procedures may be used within an analytical method.

Validation Is the determination of the attributes, or figures of merit, of an analytical method for one or more analytes in one or more sample matrices by one or more analysts in one or more analytical laboratories and the acceptance of the attributes as reasonable and useful by the users of the data. There are many levels of analytical method validation ranging from the validation of a method for a single analyte in a single matrix by a single analyst in a single laboratory to a multi-analyte, multi-matrix, multi-analyst, and multi-laboratory validation.

References

1. Budde WL (2001) Analytical mass spectrometry: strategies for environmental and related applications. Oxford University Press, New York
2. http://www.epa.gov/nerlcwww/methmans.html
3. Clesceri LS, Greenberg AE, Eaton AD (eds) (1998) Standard methods for the examination of water and wastewater, 20th edn. American Public Health Association, Washington, DC
4. Annual Book of ASTM Standards, Section 11, Water and Environmental Technology. American Society for Testing and Materials, West Conshohocken, PA
5. Standard D5466-95, Standard Test Method for Determination of Volatile Organic Chemicals in Atmospheres (Canister Sampling Methodology), Annual Book of ASTM Standards, vol 11.03. American Society for Testing and Materials, W. Conshohocken, PA; 1998 and later editions
6. Parker LV (1994) The effects of ground water sampling devices on water quality: a literature review. Ground Water Monitoring Remediation 14:130–141

Suggested Additional Reading

Bloemen HJ Th, Burn J (eds) (1993) Chemistry and analysis of volatile organic compounds in the environment. Blackie Academic & Professional, London

Keith LH (ed) (1996) Principles of environmental sampling, 2nd edn. American Chemical Society, Washington, DC

Schwarzenbach RP, Gschwend PM, Imboden DM (1993) Environmental organic chemistry. Wiley, New York

Taylor JK (1987) Quality assurance of chemical measurements. Lewis, Chelsea, MI

Nanotechnology: Fundamental Principles and Applications

8

Ranjit T. Koodali and Kenneth J. Klabunde

Introduction

Research in nanotechnology has seen an explosive growth during the past decade, fueled by significant investment in research and development and the appearance of a cornucopia of consumer-based products. As of August 2009, an inventory conducted by the Project on Emerging Nanotechnologies has identified over 1,000 nanotechnology consumer products [1]. Based on current growth rates, it is estimated that nano-related goods could be in the range of $1–2.5 trillion market by 2015 globally.

This chapter is not intended to be a comprehensive summary of nanotechnology; however, the basic principles and selected applications will be discussed for the reader to gain an appreciation and basic understanding of the importance of nanotechnology and how it has already made an impact in our day-to-day lives. The reader is also advised to refer to some excellent books published in the past few years that could provide more detailed and specific information [2–20].

What Is Nano?

The prefix *nano* literally means one billionth of a meter and one nanometer is thought to be a “magical point on the dimensional scale” [18]. This is because these are the smallest solid materials that can be fabricated since anything smaller than nanometer would essentially be a small molecule or an atom and not condensed matter. Thus, nanostructures (materials with one dimension in the range between 1 and 100 nm) are the smallest solid things possible to make. To put things in context, ten hydrogen atoms make up one nanometer. A better appreciation of the nanoscale may be obtained from Fig. 8.1 which contains images of manmade and natural substances spanning several length scales. For example, the width of a human hair is between 60 and 120 μm.

The corresponding manmade devices are microelectromechanical (MEMS). MEMS is a mature field of nanotechnology and finds extensive uses in manufacturing and assembly processes. MEMS deals with devices primarily in the 10–100 μm range. Advances in the fields of nanotechnology and MEMS have led to the creation of Nanoelectromechanical Systems (NEMS), a field that bridges and borrows principles from both nanotechnology and MEMS.

Nanotechnology is a broad encompassing area incorporating principles of sciences, engineering and technology to understand, control, manipulate, and construct matter in the 1–100 nm dimension range. A widely accepted definition is the one proposed by the National Nanotechnology Initiative (NNI), a Federal research and development program approved by the Congress in 2001. This defines nanotechnology only if all three following aspects are involved: (1) research and technology development at the atomic, molecular, or macromolecular levels, in the length scale of approximately 1–100 nm range, (2) creating and using structures, devices, and systems that have novel properties and functions because of their small and/or intermediate size, and (3) ability to control or manipulate on the atomic scale. Thus, one can define nanotechnology to be a technology based on the manipulation of individual atoms and molecules to build complex structures that have unique physical, chemical, and/or biological properties different from bulk properties. Estimates based on key indicators such as publications, patent applications, products, and research and development funding suggest that the average growth in this area has been 23–35% since 2000. In 2008 alone, over 15,000 manuscripts in the broad area of

R.T. Koodali (✉)
Department of Chemistry, Churchill and Haines Laboratory, University of South Dakota, 414 Clark Street, Vermillion, SD 57069, USA
e-mail: Ranjit.Koodali@usd.edu

K.J. Klabunde
Department of Chemistry, Kansas State University, Manhattan, 66506, KS
e-mail: kenjk@k-state.edu

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_8,

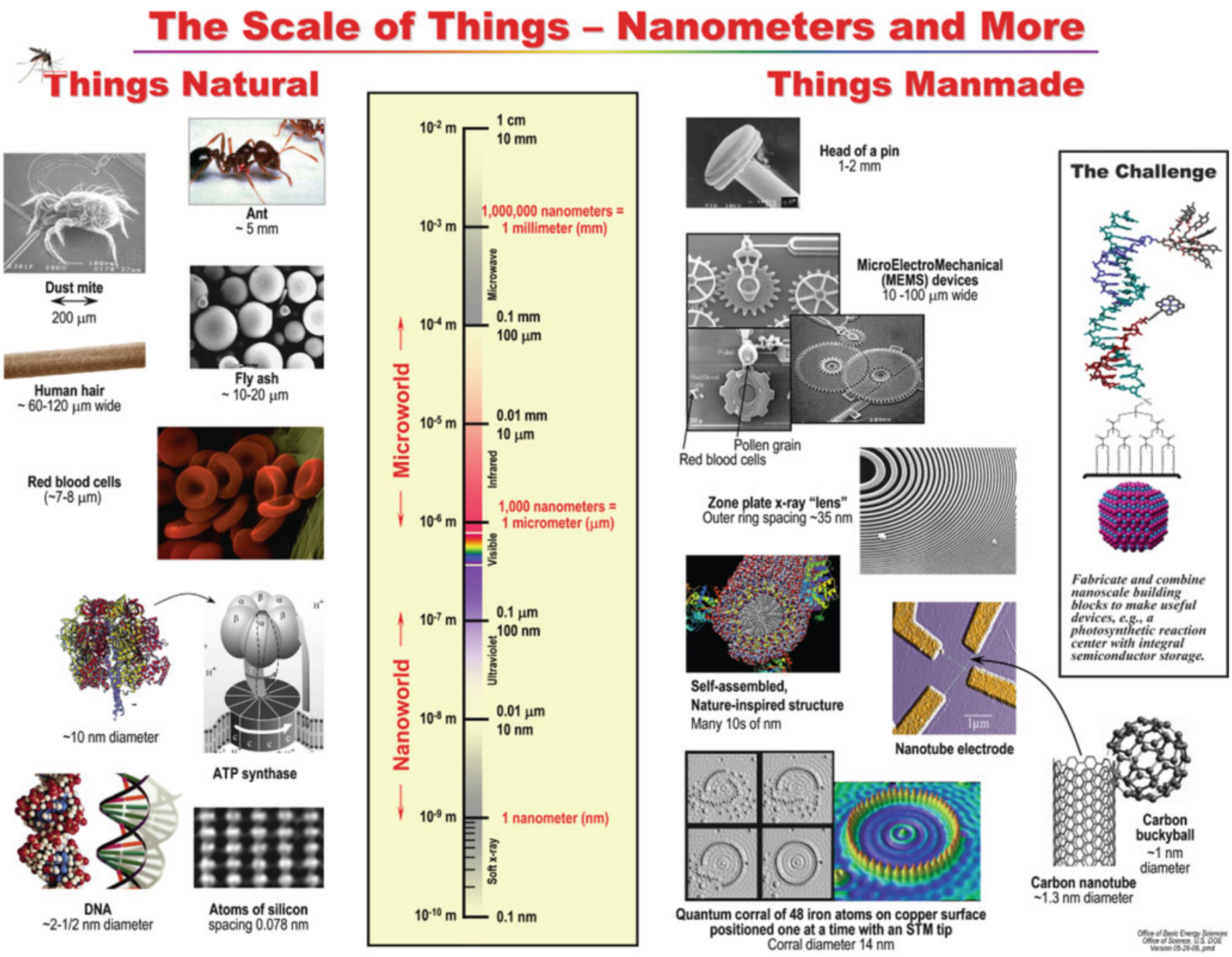


Fig. 8.1 Scale of things, Office of Basic Energy Sciences, Office of Science, Department of Energy (DOE)

nanoscience/nanotechnology have been published, compared with ~5,000 in the year 2000, illustrating robust growth in this fertile area of research.

Nanotechnology is a technology that is expected to create products or devices that are lighter, robust, safer, reliable, and durable. At the core of nanotechnology is what is called "molecular technology" or "molecular manufacturing." Molecular technology refers to manufacturing processes using molecular machinery, i.e., obtaining molecule-by-molecule control of products and by-products via positional chemical synthesis. Most of the current technologies rely on bulk technology or a top-down approach such as the ancient technology of chipping a large rock to smaller particles or the modern technology of making Silicon chips for microprocessors. In contrast, nanotechnology builds materials and products from the bottom-up. Nanotechnology handles individual atoms or molecules with control and precision.

History

The term *nano-technology* was coined by the Japanese scientist Norio Taniguchi in 1974 and then by Eric Drexler (who was unaware of the usage of the term by Taniguchi) who popularized the concept of nanotechnology [21, 22]. However, the dawn of nanotechnology was envisioned by the late Nobel laureate Richard Feynman much earlier [23]. In his seminal talk at the annual meeting of the American Physical Society in December 1959, titled *There's plenty of room at the bottom—An Invitation to Enter a New Field of Physics*, Feynman discussed the advantages that could be obtained through precise control of atoms. Feynman, had stated in his talk "*But I am not afraid to consider the final question as to whether, ultimately—in the great future—we can arrange the atoms the way we want; the very atoms, all the way down.*" There is no doubt that the seeds for the foundation of nanotechnology was laid by Feynman more than 50 years ago.

The far-reaching idea proposed by Feynman in 1959 did not immediately spur research in this area and it took time before scientists had the ability to synthesize new materials and image them by microscopic tools. In 1985, a team of scientists at Rice University, Houston, USA discovered fullerenes (Buckyballs), a new allotrope of carbon in which all the C atoms were arranged in closed shells. The American Chemical Society (ACS) honored the discovery of fullerenes as a National Historic Chemical Landmark in 2010. A book written by Eric Drexler in 1986 titled *Engines of creation—The coming era of nanotechnology* created considerable excitement in the scientific community [22]. In his book, Drexler had stated that "*Our ability to arrange atoms lies at the foundation of technology and for better or for worse; the greatest technological breakthrough in history is still to come*." The publication of this book sparked imagination in the minds of scientists. Coupled with technological progress in instrumentation, research in the area of nanotechnology attracted the attention of several scientists. Throughout the late 1980s and early 1990s, research in the area of nanoscience and nanotechnology blossomed and scientists were able to synthesize novel materials. In September 1989, IBM Fellow, Don Eigler realized the vision of Feynman, by being able to move and control individual atoms. In November of the same year, the research team at IBM showed that they could use a Scanning Tunneling Microscope (STM) tip to move atoms, and they spelled out the letters "IBM" using 35 individual Xenon atoms on a nickel surface. This pioneering work opened new vistas and heralded the coming of nanotechnology. This was made possible by the invention of the STM and the Atomic Force Microscope (AFM) in 1986 and 1989, respectively. In 1991, Iijima reported the synthesis of carbon nanotubes, cylindrical structures that have novel properties such as extraordinary strength, efficient thermal conductivity, and high electrical conductivity.

In the United States in 1998, several federal agencies formed a group called "Inter-agency Working Group on Nanotechnology (IWGN)" to discuss future plans in the area of nanotechnology and sponsor research funding in this area. IWGN was the forerunner to the current NNI established in 2001 to foresee research and development in nanotechnology and to forecast future directions. The federal government initiative includes activities ranging from basic and fundamental nanoscience to the development of specific nanotechnology devices and applications. They include the design and manufacture of nanostructured materials that are correct and precise at the atomic or molecule level. These advances are aimed at applications such as cost-effective manufacture of nanoscale microelectronics, more efficient and cost-effective energy conservation and storage devices, and biological sensors with applications to both health care and chemical and biological threat detection. In 2003, the US Congress passed the Twenty-first Century Nanotechnology Research and Development Act which provided statutory funding for NNI. In pursuance to the environmental and safety concerns, the Environmental Protection Agency (EPA) issued guidelines in 2009 to better understand the impact of manufactured nanomaterials.

Nanotechnology has now become a mature field with the appearance of over 1,000 verified products that incorporate some element or aspect of nanotechnology. The various product categories containing nanotechnology are health and fitness, computers and electronics, home and garden, food and beverage, and consumer appliances. The health and fitness accounts for most of the products and it includes cosmetics (137 products), clothing (155), personal care (193), sporting goods (93), sunscreens (33), and filtration units (43).

A New Realm of Matter

Chemistry is the study of atoms and molecules whose dimensions are generally less than one nanometer (e.g., a carbon-carbon single bond length is only 0.154 nm), while the majority of physics, particularly the area of condensed matter, deals with solids that are essentially an infinite array of bound atoms or molecules greater than 100 nm. A considerable gap exists between these two extreme regimes. The regime called the nanoscale regime lies intermediate between the realms of quantum chemistry and solid state or condensed matter physics and deals with particle sizes of 1–100 nm. In this regime, classical laws of physics or quantum chemistry rules are not strictly applicable. Hence, this regime represents a new realm of matter in which the size scale creates/imparts unique physicochemical properties [6]. For example, the melting point of gold (Au) nanoparticle is lower than the melting point of the bulk Au metal and does not necessarily vary linearly with particle size. Also, a nanoscale wire may not necessarily obey Ohm's law that is the foundation and cornerstone of the electronics industry. Thus, the nanoscale size imparts unique and often unpredictable physical, chemical, and/or biological properties distinct from the bulk.

The field of nanomaterials is not unique to manmade substances; nature has been utilizing nanomaterials for millions of years. Several biological systems contain several nanoscale materials. For example, bones, teeth, and shells are molecular composites of proteins and biominerals that have superior strength and toughness. Human bones contain minerals with particle sizes in the nanoscale regime. Another example is found in certain aquatic bacteria that are able to orient themselves using the Earth's magnetic field. This is possible because they contain chains of nanosized, single-domain magnetite (Fe_3O_4) particles. Another example is the teeth of herbivorous mollusks. Their teeth have an intricate structure that contains nanocrystalline needles of goethite (FeO(OH)). It is quite a marvel that nature produces such tough materials out of protein constituents as soft as human skin and mineral constituents as brittle as a classroom chalk. In fact, nature can be considered as providing model systems for developing technologically useful nanomaterials. According to Dickson, *Life itself could be considered as a nanophase system!* [24]

Synthesis

In this section, we will briefly examine the various synthetic methods available for the preparation of nanomaterials. Our focus in this section will be on metals (Gold and Silver) and metal oxides (Magnesium oxide and Calcium oxide), the two important classes of materials that have received considerable attention because of their widespread applications.

Several innovative and "bottom-up" approaches have been developed to prepare nanomaterials with desired physical and chemical properties. In this chapter, we shall discuss two methods that have been developed and successfully commercialized in the author's laboratory for the production of nanoparticles.

The first method, called Solvated Metal Atom Dispersion (SMAD) method, allows the preparation of gram scale and higher quantities of metal nanoparticles such as Au, Ag, Cu, Pd, etc., metal supported catalysts, bimetallic alloy materials such as Au-Ag, Au-Cu, Pt-Sn, Ag-Au, Mn-Co, etc., and also semiconductors such as CdSe, CdS, PbS, etc., that have uniform particle sizes in nanoscale dimensions. A further important aspect of the SMAD method is that it leads to the production of high purity materials (with no formation of by-products) and most importantly the materials produced are highly monodisperse (uniform particle sizes).

Solvated Metal Atom Dispersion Method for the Preparation of Nanoparticles

Before we discuss the SMAD method, we will briefly discuss other methods for the preparation of nanoparticles with emphasis on Au nanoparticles. A popular method of preparation of Au nanoparticle was first reported by Turkevich in 1951. This method relied on the use of a reducing agent, citrate ions, to reduce Au^{3+} ions in aqueous solution. The original recipe reported by Turkevich led to the synthesis of Au nanoparticles that were spherical in shape with a particle diameter of ~20 nm [25]. An improvement in the synthetic method was reported in 1973 in which Au nanoparticles with varying particle sizes (16–147 nm) could be prepared by careful control of the experimental conditions. A new method was developed by Brust and Schiffrin in which a thiol containing compound was simultaneously added to the citrate and gold ions [26]. This work is considered to be important in the development of Au nanoparticles since the synthesis was facile and led to the formation of Au nanoparticles that were thermally stable and considerably resistant to oxidation by air. The Brust-Schiffrin method led to the formation of Au nanoparticles ranging in diameter from 1.5 to 5.2 nm. Another key feature of this synthetic method was the ability to isolate and re-disperse the Au nanoparticles in common organic solvents without decomposition or irreversible aggregation of the nanoparticles. Although this work led to the development of other stable Au nanoparticles, a drawback of this method was the inability to prepare them in large amounts, i.e., gram *or* kilogram scale. It is usually challenging to prepare gram scale amounts of pure nanomaterials that are monodisperse.

The importance in the preparation of metal nanoparticles lies in the fact that metal supported catalysts are among the important class of synthetic materials developed in the past 50 years. Approximately, 20 % of the US national gross product is dependent on the use of these catalysts. Thus, the preparation of metal-supported catalysts is very vital to the economy of any nation. It is worthwhile to briefly review the traditional methods of preparation of metal-supported catalysts.

Metal ion reduction has been the principal procedure for preparing metal-supported catalysts. In this method, metal nitrates or metal halides are dissolved in water and added to a high surface area solid support such as alumina or silica. Then, water is evaporated and the dispersed metal salt solution is converted into the oxide form by drying the slurry in air. Finally, the metal oxide is reduced to the metallic state by passing a stream of hydrogen gas to give the final metal-supported catalyst. During the reduction step, the reduced metal atoms sinter to form larger particles leading to difficulties in controlling their particle sizes. The advantages of this method of preparation are the ease of preparation, easy scale-up for industrial applications, and the relative low cost involved. However, there are several disadvantages: (1) incomplete reduction of the metal ions to the metallic state, (2) sintering of the metal atoms during the high temperature treatment leading to broad particle size distribution, and (3) difficulty in reducing two metal particles simultaneously to prepare bimetallic particles supported oxide catalysts [27, 28].

A second popular method of preparation of metal supported catalysts is decomposition of organometallic compounds on catalyst supports. In this procedure, stable organometallic compounds are adsorbed from either solution or gas phase onto a catalyst support. Controlled thermolysis of the organometallic compounds leads to loss of the organic ligand molecules and the metal particles grow on the support surface. In principle, by using an organometallic cluster compound, it is possible to form a well-defined cluster of metal atoms on the support surface; however in practice, the removal of the ligands in a well-organized manner to deposit metal nanoparticles with pre-determined particle sizes that are highly uniform has been a challenge. Cluster aggregation and decomposition are frequently encountered and are the major disadvantages of this approach. A related method that has been employed is the use of plasma discharge to decompose the organometallic compound in a controlled manner. However, this method suffers from the disadvantage that only a limited number of organometallic compounds can be utilized for the preparation of metal supported catalysts; this severely limits the range of nanomaterials that can be synthesized.

A novel approach was developed in 1976 and makes use of weakly stabilized metal atoms or “solvated metal atoms” [29]. The nature of the solvent and the metal determine the stability of the metal solvate. Such metal solvates are normally thermally unstable and hence can be used as precursors to metal nanoparticles. On warming the metal solvates, metal cluster nucleation takes place and in the presence of a support, both nucleation and particle growth occur that lead to the formation of a highly dispersed metal-supported heterogeneous catalyst. This preparation method has been termed by the authors as “Solvated Metal Atom Dispersed” (SMAD) procedure. The numerous advantages of this method are: (1) unusual metal morphologies that are highly reactive are often encountered since the metal particle deposition occurs at low temperature, (2) very high dispersion of the zero valent metal is achieved, (3) no reduction step is necessary and sintering or agglomeration of metal particles is avoided, (4) many metal–solvent combinations can be used, (5) bimetallic metal or alloy particles can be prepared, (6) metal particles are securely anchored to catalyst support by the reaction of solvated atoms with surface hydroxyl groups, and (7) handling of toxic organometallic compounds can be avoided. However, careful and skillful operation of the experimental setup is required, and although a number of metal–solvent combinations can be used, only some of the combinations have been found to be of synthetic utility.

SMAD catalysts have been generally prepared in the following manner. An aluminum oxide coated tungsten crucible is first degassed by placing it into a reaction vessel and heating to ~1,400°C in vacuum overnight. Then, after cooling the crucible the reactor is filled with air and a gram scale quantity of metal is placed in the crucible. About 100 mL of organic solvent (usually a non-polar solvent such as toluene or pentane) is degassed and placed in Schlenk tubes and attached to the SMAD reactor. This whole system is sealed under vacuum and the reactor flask is immersed in liquid nitrogen and evacuated for few hours (typically at least 3 h) reaching final pressures of nearly 1×10^{-3} Torr. After approximately 30 min, the solvent is deposited on the cold walls of the reactor flask forming a layer of frozen solvent. At the same time, the crucible is slowly heated to the evaporation temperature of the metal. The metal vapor and solvent are co-deposited for a period of 3 h. After the deposition of the solvent layer and the evaporation of the metal are completed, the frozen matrix is warmed from $-196°$C to room temperature until the solvent melts to form a solvent-solvated metal solution. The solution is stirred and warmed further in the presence of a support (silica or alumina, ~20 g preheated at 500°C for 3 h in dry air and cooled and transferred to the reaction chamber and placed in vacuum prior to the metal vapor reaction). After reaching room temperature, the mixture is siphoned in vacuum and placed under an atmosphere of nitrogen in glassware to prevent any oxidation of the metal. The solvent is then slowly removed under vacuum and the resulting powder is then outgassed at room temperature to 1×10^{-5} Torr and used without any further treatment as catalyst. The metal concentration, temperature, and time of impregnation have an effect on the final size and shape of the nanoparticle. Support effects are not pronounced, but the presence of acid sites and higher specific surface areas of the support tend to allow better dispersion of the metal particles. The following section describes some recent results obtained for the preparation of Au, Ag, and semiconductor nanoparticles (CdS, PbS, CdSe, etc.) using the SMAD method.

Gold Nanoparticles

Gold has received the attention of mankind since ancient times and continues to be an important metal. Au nanoparticles or colloids had been inadvertently incorporated into glass several centuries ago. The Lycurgus cup is a classic and outstanding work of Roman technology belonging to the fourth century. The cup shows a spectacular color change from green to red depending on whether light is reflected or transmitted. This dichroic property has been attributed to the presence of both Au and Silver (Ag) nanoparticles in trace amounts that had been inadvertently introduced into the glass during the manufacturing process by the Romans. Transmission electron microscopic (TEM) studies conducted in recent years indicate the presence of 50–100 nm of Au and Ag nanoparticles, key to producing the unusual optical effect. Thus, the process of making metallic nanoparticles is not entirely new. The stained glass windows found in many churches built during the medieval period is attributed to the presence of gold nanoparticles. Thus, the first nanotechnologists were glass workers, although they were not fully sure what caused the various colors.

An interesting feature is the arrangement of ligand-stabilized, monodisperse particles into regular and periodic two- or three-dimensional structures called nanocrystalline superlattices (NCSs). The formation of such two-dimensional (2D) or three-dimensional (3D) NCS leads to the formation of materials characterized by very different properties compared to those of the discrete species. The collective interactions of the individual nanoparticles in the NCS and the individual particles themselves lead to the manifestation of unique properties compared to the bulk particles [30].

Several methods have been reported for the synthesis of 2D and 3D NCSs [31]. The most common method is to reduce the metal salt in the presence of stabilizing agents [32]. However, they invariably end up in the formation of by-products and an additional step is required for the isolation of samples. Thus, the traditional methods of preparation are quite tedious and time consuming and the possibility of scale-up is fraught with more difficulties and challenges. An important advantage of the SMAD method is the possibility of scaling up the process and the ability to coat substrates uniformly with continuous electrically conductive metals [33]. Industrial applications of nanoparticles require their preparation in large quantities and reproducible quality. Thus, the SMAD method seems to be a very facile method for the preparation of large-scale quantities of metal nanoparticles.

Colloidal solutions of gold in acetone have been well documented in the literature. A combination of two solvents such as acetone and toluene was used in the preparation of Au nanoparticles by the SMAD method [34]. Acetone, being a polar solvent, solvates the nascent Au metal clusters during the warming stage of the SMAD method and acts as a stabilizing agent. After the removal of acetone from the system, the Au nanoparticles are stabilized by dodecanethiol molecules, which enable their dispersion in non-polar solvents such as toluene. A second step called the digestive ripening is carried out in the presence of the surface active agent and a solvent (typically toluene) [35, 36]. This step is the key for the formation of highly monodisperse Au nanoparticles from the polydisperse Au–toluene–thiol colloid [37, 38]. In this step, the Au–toluene–thiol colloid is heated to reflux under Ar atmosphere for a couple of hours. This leads to a narrowing of the particle size distribution and monodisperse colloids are obtained. This particle narrowing is caused by the transfer of gold atoms/clusters from particle to particle until a thermodynamic equilibrium is achieved between the formation of more gold–sulfur bonds as shown by the following equations:

$$\text{Au vapor} + \text{acetone} \xrightarrow{77\text{K}} (\text{Au})(\text{acetone})\ \text{solvate} \tag{8.1}$$

$$\text{Au(acetone)} + \text{C}_{12}\text{H}_{25}\text{SH} \rightarrow (\text{Au})_x(\text{HSC}_{12}\text{H}_{25})_y \tag{8.2}$$

$$(\text{Au})_x(\text{HSC}_{12}\text{H}_{25})_y \xrightarrow[\text{reflux}]{\text{toluene}} (\text{Au})_{5000}(\text{HSC}_{12}\text{H}_{25})_{300} \tag{8.3}$$

Although, the exact mechanism of the digestive ripening is not fully clear at this moment, the process probably involves dissolution of surface atoms or clusters of atoms by the ligand (dodecanethiol) molecules [39]. Hence, it is hypothesized that reactive sites such as corners and edges (since these are the most energetic ones) are probably the first that are prone to dissolution-reprecipitation process. The dry product (usually obtained as a fine powder) is then obtained by cooling the colloidal solution to room temperature and precipitating by adding absolute ethanol. The final dry product consists of Au nanoparticles stabilized by dodecanethiol molecules that are readily soluble in non-polar solvents such as toluene or hexane. Treatment with a long-chained thiol in refluxing toluene causes a remarkable particle size narrowing such that each final gold nanoparticle is 4.6 ± 0.1 nm (about 5,000 atoms/particle and 300 $\text{C}_{12}\text{H}_{25}\text{SH}$ ligands/particle). A remarkable property of these colloidal solutions is their tendency to organize into 2D and 3D NCSs [34]. The TEM figures of the Au colloids obtained after digestive ripening and cooling to room temperature show remarkable features. After just 15 min, these Au nanoparticles organize and form predominantly 3D NCSs and some 2D NCSs when deposited on a support. Interestingly these NCSs are stable and even after 2 months, distinct 2D structures are seen.

High-resolution TEM studies indicate that the Au nanoparticles prepared by the SMAD procedure lack faceting, have defective core structures, and behave like “hard spheres.” They self-assemble and organize into hcp-type superlattices with long-range translational order. The Au nanoparticles have predominantly spherical shapes and the TEM of the Au nanoparticles showed reproducibly hcp structures indicating the versatility of the SMAD method [40]. An interesting observation from this study was the fact that Atomic Force Microscopy (AFM) studies conducted on Au nanoparticles deposited on surfaces indicated an fcc-type superlattice packing indicating that the superlattices can undergo reconstruction when deposited onto surfaces. An AFM image of Au superlattice surface is shown in Fig. 8.2.

Silver (Ag) Nanoparticles

Silver nanoparticles are of great importance due to their ability to efficiently interact with light because of plasmon resonances [15]. These are due to the collective oscillations of the conducting electrons in the metal [41]. Ag nanoparticles are envisaged to be vital components of optical and photonic devices in the future and have already been

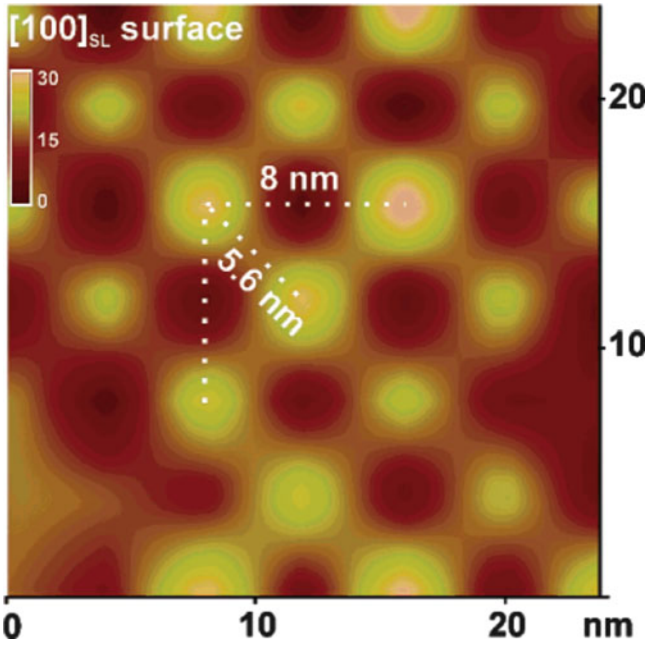


Fig. 8.2 Tapping-mode AFM image of a superlattice surface of Au nanoparticles synthesized by the SMAD method (Stoeva et al. [40], Fig. 13, by permission of American Chemical Society)

incorporated in a variety of consumer products because of their excellent antimicrobial properties.

Over the years, several methods have been reported for the synthesis of Ag nanoparticles. These include the Creighton method [42] in which sodium borohydride is used for the reduction of silver nitrate, or the Lee-Meisel method [43] in which silver nitrate is reduced by sodium citrate. The most common source of silver ions is silver nitrate; however, other silver salts such as silver perchlorate, silver sulfate, or silver 2-ethylhexonate have also been used. The high reduction potential of Ag ions (0.799 V) allows it to be reduced by several inorganic and organic compounds. Several reducing agents have been employed for the reduction of silver ions, $NaBH_4$, sodium citrate, ascorbic acid, ethanol, pyridine, *N*,*N*-dimethylformamide, hydrazine hydrochloride, poly(ethylene glycol), etc. Unlike Au, Ag is more prone to oxidation. Thus, for the long-term stability of the Ag nanoparticles, surfactants have been used as protecting agents [44]. Surfactants such as sodium 2-diethylhexyl sulfosuccinate (AOT), cetyltrimethylammonium bromide (CTAB), and stabilizers like poly(vinyl pyrrolidone) (PVP) have been successfully employed for stabilization of Ag nanoparticles and to also direct the formation of Ag nanoparticles into different shapes such as plates, rods, discs, etc. [45]. All the above methods, in general, produce nanoparticles of Ag; however, very careful control of conditions is necessary in order to obtain Ag nanoparticles with narrow particle size distribution. Scalability and the purity of the Ag nanoparticles are also a problem. The preparation of metallic silver nanoparticles is a challenge because of its relative ease of oxidation compared to Au nanoparticles as stated previously. Very often, after several days of standing under ambient conditions, the surface of the Ag nanoparticles becomes oxidized due to the formation of a thin oxide layer coating.

A key requirement for applications is a reproducible and reliable method that can produce large quantities of Ag nanoparticles with a narrow size distribution. The SMAD method has the potential for preparation of large-scale quantities; a 5,000-fold scale up is possible by the SMAD method compared to the inverse micelle method of production of metal nanoparticles. Gram scale quantities of monodisperse spherical Ag nanoparticles have been successfully prepared using the SMAD method followed by digestive ripening [46]. Two different capping agents or stabilizing agents were evaluated and it was found that the use of dodecanethiol led to Ag nanoparticles with a mean diameter of 6.6 ± 1 nm, while trioctyl phosphine capped particles were 6.0 ± 2 nm. Like the Au nanoparticles, the Ag nanoparticles were also found to organize into 2D and 3D superlattices with a well-defined geometry through self-assembly in solution. 2D arrays were found when either dodecanethiol or trioctyl phosphine were used as ligands; interestingly, only circular 3D superlattices were observed when trioctyl phosphine was used as the protecting ligand and triangular 3D superlattices were observed when dodecanethiol was used as the protecting ligands. The synthetic procedure for the preparation of Ag nanoparticles was similar to that used for Au; however, some changes were needed to prevent the oxidation of Ag. Ag nanoparticles are normally resistant to the formation of 3D structures and the structures reported earlier in the literature were found to be amorphous. Thus, the observation of crystalline 3D Ag superlattice structures formed via the SMAD method and subsequent digestive ripening is unique. The versatility of SMAD method is indicated by its successful application for the preparation of bimetallic nanoparticles. Noble metal intermetallic particles containing Au and Ag or Au and Cu are formed in gram quantities via digestive ripening in as little as 1 h. Highly monodisperse particles of 5.6 ± 0.5 nm diameters for Au–Ag and 4.8 ± 1.0 nm diameter for Au–Cu are formed and they undergo facile self-assembly to form 3D superlattice ordering [47].

Semiconductor Nanoparticles

Semiconductors are an important class of materials that are ubiquitous in our daily lives. Microprocessors, LEDs, transistors, and solar cells are all made of semiconducting materials. The size-dependent properties of semiconductor nanoparticles were first reported by several groups in the early 1980s. In their pioneering works, Brus, Henglein, and

Grätzel independently developed liquid phase synthesis of stable semiconductor nanoparticles suspended in solvents such as acetonitrile and water [48–50]. Several nanocrystalline semiconductors such as II-VI materials (CdS, HgS, and CdSe), I-VII materials (AgBr, AgCl), and III-V materials (GaAs) that are narrowly dispersed have been prepared. The strategy adopted by Brus was to separate the temporal growth and nucleation sites so that a highly crystalline and monodisperse semiconductor could be obtained. This is important since the presence of amorphous phase and/or trace amounts of impurities can significantly alter the physicochemical properties of these materials. Thus, synthetic strategies developed should be facile, reproducible, and most importantly lead to consistent and predictable electronic properties. A typical synthesis of a semiconductor nanoparticle is conducted in the following manner. A room-temperature solution containing the inorganic precursors is rapidly injected into a preheated solvent (350°C) to facilitate rapid nucleation at the same time. This process is called "burst nucleation" since all the particles nucleate almost simultaneously. This is followed by lowering the temperature to around 300°C and this allows for a slow growth of the nanocrystals until all the reactants are consumed. The size and shape of the semiconductor nanocrystals can be controlled by adjusting the ratio of the concentration of the inorganic precursors and the concentration of the stabilizing or capping agents (surfactants, dendrimers, and polymers). This type of synthesis is called the bottom-up method. In contrast, in the top-down method, the macroscopic or bulk material is "machined" down to nanometer length scale by laser ablation–condensation or lithographic techniques.

In bulk semiconductors, the conduction band and the valence band are separated by a band gap. On excitation, the electrons are promoted to the conduction band and holes are produced in the valence band. The charge carriers (electrons and holes) are separated by distances that encompass a number of ions or molecules that constitute the semiconductor material. This separating distance is called the Bohr radius and normally has dimensions on the nanometer scale. When the size of the semiconductor material becomes similar to the Bohr radius or smaller, it leads to a situation in which the charge carrier excitons have a restricted space or volume to move and thus their motion is confined. Like the motion of an electron in a box, the kinetic energy and the excitation energy of the electron increase as the size of the box decreases. Similarly, the band gap of the semiconductor increases as the size of the particles becomes smaller than their Bohr radius. Along with the band-gap, the color and emission of the nanoparticle change with change in the size of the nanoparticle. The physical, chemical, and optical properties of semiconductor nanoparticles thus are sensitive to both the size and shape of the nanoparticles. The surface-to-volume ratio increases as the semiconductor nanoparticle decreases in size and since the properties of the semiconductor depend on the size and shape of the nanoparticle, semiconductor nanoparticles exhibit very different properties compared to bulk semiconductors [51]. Several applications are envisaged for the use of these semiconductor particles; these include photovoltaics, nanoelectronics, and optoelectronics. Other applications where these semiconductor nanoparticles find extensive use are in the fields of photovoltaics and photocatalysis [52, 53].

The SMAD method has been successfully developed to synthesize gram scale quantities of high-purity metal sulfide semiconductor nanoparticles such as ZnS, PbS, CdS, SnS, and CdSe [54, 55]. The advantages of the SMAD method compared to other methods reported in the literature are: (1) gram scale quantities of stable nanocrystalline (1–10 nm) particle sizes can be prepared without any stabilizers with relative ease, (2) the synthesis is very reproducible and yields are typically >90%, (3) the nanoparticles have fairly uniform particle size distribution and hence are highly monodisperse, (4) no purification step is necessary and hence easy work-up affords high-purity metal sulfide nanoparticles, and (5) the surfaces of the metal sulfide nanoparticles are "clean," i.e., no oxide coating is found on the surface of the nanoparticles. In addition, the SMAD method for the synthesis of metal sulfide nanoparticles allows one to tune the particle size and hence its band gap by appropriate choice of solvents and warm-up time of the matrix consisting of frozen solvent and metal sulfide vapors. The textural properties such as specific surface area and pore volume of the nanoparticles are much higher compared to the commercial samples. For example, ZnS prepared by the SMAD method using pentane as the solvent leads to the production of ZnS nanoparticles with surface area as high as 237 m^2/g whereas a commercial ZnS (Fluka) has a surface area of only about 10 m^2/g. Similarly PbS and CdS nanoparticles prepared by the SMAD method have surface areas ~120 m^2/g. In the case of CdSe nanoparticles prepared by the SMAD method, broad photoluminescence (PL) was observed; however, the digestive ripening process led to an increase in the fluorescence intensity and very narrow PL spectra. Figure 8.3 shows the variation in luminescence with change in time of reflux.

Thus, the SMAD method of preparation of metal sulfide nanoparticles leads to the production of high surface area metal sulfides with unique physicochemical and optical properties.

Metal Oxide Nanoparticles

Metal oxides are ubiquitous in catalysis and are key components in several catalytic reactions. They function directly as catalytic reactive centers or serve as high surface

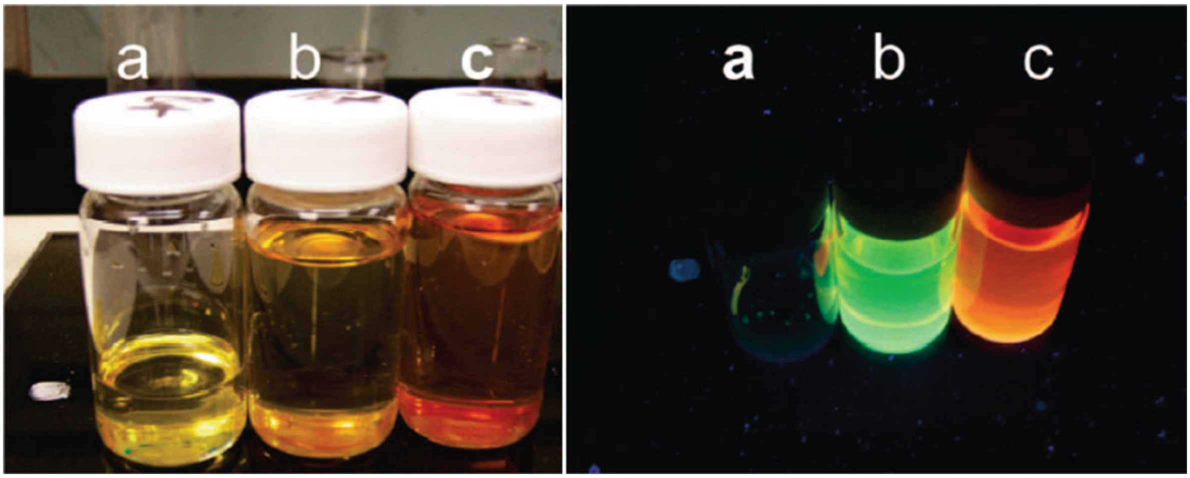


Fig. 8.3 Luminescence of CdSe samples prior to digestive ripening (*left*) and after digestive ripening (*right*). ((**a**–**c**) refer to samples before reflux, after 1 h reflux, and after 16 h of reflux. Reprinted with permission from Cingarepu et al. Chem. Mater. 2009, 21, 1248–1252, Fig. 2 copyright (2009) American Chemical Society)

area supports to disperse active metal centers or as promoters to enhance the rate of catalytic reactions. Many commercial catalysts consist of zero valent metal atoms dispersed finely on a high surface area metal oxide support such as silica or alumina.

Surface chemistry is of importance in numerous processes such as catalysis, corrosion, and adsorption. When the particles sizes are in the 1–10 nm range, a whole new field of surface chemistry is realized. In recent years, it has been possible to successfully produce nanocrystalline metal oxides of MgO, CaO, ZnO, TiO_2, CuO, CeO_2, and other binary metal oxides by a specially designed sol–gel-based process developed in the author's laboratory [56, 57]. The results obtained so far indicate that: (1) the intrinsic reactivities are higher per unit surface area for nanocrystals compared to microcrystals and (2) consolidation of the nanoparticles into pellets does not significantly lower the surface reactivity and the surface area when moderate pressures are used. The high surface area of the nanomaterial ensures that a high percentage of the atoms are on the surface. Thus, reactant–surface interactions can reach stoichiometric ratios. For spherical Fe atoms it has been calculated that for a 3 nm particle, 50% of the atoms lie on the surface while a 20 nm particle has less than 10% of the atoms on the surface. Another important feature that has larger practical implications is that the nanoparticles exhibit higher intrinsic chemical reactivities as the particle size becomes smaller. The reason for the dramatic increase in the activities is most probably due to changes in crystal shape and the presence of defect sites.

Until a few years ago, the crystal shape of solid materials was of academic curiosity only and shape was not considered to have a significant effect on the chemical properties and reactivities of a material. However, recent results obtained in the author's laboratory clearly indicate that the shape of nanocrystals influences the chemical reactivity significantly. For example, it has been shown that 4 nm sized nanocrystalline MgO particles adsorbs as much as six molecules of SO_2 per nm^2 at room temperature and 20 Torr pressure [58]. On the other hand, microcrystalline MgO adsorbs only two molecules of SO_2 per nm^2 under similar conditions. Similarly, an MgO nanocrystalline aerogel (AP-MgO) material adsorbs four times as much CO_2 as the MgO microcrystals. There are not only differences in the amounts of gaseous molecules adsorbed on these surfaces, but also the mode of surface binding can be different in the nanocrystalline materials. For example, SO_2 binds more predominantly as a monodentate species on the AP-MgO crystal but favors a bidentate geometry on conventionally prepared (CP-MgO) microcrystals. Clearly, these results indicate that the shape and size of the crystals affect the adsorptive properties of the MgO surfaces. Thus, the modified sol–gel process leads to the production of a new family of porous inorganic metal oxides that exhibit unique physicochemical properties [59].

Modified Aerogel Procedure

Aerogels are porous materials that have nanoscale dimensions and have low density and high surface area [60]. They are widely used as insulators and increasingly as heterogeneous catalysts. An aerogel procedure involving supercritical drying was first reported by Kistler in 1932 for the preparation of high surface area silica. However the procedure developed by Kistler was time consuming [61]. Teichner and his co-workers used organic solvents that dramatically decreased the processing time from weeks to hours [62].

A further modification of the aerogel method involved the addition of large amounts of aromatic hydrocarbons such as toluene to the alcohol–alkoxide mixture [63]. The addition of toluene dramatically increased the surface area of the resulting metal oxide and although the exact reason is not fully understood, it is believed that incorporation of a less polar solvent such as toluene could reduce the surface tension at the pore walls and thus minimize pore collapse and retain the high porosities of the "wet gel."

The modified aerogel procedure (MAP) for the preparation of inorganic metal oxides has four steps. Let us consider the preparation of MgO by MAP. The first step is the preparation of Mg alkoxide by the reaction of Mg metal with alcohols such as methanol or ethanol. The second step is the hydrolysis of the Mg alkoxide conducted in the presence of a non-polar solvent such as toluene. The third step is the supercritical or hypercritical drying procedure. In this step, the gel obtained in the hydrolysis step is placed in an autoclave and the wet gel is heated at a controlled rate so that the pressure and the temperature exceed the critical temperature (T_c) and the critical pressure (P_c) of the solvent entrapped inside the pores of the gel. The supercritical conditions vary depending on the nature of the solvent employed. The autoclave is slowly heated from room temperature to the critical temperature and after about 15 min of standing at T_c, the solvent is vented quickly through a port to give a dry solid powder. This drying procedure preserves the texture of the wet gel by avoiding the collapse of the pores and hence a highly porous, low-density material having very high specific surface area and large pore volume is obtained. In the final and fourth step, the metal hydroxide, obtained previously is heated under vacuum or air (typically at 500°C) to convert it into metal oxide. This step also gets rid of surface carbonates and/or other species such as alkoxy groups adsorbed on the surface.

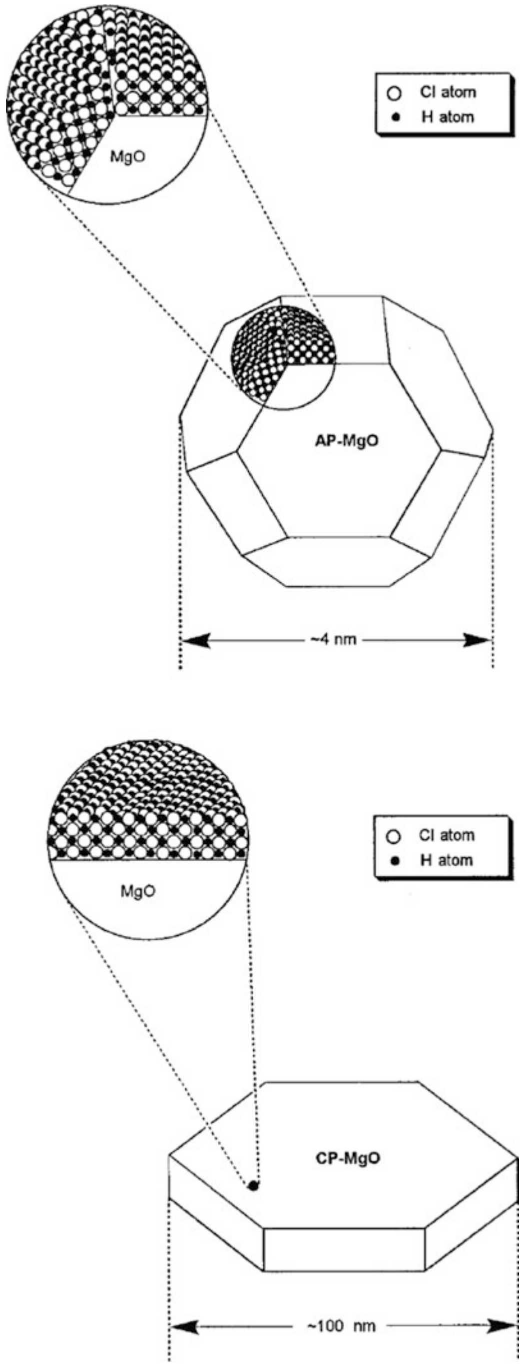


Fig. 8.4 Models of AP-MgO and CP-MgO crystallites. Insets illustrate multilayer HCl adsorption, suggesting more disorder in the AP-MgO case (reprinted with permission from Klabunde et al. [65], Fig. 5 copyright (1996) American Chemical Society)

The following section describes some of the properties of the metal oxide nanoparticles prepared by the MAP. The crystallite sizes of the metal oxide nanoparticles prepared by MAP are remarkable. For example, MgO have crystallite sizes ~4 nm, AP-CaO ~7 nm, AP-TiO_2 ~10 nm, AP-Al_2O_3 ~2 nm and AP-ZrO_2 ~8 nm [64–68]. The specific surface area of AP-MgO obtained is ~500 m^2/g, whereas a MgO sample that is conventionally prepared (referred to as CP henceforth) has specific surface area of ~200 m^2/g, while the specific surface area of a commercial (referred as CM) MgO sample is only around 30 m^2/g. The surface areas of AP-CaO, CP-CaO, and CM-CaO are typically ~150, ~100, and ~1 m^2/g, respectively. Similarly, AP-Al_2O_3 possesses specific surface areas as high as 810 m^2/g, while a commercial Al_2O_3 has a surface area of only about 100 m^2/g. These examples clearly indicate that the MAP results in the formation of ultrafine particles that have very high specific surface areas compared to commercial samples. Also, as the particle size decreases, the surface area increases and the reactivity is considerably enhanced. The increased reactivity is not simply due to the increased surface areas alone. These nanoparticles contain numerous defect sites such as crystal corners, edges, kinks, and ion vacancies. Figure 8.4 illustrates the differences in AP and CP MgO. The high

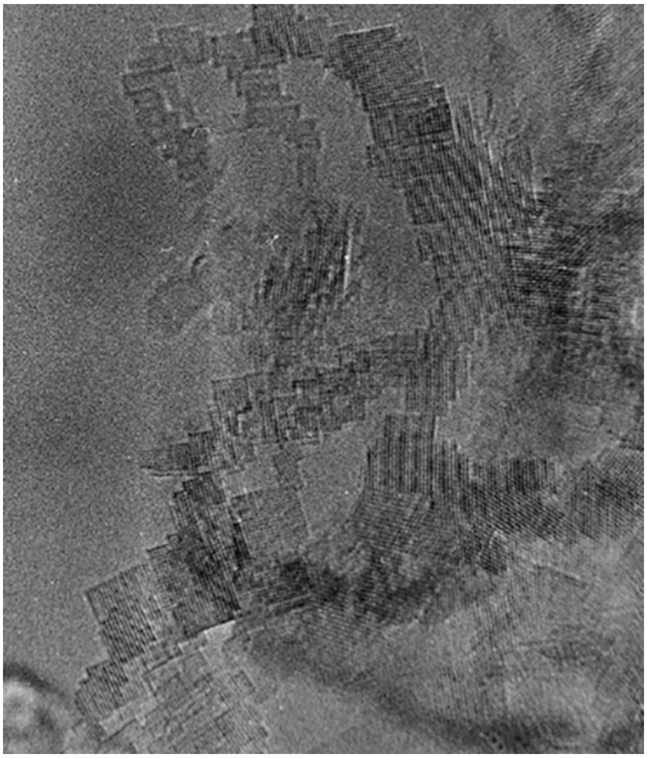


Fig. 8.5 High-resolution transmission electron microscopic image of aerogel prepared AP-MgO (reprinted with permission from Richards et al. [69], Fig. 2 copyright (2000) American Chemical Society)

reactivities of the polyhedral-shaped MgO and CaO are attributed to the higher percentage (~20%) of defect sites such as corner and edges on the surfaces of AP-MgO and AP-CaO. In contrast, a conventionally prepared CP-MgO possesses less than 0.5% of defect sites while a commercial sample of MgO and CaO essentially contains very little defect sites since it is highly crystalline. In addition, surface hydroxyl groups which can be either isolated or lattice bound add to the rich surface chemistry exhibited by these metal oxide nanoparticles.

The MAP process leads to the formation of free flowing ultrafine powders rather than rigid monoliths. TEM studies of AP-MgO show porous web-like aggregates in the range of about 1,400 nm. These are formed by the interaction of the 4 nm (average) polyhedral crystallites and their overall size distribution is narrow. A high-resolution TEM image of AP-MgO is shown in Fig. 8.5.

The TEM picture clearly shows that the cube-like crystallites aggregate into polyhedral structures that have numerous edge and corner sites. Also, interesting is the formation of pores between the crystalline structures that are visible at this magnification. Essentially, the TEM picture is an atomic resolution image of MgO nanocrystallites with lattice planes of Mg and O ions clearly revealed.

An interesting behavior is exhibited by the MgO when the loose powders of it are consolidated into pellets. AP-MgO powder has a low density of 0.30 cm^3/g. Upon compression at low pressure (3,000 psi) this increases to 0.58 cm^3/g, and at 20,000 lb, the density increases to 1.0 cm^3/g. The surface area of AP-MgO powder is 364 m^2/g. Upon compaction at a low pressure of 1,000 psi, the surface area was found to reproducibly increase to ~370 m^2/g [69]. On compaction at higher pressures of 5,000 and 20,000 psi, the specific surface area was found to be 366 and 342 m^2/g, respectively. Another interesting aspect is that these powders exhibit adsorption/desorption isotherms typical of bottleneck pores. However, after compression at 1,000 psi, the samples of $Mg(OH)_2$ and MgO exhibit bottleneck and cylindrical pores, open at both ends. For comparison purposes the conventionally prepared (CP) samples too were subjected to compaction and their textural properties measured. The CP samples of MgO and CaO were more susceptible to collapse of pore volume, especially the CaO samples. However, the AP-CaO exhibited more resistance. The AP oxides, due to their polyhedral nanocrystal shapes and their tendencies toward forming porous web-like aggregates, are the most resistant to collapse under pressure. The CP oxides, due to their more ordered hexagonal platelet shapes, more easily compress into denser structures.

The aerogel-prepared metal oxide nanoparticles constitute a new class of porous inorganic materials because of their unique morphological features such as crystal shape, pore structure, high pore volume, and surface areas. Also, it is possible to load catalytic metals such as Fe or Cu at very high dispersions on these oxide supports and hence the nanocrystalline oxide materials can also function as unusual catalyst supports. Further, these oxides can be tailored for desired Lewis base/Lewis acid strengths by incorporation of thin layers of other materials or preparation of mixed metal oxides [70, 71].

Mixed Metal Oxide Nanoparticles

Many methods have been reported for the synthesis of mixed metal oxides. Generally they have been prepared by co-precipitation of metal hydroxides followed by high temperature treatment; solid-state physical mixing of hydroxides, oxides, or nitrates followed by high temperature treatment or by hydrolysis of bimetallic bridged alkoxides. However, all these methods of preparation lead to partial segregation of the individual oxides as final products. This is because the starting precursors tend to hydrolyze at different rates leading to phase segregation. Further, the very high temperature treatment leads to sintering of the particles resulting often in low surface area oxide materials. The MAP method avoids the problems inherent in the synthetic methods and provides an ideal way to prepare intermingled mixed metal oxide nanoparticles that are essentially mixed at the molecular level. In the MAP method, a large portion of a spectator solvent (toluene) is added that greatly enhances gelation rates, thereby increasing the chances that the two hydrolyzing metal alkoxides gel together. This results in the formation intermingled mixtures or gels that are essentially molecular in nature. This technique was used to prepare a series of intimately intermingled mixed metal oxide nanoparticles [72, 73]. These mixed oxides were composed of alkaline earth oxide and alumina. It was found that only aerogel-prepared AP-$MgAl_2O_4$ and AP-$BaAl_2O_4$ showed peaks in the powder X-ray diffractometry (XRD) studies indicating high crystallinity. In the XRDs for AP-$MgAl_2O_4$ and AP-$BaAl_2O_4$, it was possible to identify peaks corresponding to MgO and BaO. This clearly indicates that the samples are intimately mixed. The intimately intermingled mixed metal oxide samples retain the high surface area and relatively large pore volumes when compared with the individual metal oxides. High-resolution TEM studies clearly indicate the extent of the intermingling of the two oxides. Individual and aggregated Al_2O_3 boehmite planes mixed with AP-MgO nanocrystals could be observed. This gives a clear indication that the MgO and Al_2O_3 are intimately mixed throughout the entire material resulting in some unique structures. The presence of MgO "guest" planes in the "host" AlO(OH) planes has led to an increase in the distance between the planes. Typically in AP-Al_2O_3 the spacing between planes is 6 Å. MgO is now sandwiched between the boehmite planes and the spacing has increased to 15 Å.

Thus, an interesting facet of the MAP procedure is that one can engineer and tune acid/base sites in them. The intimately intermingled MgO-Al_2O_3 nanostructures discussed earlier exhibit enhanced capacity/activity over the pure forms of either AP-Al_2O_3 or AP-MgO. The enhanced reactivity is attributed to the Lewis base nature of the very small and very well dispersed MgO crystallites that are "housed" within the large pores of Al_2O_3. In addition, such intimately intermingled mixed oxides are highly thermally stable with minimal sintering after heating up to 700°C.

Selected Applications

As stated earlier, catalytic processes contribute to at least 20% of the GDP of the USA. Thus, an important application of nanostructured materials lies in the field of heterogeneous catalysis. It is beyond the scope of this chapter to discuss the numerous applications of nanoparticles as catalysts. However, we shall confine our discussion to a select few examples that demonstrate the importance of nanoparticles in the field of environmental catalysis.

Catalysis: Dechlorination and Dehydrochlorination

The disposal and destruction of chlorinated compounds is a subject of great importance. The cost of complete elimination of chlorinated compounds is quite staggering and it is estimated to be over $100 billion/year. In 1993 some environmental agencies had proposed the need for a "chlorine free economy" [74]. The most widespread group of chlorocarbons includes polychlorinated biphenyls (PCBs) and cleaning solvents such as CCl_4 and C_2Cl_4. The most common method used to destroy chlorocarbons is high temperature thermal oxidation (incineration). The toxic chlorinated compounds seem to be completely destroyed at high temperatures; however, there is great concern about the formation of toxic by-products such as dioxins which are far more toxic than the chlorinated compounds themselves. The desired reaction is the complete oxidation of chlorocarbons to water, CO_2, and HCl without the formation of any toxic by-products.

Chlorocarbons are destroyed during reactions with nanocrystalline metal oxides by becoming mineralized as environmentally benign metal chlorides and CO_2 gas [66, 75, 76]. The process can be written as

$$2MO + CCl_4 \rightarrow CO_2 + 2MCl_2 \tag{8.4}$$

From a purely thermodynamic perspective, the reaction of metal oxides with chlorocarbons with MgO and CaO are energetically favorable according to the following equations:

$$\begin{gathered} 2MgO_{(s)} + CCl_4 \rightarrow 2MgCl_{2(s)} + CO_{2(g)} \\ \Delta H^0 = -334\ \text{kJ/mol} \end{gathered} \tag{8.5}$$

$$\begin{gathered} 2CaO_{(s)} + CCl_4 \rightarrow 2CaCl_{2(s)} + CO_{2(g)} \\ \Delta H^0 = -573\ \text{kJ/mol} \end{gathered} \tag{8.6}$$

CaO and MgO can destroy chlorocarbons such as CCl_4, $CHCl_3$, and C_2Cl_4 at temperatures around 400–500°C in the absence of an oxidant, yielding mainly CO_2 and the corresponding metal chlorides. Also, it has been demonstrated that if MgO or CaO nanoparticles were coated with transition metal oxides (e.g., Mn_2O_3, Fe_2O_3, CoO, or NiO), the reactivity could be enhanced substantially and a kinetic advantage could be obtained [77]. The dehydrochlorination is represented as shown by (8.7) below:

$$R\text{-}CH_2\text{-}CH_2\text{-}Cl \rightarrow R\text{-}CH = CH_2 + HCl \tag{8.7}$$

Nanoscale AP-MgO and AP-CaO were found to be superior to conventionally prepared CP-CaO, CP-MgO, and commercial CaO and MgO catalysts for the dehydrochlorination of several toxic chlorinated substances. The interaction of 1-chlorobutane with nanocrystalline MgO at 200–350°C results in both stoichiometric and catalytic dehydrochlorination of 1-chlorobutane to isomers of butene and simultaneous topochemical conversion of MgO to $MgCl_2$ [78]. During the course of the study, it was found that the surface of MgO was converted to $MgCl_2$ and a steady performance of the catalyst was observed only after the formation of the surface $MgCl_2$ phase. Also, a decrease in the surface area during the transformation of MgO to $MgCl_2$ was found due to aggregation of the nanoparticles. The study demonstrates that multi-component systems comprising nanoparticles of $MgO/MgCl_2$ have unusual catalytic properties in acid–base reactions.

Destructive Adsorption of Chemical Warfare (CW) Agents

Chemical warfare agents constitute one of the biggest threats to mankind along with biological warfare agents such as bacteria, fungi, and viruses. The huge stockpiles of CW nerve agents in the USA and Russia and other nations such as North Korea is a matter of grave concern to the public in these countries. The CW agents are known to irreversibly react with the enzyme acetylcholinesterase (AChE) which prevents the enzyme from controlling the central nervous system and hence cause irreversible damage to humans. Hence, many methods have been attempted to decontaminate or neutralize these CW agents [79]. However, only a few reactions are feasible for practical neutralization. Examples of CW agents include VX [*O*-ethyl S-(2-diisopropylamino) ethylmethyl-phosphonothioate], GD (pinacolyl methylphosphono-fluoridate), and HD [bis(2-chloroethyl) sulfide].

The traditional method of decontaminating CW agents is to neutralize them in basic media in solution phase. An approach in this direction has been to use oxidizing reagents such as aqueous bleach (NaOCl or $Ca(OCl)_2$) or hypochlorous acid (HClO). However, these reactions require careful control of pH and solubility problems arise frequently; for example, in basic bleach, VX has solubility problems. Also, large quantities of liquids have to be used for complete neutralization and hence aqueous phase methods are cumbersome. Dry powders have numerous advantages: (1) they are non-toxic, (2) easy to handle and store, (3) waterless and hence have low logistical burden and easy to transport, and finally (4) there is no liquid waste stream. Nanosize MgO, CaO, and Al_2O_3 have been used in room temperature reactions with chemical warfare agents such as VX, GD, and HD [80–82]. Reaction with nerve agents VX and GD and metal oxides results in the nerve agents being hydrolyzed. There are two important differences between their solution behavior and their destructive adsorption on nanoscale metal oxides. The first difference is that the non-toxic phosphonate products reside as surface-bound complexes. The second important difference is that toxic EA-2192, which is known to form under basic hydrolysis in aqueous phase, is not observed on the surface of either MgO or CaO.

Commercial Products

A comprehensive list of consumer products based on nanotechnology is available in the Internet [1]. As of early 2011, there are 1,015 products produced by 484 companies located in 24 countries. United States continues to be the leader in the commercialization with over 500 products. The inventory was first created in 2006 with a list containing initially 215 products. Of the 1,015 products, health and fitness accounts for over 60% of the products with 605 listed. In this category, the sub-category of personal care and clothing contains the largest number of products. When the products are examined in terms of the chemical constituent responsible for the product being listed in this inventory, silver nanoparticles account for over 250 products. The other two important classes of materials are carbon and titanium dioxide. Silver nanoparticles have been incorporated into a wide variety of products that include socks, wound dressings, shirts, toothpastes, air filters and purifiers, refrigerators, washing machines, and vacuum cleaners. The second widely used material, carbon is found in tennis rackets, bicycles, and tires to name a few products. Titanium dioxide nanomaterial is a popular choice of sunscreen lotions and creams because it has the ability to absorb harmful UV rays.

Conclusions

It is quite clear from the various applications listed above that consumer products based on nanotechnology have made inroads into our daily lives. As the technology becomes more mature and more products based on the principles of nanotechnology appear in the market, societal and environmental implications have also to be considered. Without a doubt, the economic potential of nanotechnology is tremendous; however, it is best to balance the benefits with potential risks and identify strategies to mitigate these. As discussed in this chapter, the reactivity of nanoparticles is significantly higher compared with bulk materials. Although the enhanced reactivity is useful for most applications, environmental effects are not known at this moment. Several properties such as solubility, permeability, and adsorption need to be studied in greater detail and their interaction with aquatic and ecological ecosystems needs to be monitored closely when they are released into the environment. For example, Ag nanoparticles have been detected in several wastewater treatments plants and it has been documented that at high concentrations, the growth of nitrifying bacteria is inhibited. This has prompted the EPA to take action to issue guidelines and policies that protect against unreasonable risk to human health and environment. It is worth concluding this chapter by quoting the Nobel laureate Richard Smalley, "*Just wait—the next century is going to be incredible. We are about to be able to build things that work on the smallest possible length scales, atom by atom. These little nanothings will revolutionize our industries and lives*" [83].

References

1. The project on emerging nanotechnologies (2011) Woodrow Wilson International Center for Scholars. http://www.nanotechproject.org/inventories/consumer/. Accessed Jan 2011
2. Grätzel M, Kalyanasundaram K (eds) (2011) Nanotechnology: light and energy. Wiley-VCH, Berlin, Germany
3. Wehrspohn R (ed) (2011) Ordered porous nanostructures and applications. Springer, New York
4. Welles AE (ed) (2010) Silver nanoparticles: properties, characterization and applications. Nova Science Publications, Hauppauge, New York
5. Rogach A (ed) (2010) Semiconductor nanocrystal quantum dots: synthesis, assembly, spectroscopy and applications. Springer, New York
6. Klabunde KJ, Richards RM (2009) Nanoscale materials in chemistry. Wiley-VCH, Berlin, Germany
7. Ozin GA, Arsenault AC, Cademartiri L (eds) (2009) Nanochemistry: a chemical approach to nanomaterial. Royal Society of Chemistry, Cambridge, UK
8. Niederberger M, Pinna N (eds) (2009) Metal oxide nanoparticles in organic solvents: synthesis, formation, assembly and application. Springer, New York
9. de Villiers MM, Aramwit A, Kwon GS (eds) (2008) Nanotechnology in drug delivery. Springer, New York
10. Mitin VV, Kochelap VA, Stroscio MA (eds) (2008) Introduction to nanoelectronics: science, nanotechnology, engineering, and applications. Cambridge University Press, Cambridge, UK
11. Wiesner M, Bottero J-Y (eds) (2007) Environmental nanotechnology: applications and impacts of nanomaterials. McGraw-Hill, New York
12. Hutchison J, Kirkland A (eds) (2007) Nanocharacterisation. Royal Society of Chemistry, Cambridge, UK
13. Wolf EL (ed) (2006) Nanophysics and nanotechnology: an introduction to modern concepts in nanoscience. Wiley-VCH, Berlin, Germany
14. Cavaleiro A, de Hosson JT (eds) (2006) Nanostructured coatings. Springer, New York, USA
15. Wang ZL (ed) (2005) Nanowires and nanobelts: materials, properties and devices: volume 1: metal and semiconductor nanowires. Springer, New York, USA
16. Miziolek AW, Karna SP, Vaia RA (eds) (2005) Defense applications of nanomaterials. American Chemical Society Symposium Series, Washington DC, USA
17. Niemeyer CM, Mirkin CA (2004) Nanobiotechnology: concepts, applications and perspectives. Wiley-VCH, Berlin, Germany
18. Ratner MA, Ratner D (eds) (2002) Nanotechnology: a gentle introduction to the next big idea. Prentice Hall, New Jersey, USA
19. Ajayan PM, Schadler LS, Braun PV (2003) Nanocomposite science and technology. Wiley-VCH, Berlin, Germany
20. Wilson M, Kannangara K, Smith G, Simmons M, Raguse B (2002) Nanotechnology: basic science and emerging technologies. Chapman and Hall, CRC Press, Boca Raton, Florida, USA
21. Taniguchi N (1974) On the basic concept of nano-technology. Proc. Intl. Conf. Prod. Eng. Part II, Japan Society of Precision Engineering, Tokyo
22. Drexler KE (1986) Engines of creation. Anchor Publishers, Garden City, NY
23. Feynman RP (1959) Annual meeting of the American Physical Society. California Institute of Technology, Pasadena, CA
24. Dickson DPE (1996) In: Edelstein AS, Cammarata RC (eds) Nanomaterials: synthesis, properties and applications. Institute of Physics Publishing, Bristol, UK, pp 459–476
25. Turkevich J, Stevenson PC, Hillier J (1951) Discuss Faraday Soc 11:55–75
26. Brust M, Walker M, Bethell D, Schiffrin DJ, Whyman RJ (1994) J Chem Soc Chem Commun 801–802
27. Klabunde KJ, Li YX, Tan BJ (1991) Chem Mater 3:30–39
28. Li YX, Klabunde KJ (1990) J Catal 126:173–186
29. Klabunde KJ, Efner HF, Murdock TO, Ropple R (1976) J Am Chem Soc 98:1021–1023
30. Schmid G, Chi L (1998) Adv Mater 10:515–525
31. Bethell D, Schiffrin DJ (1996) Nature 382:581
32. Green M, O'Brien P (1999) J Chem Soc Chem Commun 2235–2241
33. Klabunde KJ (1989) US Patent. 4,87,677
34. Stoeva S, Klabunde KJ, Sorensen CM, Dragieva I (2002) J Am Chem Soc 124:2305–2311
35. Lin XM, Sorensen CM, Klabunde KJ (2000) J Nanopart Res 2:157–164
36. Prasad BLV, Stoeva SI, Sorensen CM, Klabunde KJ (2003) Chem Mater 15:935–942
37. Klabunde KJ, Stoeva SI, Sorensen CM (2003) US Patent. 6,562,403
38. Prasad BLV, Stoeva SI, Sorensen CM, Klabunde KJ (2002) Langmuir 18:7515–7520
39. Lin XM, Sorensen CM, Klabunde KJ (1999) Chem Mater 11:198–202
40. Stoeva SI, Prasad BLV, Uma S, Stoimenov PK, Zaikovski V, Sorensen CM et al (2003) J Phys Chem B 107:7441–7448
41. Smetana AB, Klabunde KJ, Marchin GR, Sorensen CM (2008) Langmuir 24:7457–7464
42. Creighton JA, Blatchford CG, Albrecht MG (1979) J Chem Soc Faraday Trans 2(75):790–798
43. Lee PC, Meisel D (1982) J Phys Chem 86:3391–3395
44. Leopold N, Lendl B (2003) J Phys Chem B 107:5723–5727
45. Caswell KK, Bender CM, Murphy CJ (2003) Nano Lett 3:667–669
46. Smetana A, Sorensen CM, Klabunde KJ (2005) J Colloid Interface Sci 284:521–526
47. Smetana A, Klabunde KJ, Sorensen CM, Ponce AA, Mwale B (2006) J Phys Chem B 110:2155–2158
48. Rossetti R, Ellison JL, Gibson JM, Brus LE (1983) J Chem Phys 80:4464–4469
49. Fojtik A, Weller H, Koch U, Henglein A (1984) Ber Bunsenges Phys Chem 88:969–977
50. Ramsden JJ, Grätzel M (1984) Faraday Trans 80:919–933
51. El-Sayed ME (2004) Acc Chem Res 37:326–333
52. Kamat PV (2007) J Phys Chem C 111:2834–2860
53. Kaneko M, Okura I (eds) (2002) Photocatalysis: science and technology. Springer, New York

54. Cingarapu S, Yang Z, Sorensen CM, Klabunde KJ (2009) Chem Mater 21:1248–1252
55. Sorenson CM (2010) Nanoscale materials in chemistry: environmental applications. American Chemical Society Symposium Series, Washington, DC, pp 35–49
56. Lucas E, Decker S, Khaleel A, Seitz A, Fultz S, Ponce A et al (2001) Chem Eur J 7:2505–2510
57. Rajagopalan S, Koper O, Klabunde KJ, Malchesky PS, Winecki S (2005) US Patent. 6,887,302
58. Stark JV, Park DG, Lagadic I, Klabunde KJ (1996) Chem Mater 8:1904–1912
59. Choudary BM, Mulukutla RS, Klabunde KJ (2003) J Am Chem Soc 125:2020–2021
60. Pierre AC, Pajonk GM (2002) Chem Rev 102:4243–4265
61. Kistler SS (1932) J Phys Chem 36:52–64
62. Teichner SJ (1986) Aerogel. In: Fricke J (ed) Proceedings of the First International Symposium, Wurzburg. Springer, Berlin, p 22
63. Utamapanya S, Klabunde KJ, Schlup JR (1991) Chem Mater 3:175–181
64. Koper OB, Lagadic I, Volodin A, Klabunde KJ (1997) Chem Mater 9:2468–2480
65. Klabunde KJ, Stark JV, Koper O, Mohs C, Park DG, Decker S et al (1996) J Phys Chem 100:12142–12153
66. Koper O, Klabunde KJ (1993) Chem Mater 5:500–505
67. Carnes CL, Kapoor PN, Klabunde KJ (2002) Chem Mater 14:2922–2929
68. Bedilo A, Klabunde KJ (1997) Nanostructured Mater 8:119–135
69. Richards R, Li W, Decker S, Davidson C, Koper O, Zaikovski V et al. (2000) J Am Chem Soc 122:4921–4925
70. Melgunov MS, Melgunova EA, Zaikovskii VI, Fenelonov VB, Bedilo AF, Klabunde KJ (2003) Langmuir 19:10426–10433
71. Demydov D, Klabunde KJ (2004) J Non Cryst Solids 350:165–172
72. Carnes CL, Kapoor PN, Klabunde KJ, Bonevich J (2002) Chem Mater 14:2922–2929
73. Medine GM, Zaikovskii V, Klabunde KJ (2004) J Mater Chem 14:757–763
74. Amato I (1993) Science 261:152–154
75. Koper O, Lagadic I, Klabunde KJ (1997) Chem Mater 9:838–848
76. Koper O, Klabunde KJ (1997) Chem Mater 9:2481–2485
77. Decker S, Klabunde KJ (1996) J Am Chem Soc 118:12465–12466
78. Fenelonov VB, Mel'gunov MS, Mishakov IV, Richards RM, Chesnokov VV, Volodin AM et al (2001) J Phys Chem B 105:3937–3941
79. Yang YC (1999) Acc Chem Res 32:109–115
80. Wagner GW, Bartram PW, Koper O, Klabunde KJ (1999) J Phys Chem B 103:3225–3228
81. Wagner GW, Koper OB, Lucas E, Decker S, Klabunde KJ (2000) J Phys Chem B 104:5118–5123
82. Wagner GW, Procell LR, O'Connor RJ, Munavalli S, Carnes CL, Kapoor PN et al (2001) J Am Chem Soc 123:1636–1644
83. Smalley R (1999) Nanotechnology, Congressional Hearings—Emerging Technologies in the New Millennium. The U.S. Senate Committee on Commerce, Science and Transportation, May 12

9

Nanostructured Materials: Industrial Applications

Kristin Clement, Angela Iseli, Dennis Karote, Jessica Cremer, and Shyamala Rajagopalan

Introduction

Nanoscience and nanotechnology are transforming materials science in a broad way, in a manner similar to polymer chemistry's transformation of materials science over the preceding century. The continuous development of novel nanostructured materials and the extensive study of physico-chemical phenomena at the nanoscale are creating new approaches to innovative technologies that are constantly resulting in products with a wide range of applications [1–5].

Nanoscience is generally defined as the study of phenomena and manipulation of materials at slightly above atomic or molecular scale, where properties of matter differ significantly from those at the macro scale. The term nanotechnology generally refers to techniques capable of designing and synthesizing nanomaterials that offer advanced material properties for novel applications. The US National Nanotechnology Initiative (NNI) defines a technology as nanotechnology only if it involves all of the following [6]:

- Research and technology development involving structures with at least one dimension in approximately the 1–100 nm range, frequently with atomic/molecular precision.
- Creating and using structures, devices, and systems that have novel properties and functions because of their nanometer scale dimensions.
- Ability to control or manipulate on the atomic scale.

History of Nanomaterials

It should be pointed out that nanomaterials are not entirely new. Carbon black, a material discovered in the early 1900s, is a nanomaterial that is used to increase the life of car tires and to provide the black color. Fumed silica, a component of silicone rubber, adhesives, coatings, and sealants, is a nanomaterial that has been in the market since the 1940s. Several chemicals and chemical processes possess nanoscale features—for example, large polymer molecules made up of tiny nanoscalar subunits. The first nanotechnologists were the Roman glass workers who invented dichroic glass. Other early nanotechnologists were the medieval glass workers who were responsible for providing the nanomaterials that gave stained glass and glazes their bright and vibrant colors. Similarly, nanotechnology has been used to create the tiny features on computer chips for the past several years. Over the last 20 years, the understanding of some preparations of metals, oxides, and other substances as nanomaterials has emerged through a vast body of research and development efforts. Additionally, with the development of more advanced analytical techniques, new nanomaterials are being developed in a systematic manner and with a greater understanding.

The origin of modern nanotechnology is mainly attributed to Professor Richard Feynman's speech "There's Plenty of Room at the Bottom," which was delivered in 1959 at the annual meeting of the American Physical Society at Caltech [7]. At the time, Professor Feynman's predictions were based on theoretical speculation that the principles of physics should allow the possibility of manipulating things atom by atom. He described such atomic scale fabrication as a bottom-up approach, as opposed to the top-down approach that is generally employed in manufacturing. However, developments such as the invention of the Scanning Tunneling Microscope in 1981 have since made nanoscale science a reality. Nanotechnology is now a rapidly growing field of research and development that is cutting across many traditional scientific boundaries.

K. Clement • A. Iseli • D. Karote • J. Cremer • S. Rajagopalan (✉)
NanoScale, Corporation 1310 Research Park Drive, Manhattan,
KS 66502, USA
e-mail: srajagopalan@NanoScaleCorp.com

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_9,

Nanomaterials: What and Why

In typical nanomaterials, a large fraction of the atoms are located on the surface of the particles, whereas in conventional materials, they are located in the bulk. As a result, many of the intrinsic properties of nanomaterials are different from conventional materials, since many of the atoms are in a different environment. Two key factors cause the properties of nanomaterials to differ significantly from other materials: increased surface area and quantum effects. Nanomaterials have much greater surface area per unit mass compared to conventional materials. As chemical reactions occur at surfaces, this means that a given mass of nanomaterials will be more accessible than the same mass of material made up of larger particles. In addition, higher surface energies and the presence of many corners, edges, and surface defects cause additional intrinsic reactivity. Quantum effects also begin to dictate the properties of matter as size is reduced to the nanoscale, which in turn affects the optical, electrical, and magnetic behavior of materials.

Since industrial sectors such as aerospace, biotech, energy, healthcare, and transportation depend on materials and devices made of atoms and molecules, they can all be improved by application of nanomaterials. Indeed, this rapidly advancing field of science has afforded a variety of commercially available products including catalysts, cosmetics, electronics, paints, self-cleaning windows, suntan lotions, and stain-resistant clothing [8]. Of all the materials, the most commonly mentioned materials are carbon based on fullerenes and nanotubes, followed by silver, silica, titanium dioxide, zinc oxide, and cerium oxide. Iron nanoparticles appear to have potential for environmental applications, particularly for groundwater remediation [9, 10]. Nanoparticles of titanium dioxide and zinc oxide are included in personal care products such as toothpaste, beauty products, sunscreens, and textiles [11, 12]. Silver nanoparticles are used as antimicrobials in textiles such as socks and underwear and are recommended for use in food packaging and in detergents [8]. Areas producing the greatest revenue for nanoparticles reportedly are chemical–mechanical polishing, magnetic recording tapes, sunscreens, automotive catalyst supports, biolabeling, electroconductive coatings, and optical fibers. The number of consumer products on the market containing nanomaterials (as particles or fibers) now exceeds 1,000 and is growing quickly. This count includes nanomaterials containing carbon, cerium oxide, silver, silica, titanium dioxide, magnesium oxide, and zinc oxide. The remarkable advances in current nanotechnology are mirrored in our capability to design and control the chemical composition, size, shape, and assembly structure for various applications.

The advantages of nanotechnology are spreading to numerous industries and applications. The ability of nanoparticles to make products lighter, stronger, and less expensive will infiltrate nearly every part of our lives from transportation to what we wear to treatments of the human body. Products using nanotechnology are already on the market and include stronger tennis racquets and golf clubs, cleaning products with powerful bactericidal and odor removing capabilities, and fabrics treated with nanoparticles to provide odor, stain, or liquid-resistant clothing. Cosmetic companies are also finding nanomaterials to be increasingly effective in anti-wrinkle creams and sunscreens, along with many other applications that will be discussed later in the chapter.

Governments and global corporations are investing billions of dollars to grow nanotech research and development for many other applications that will be greener and more sustainable than current technologies. This investment will fuel the $1.6 trillion of nanomaterials that are expected to be incorporated into manufactured goods in 2013 [13].

Quite recently, the President's Council of Advisors on Science and Technology (PCAST) released their review of the US NNI. Among the major recommendations is enhanced support of commercialization efforts including nanomanufacturing along with continued support of science and engineering, better coordination between US agencies, and increased efforts in nanoscience education and in clearly understanding the societal impacts of nanotechnology. Most importantly, a forward-looking view of environmental health and safety for nanomaterials is promoted [14].

Worldwide Initiatives on Nanoscience and Nanotechnology

As mentioned earlier, the existence of nanosized materials has been known for centuries, but only recently has characterization and manipulation of nanoscale systems opened a new era of nanoscience. Worldwide initiatives in nanoscience have gained great momentum from public and private investors due to economic benefits. In the present global market, a market dominated by tough competition, and a taste for superior products and services, experts believe nanotechnology advancements promise a stronger economy, greater return on investments, and soaring job creation. Countries have realized the importance of nanotechnology for future economic development and its potential impact on various industries. Nanotechnology is expected to revolutionize the existing major industries to create new, enhanced products and processes. Energy, medicine, military, communication, electronics, transportation, and material sciences are just a few segments that would be improved by nanotechnology. Experts believe that all areas will be impacted by nanotechnology, which could lead to exponential economic growth and possibly be the next great accelerant for growth similar to the information technology boom.

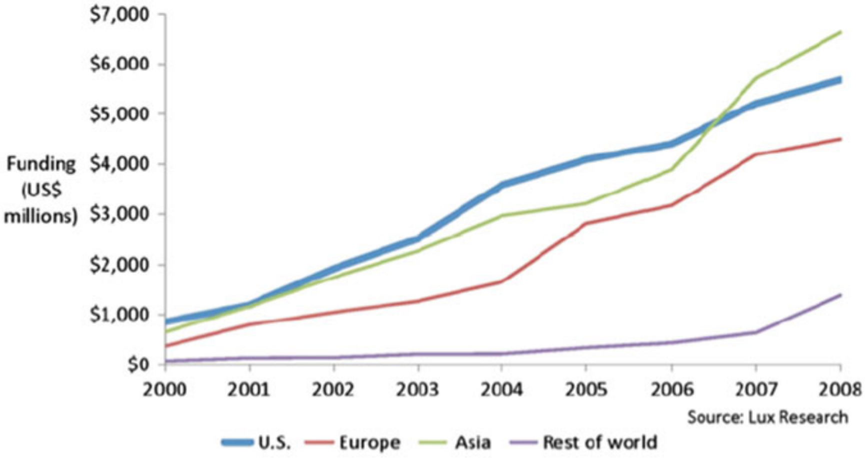


Fig. 9.1 Nanotechnology funding from 2000 to 2008 (includes public and private investments)

According to Lux Research, public and private investments totaled 18.2 billion USD globally in 2008. For the first time in 2007, private funding for nanotechnology worldwide surpassed public funding, indicating that basic research has led to commercialization and return on investments. The sales of nanotechnology related products grew 41% from 2007 to 2009, to a total of 224 billion USD [14].

Initiative and investment by governments in developed and emergent countries in this technology are an unambiguous indication of the relevance of nanotechnology. A market research firm, Cientifica, reports that in 2004, 85% of worldwide R&D spending was dominated by the European Union (27 EU member states), Japan, and the United States. In 2009, this dropped to 58% reflecting the emergence of additional countries contributing to nanotechnology [15]. Although the EU, Japan, and the US have increased their investments through 2009, other developing countries are catching up. Countries have varied approaches towards nanotechnology advancements. Some countries establish specific agencies to monitor progress, fund research, and transfer research to commercialization, while others utilize existing science technology for their advancements. According to an Organisation for Economic Co-Operation and Development (OECD) report, 17 of the 24 countries surveyed had specific governmental agencies focused on nanoscience, while six of the 24 countries invested in nanotechnology through the existing science and technology (S&T) agencies [16]. As countries continue the race to exploit nanotechnology for socioeconomic benefits, the scale of nanotechnology advancements can be measured by the number of nanotechnology centers and initiatives, public and private funding, number of active companies, number of scientific publications, patents, and number of Ph.D. graduates. The United States invests the most in the field of nanotechnology, but is losing ground against other foreign competitors (Fig. 9.1). Progress in basic and advanced research can be gauged by grouping the number of nanotechnology publications by country in the science citation index. Figure 9.2 illustrates publications reported in various geographical areas. The EU, which acts as the policy maker and funding agency, as well as one voice for Europe, has the most nanotechnology publications. China (including Taiwan) surpassed the US in 2007 for the number of nanotechnology publications [14].

Figure 9.3 shows China's dominance in patents compared to other countries. Currently, China has a larger number of patent applications than patents issued; nonetheless, China's progression in nanotechnology is impressive.

The US produces the majority of Ph.D. students in the field of science and engineering, but most of these students are foreign nationals. More than one third of these students leave the US, resulting in technology drain [14].

Governmental involvement in promoting nanoscience is the major driver for the current progress in the area of nanoscience. In the next section, various governmental initiatives introduced to promote nanotechnology growth are discussed. We will focus on key players in regions that have played a major role to date. These regions include North and South America, Europe, Eurasia (Russia), Asia, and Australia.

North and South America

United States of America

The US government, in hopes of maintaining its leadership in the field of nanotechnology, has founded the NNI to promote nanoscale science, engineering, technology research, and development programs. The NNI is a sub-branch of the National Science and Technology Council (NSTC), the cabinet level council by which the President coordinates science, space, and technology policies across the Federal Government. The program was established in 2001, and to date the cumulative investment is 12 billion USD. The proposed NNI budget for fiscal year 2010 is 1.64

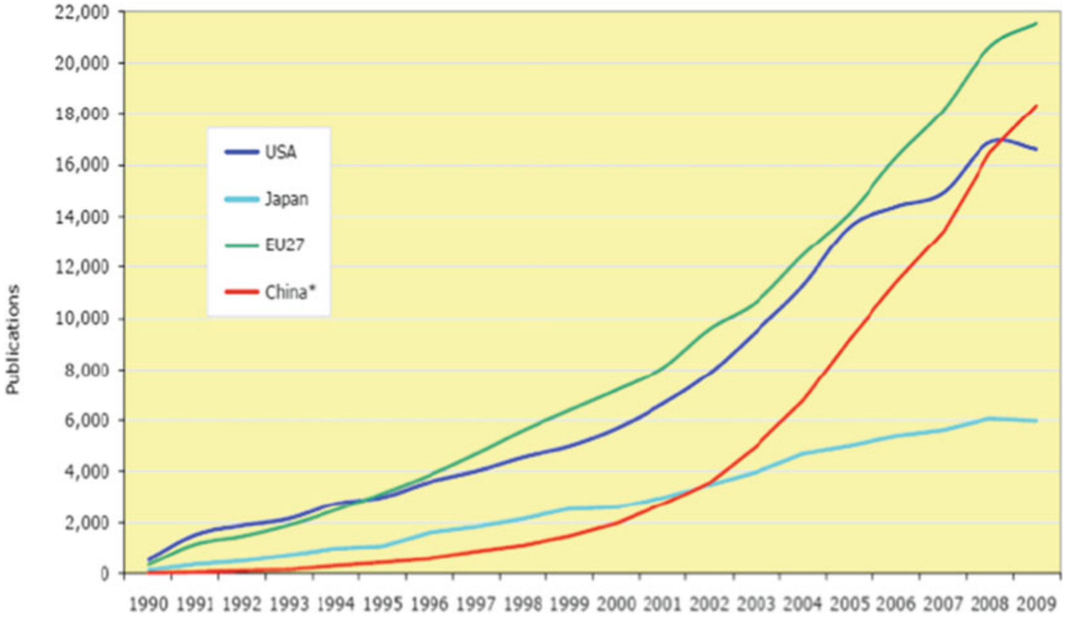


Fig. 9.2 Nanotechnology publications in science citation index by year. *Source*: Chen HC, Dang M, Roco MC. "Updated nanotechnology indicators, January 2010." Addendum to Chen H, Roco MC (eds). Mapping nanotechnology innovations and knowledge: global, longitudinal patent and literature analysis. New York: Springer, 2008

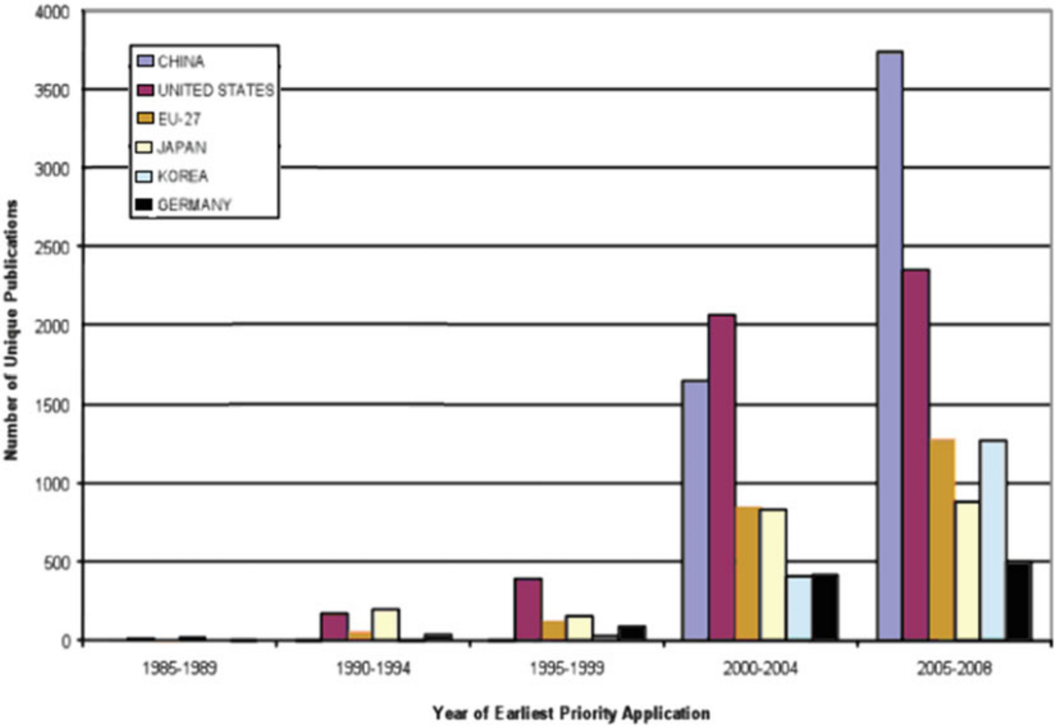


Fig. 9.3 Nanotechnology related patent applications. *Source*: Kisliuk B, USTPO, unpublished on comparative patent filings, January 2010

billion USD, with a growth of 8.5% over fiscal year 2009. Currently, the NNI involves 25 federal agencies, 13 of which had budgets for nanotechnology R&D in 2010. Table 9.1 displays NNI budgets (2008–2010) for the 13 budgeted federal agencies.

Of the total investment in the 2010 NNI proposed budget, 96% of the total is focused on five federal agencies, which are the National Science Foundation (NSF), the Department of Defense (DOD), the Department of Energy (DOE), the National Institutes of Health (NIH), and the National Institute

Table 9.1 National nanotechnology initiative (NNI) budget 2008–2010 (millions in USD)

Federal agencies	2008 Actual	2009 Estimate[a]	2010 Proposed
NSF	409	397	423
DOD	460	464	379
DOE	245	337	351
DHHS (NIH)	305	311	326
DOC (NIST)	86	87	91
EPA	12	16	18
NASA	17	17	17
DHHS (NIOSH)	7	7	12
DHS	3	9	12′
USDA (FS)	5	5	5
USDA (CSREES)	6	3	3
DOT (FHWA)	1	3	3
DOJ	0	0	0
Total	1,554	1,657	1,640

[a] 2009 Additional funds were allocated as a part of American Recovery and Reinvestment Act of 2009

of Standard and Technology (NIST). The NSF focuses on research across all disciplines of science and engineering; the DOD focuses on science and engineering research advancing defense; the DOE focuses on new and improved energy technologies; the NIH focuses on biomedical research at the intersection of life science and physical science; and the NIST focuses on fundamental research and development of measurement and fabrication tools, analytical methodologies, and metrology for nanotechnology [17].

Canada

Canada contributed 4.5% of the total global investment in 2008. The nanotechnology initiative by the Canadian government includes nine institutes of the National Research Council (NRC) distributed in several provinces. These provinces have spearheaded technological advancements in Canadian nanotechnology. Alberta, British Colombia, Ontario, and Quebec are some of the provinces central to innovation and competiveness in Canada. Currently, there are approximately 50–200 Canadian companies involved in the field of nanotechnology [18].

Brazil

Brazil, the leader of nanotechnology advancements in South America, commenced its governmental initiative in 2001. The National Council of Scientific and Technological Development (CNPq), the agency under the Ministry of Science and Technology (MCT), is dedicated to the promotion of scientific and technological research, human resources, and funding and is involved in the nanotechnological advancement in the country. In Brazil, funds for nanotechnology projects are distributed across various regions for the accepted proposals in hopes of stimulating the regional economy. In 2008, the responsible agency awarded 50% of its budget to the South-East, 15% to the South, and 35% of the budget to the North, North-East, and Center-West regions [19–21].

Europe

United Kingdom

Nanotechnology efforts in the United Kingdom were first initiated in 1986 by the National Initiative on Nanotechnology, led by the National Physical Laboratory. The initiative was jump-started with the Taylor report published in 2002, which highlighted the significance of nanotechnology. Currently, the Technology Strategy Board (TSB) is a governmental agency with a vision to be the global leader for innovation and attract inventive business. The Biotechnology and Biological Science Research Council (BBSRC), Engineering and Physical Science Research Council (EPSRC), and Natural Environment and Research Council (NERC) are the three research councils that drive nanotechnology advancement in the UK [22, 23].

Germany

The Federal Ministry of Education and Research (BMBF) has been funding nanotechnology activities since the late 1980s. The German government initiated a strategy to integrate seven ministries for the advancement of Nanotechnology in its "Nano Initiative-Action Plan 2010." The goal of the strategy can be summarized in five action priorities:

- Developing future fields and introducing new industries.
- Creating favorable framework conditions for science and economy.
- Recognizing risks and providing guidelines for responsible handling.
- Informing and integrating the public.
- Identifying research demands of tomorrow.

The seven federal ministries include the Federal Ministries for Labor and Social Affairs (BMAS), Environment, Nature Conservation and Nuclear Safety (BMU), Food, Agriculture and Consumer Protection (BMELV), Defense (BMVg), Health (BMG), Commerce and Technology (BMWi), and BMBF [24, 25].

Eurasia

Russia

The Russian government nanotechnology initiative includes the Russian Corporation of Nanotechnology (RUSNANO) with a five billon USD budget, specialized federal R&D program financing (2008–2011) for four billion USD, and Nanoindustry Infrastructure Development in the Russian

Federation for a period of 2008–2010 for one billion USD. The primary corporation, RUSNANO, was established in 2007 to advance the field of nanotechnology financed projects which include manufacturing projects to expand nanotechnological products; infrastructure projects to advance innovation, technology centers, technoparks, and information databases; and educational programs to train and educate the workforce for nano industries. The mission of RUSNANO is to make Russia the leading player in nanotechnology [26, 27].

Asia

Japan

Japan, one of the earliest players in the field of nanotechnology, leads the governmental effort through the agency MEXT (Ministry of Education, Culture, Sports, Science and Technology). MEXT is responsible for basic research and spearheads the nanotechnology advancement. The Ministry of Economy, Trade, and Industry (METI) leads the industrial development and commercialization efforts [28].

China

In 2001, an aggressive plan titled “Compendium of National Nanotechnology Development (2001–2010)” was issued to jump-start the Chinese initiative in Nanotechnology. The Chinese government spent 600 million USD in 2007 to promote nanotechnology with approximately 70 institutions engaging in nanotechnology research. The government backed funding agencies include the Chinese Academy of Science (CAS), National Science Foundation of China (NSFC), Ministry of Education (MOE), and Ministry of Science and Technology (MOST) [29, 30].

South Korea

South Korea’s aggressive plan to become one of the top three in the world by 2015 in nanotechnology competitiveness has led to the allocation of 16 billion USD for the period 2001–2010. The purpose of the Science and Technology framework law of 2001, launched by MOST, is to advance the nanotechnology effort in South Korea. MOST focuses on basic research and funds the knowledge base. The Ministry of Commerce, Industry, and Energy (MOCIE) and the Ministry of Information and Communication (MOIC) also promote nanotechnology growth in South Korea. The MOCIE focuses on commercialization efforts, whereas the MOIC funds projects that connect information technology with nanotechnology [31].

India

In 2001, India joined the race by launching the Nano Science and Technology Initiative (NSTI) under the Department of Science and Technology (DST), with a 20 million USD budget that financed infrastructure and 100 projects over 5 years. NSTI was superseded by the aggressive program Nano Mission within the DST with a budget of 254 million USD for 5 years. The objective of the Mission includes: Basic Research Promotion, Infrastructure Development for Nano Science & Technology Research, Nano Application and Technology Development Programs, Human Resource Development and International Collaborations. Other agencies involved in nanotechnology development include the Council of Scientific and Industrial Research (CSIR), the Department of Biotechnology (DBT), agencies under the Ministry of Information and Communication Technology, the Ministry of Family and Health and Welfare, the Ministry of Defense, and the Ministry of New and Renewable Energy [32–34].

Australia

Australia has made nanotechnology a national priority since 2004. A study conducted in 2004 by Australia Research Council (ARC) highlighted the shortcomings of Australia’s position in the global race. Numerous initiatives by the government were introduced to improve Australia’s status on the world stage. According to the Nanotechnology Australian Capability report, 75 nanotechnology research organizations, along with 80 nanotechnology companies, lead the nanotechnology effort. The government channels funding for nanotechnology through main agencies like the ARC, the Commonwealth Scientific and Industrial Research Organization (CSIRO), and the National Health and Medical Research Council (NHMRC). According to the Australian Office of Nanotechnology (AON), 21 of Australia’s 41 universities participate in nanotechnology advancements. For the period of 2008–2009, 107.7 million USD were allocated to nanotechnological advancement by participating governmental agencies [35, 36].

Nanotechnology Companies

Companies focusing primarily on nanotechnology products, as well as established companies with a nanotechnology subbranch, are on the rise. The drivers for private entities include return on profit, superior product/services, and monopoly. According to Nanovip International Nanotechnology Business Directory, in November 2008 there were 1,608 companies with nanotechnology products and services; however, this is only an estimate due to various standards as qualifiers in different countries. Currently, the US dominates the race for the number of companies compared to other nations [37].

Nanostructured Materials as Destructive Adsorbents

Nanostructured materials based on diatomaceous earth, carbon, zeolites, lime, and clays have been in use as adsorbents for many years to physically remove contaminants from surfaces [38]. Many of these materials were not initially recognized as being “nanostructured.” Nanostructured materials can exist as atomic, molecular, or crystallite clusters. They may form different shapes including, but not limited to, cages, sheets, tubes, spheres, and rods. In the process of forming these unique shapes, voids and defects are also formed which increase the nanostructured materials’ surface area and porosity through pits, holes, tunnels, and edges. These defects allow the nanostructured material to absorb liquids and trap gaseous molecules. The contaminants in the area are then removed by containment within the nanostructured materials’ overall structure.

Destructive adsorbents based on high surface area inorganic metal oxides act upon contacted contaminants in two ways: they contain the contaminant by trapping the liquid or gas into their structure, and they alter the original structure (destructive adsorption) such that the contaminants are immobilized. Typically immobilization is done through interaction with reactive sites on the adsorbent. The nanostructures’ inherent molecular structure discontinuity gives additional reactive sites such as terminal hydroxyls or coordination sites for anions and cations. Bonding of the adsorbate to the adsorbent through van der Waals forces usually begins the destructive process and instigates the reaction with the surface structures of the adsorbent.

Materials that combine these characteristics can be used in several ways, with the primary usages focused on air purification and destruction and immobilization of bulk quantities of hazardous chemicals [39]. Specifically, the high surface areas, increased porosity, large concentrations of surface defects, and unusual stabilized exposed planes are credited for these useful applications [40–42]. Several mechanisms can be facilitated by these surface structures including oxidation, hydrolysis, and elimination of appropriate functional groups from the adsorbed contaminant (Fig. 9.4).

Nanomaterials Based on Metal Oxides

Some of the materials that have garnered the most attention are based on metal oxides and zeolites. Each of these materials can be further modified with doping agents and functional groups to tailor the material to its intended use. Literature evidence for proving destructive adsorption is heavily directed towards high surface area metal oxides; therefore, metal oxides are the focus of this section. High temperature desorption by thermo gravimetric analysis (TGA) in combination with Fourier transform infrared (FTIR) or GC-MS show what heat releasable by-products are observable or key changes are seen in the reacted sorbent’s NMR peaks or FTIR spectral bands. Most available analyses rely on observed reduction in the challenge agent either through sorbent extraction or measurement of a gas stream before and after entry into the sorbent bed (flowthrough or breakthrough). In general, the best characteristics for a destructive adsorbent to have are high surface area, pore size, and volume in a mesoporous range for increased contact of adsorbate and a high number of surface hydroxyls to allow for a variety of species to interact with the surface groups [43].

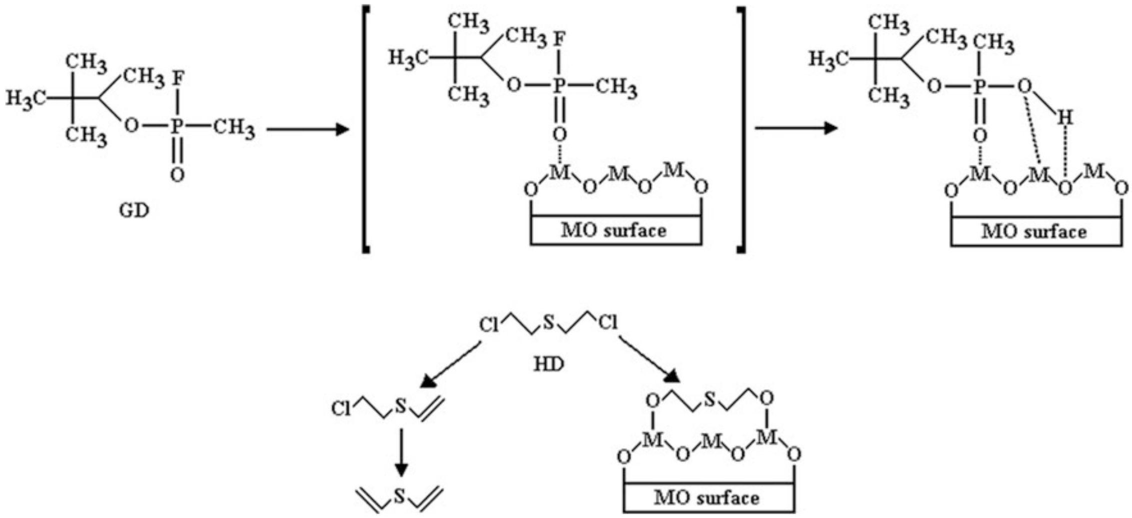


Fig. 9.4 Reaction schemes depicting destructive adsorption of GD and HD on nanocrystalline metal oxide surfaces

Nanocrystalline high surface area titanium oxide (TiO_2) has been extensively studied by several groups for its reaction with dimethyl methylphosphonate (DMMP), a common chemical warfare agent simulant for VX and G agents. Trubitsyn and Vorontsov examined ambient reactions of the anatase form by FTIR, comparing successive and simultaneous reaction processes of reactive adsorption, hydrolysis, and photocatalysis at both high and low DMMP concentrations, with and without humidity [44]. The rate of exposure to DMMP and relative humidity play a role in product formation and the degree of product binding to the surface of TiO_2 during adsorption and hydrolysis. The photocatalytic process is also affected by deactivation of sites by non-volatile products. Humidity aids in both hydrolysis and the release of converted volatiles. Interestingly, when all processes occur simultaneously, the destruction of DMMP is complete within 30 min.

Panayotov and Morris observed two paths that DMMP takes toward decomposition: molecular adsorption through hydrogen bonding at the hydroxyl groups and reactive adsorption through the Lewis acid active oxygen sites [45, 46]. They found that at low temperatures (295–400 K), nucleophilic attack from the neighboring surface hydroxyls form the Ti–CH_3 and P–O_x surface groups, while at higher temperatures (400–600 K) the lattice oxygens oxidize the Ti–CH_3. However, as the lattice oxygens are slowly poisoned and the non-volatiles accumulate, the isolated hydroxyls can be regenerated through thermal treatment in oxygen. Similar results were found during a molecular dynamics computer generated simulation performed by Quenneville et al. with varied surface hydroxylation of amorphous silica (SiO_2) and DMMP [47].

Li et al. have thoroughly investigated adsorption and decomposition of organophosphorus compounds with magnesium oxide (MgO) using an in situ pulse reactor GC-MS technique [39]. Comparison of two different MgO sorbent beds of varied surface areas (130 and 390 m^2/g), after exposure to various gaseous organophosphorus compounds, revealed several important facts related to decomposition temperature, product distribution, and capacity for destructive adsorption. Overall, MgO reduces the decomposition temperatures for these compounds via dealkylation reactions, proton abstractions, and nucleophilic displacements; the capacity for the reaction is dependent on surface area; and presence or absence of water affects product distribution.

Photocatalytic studies with high surface area TiO_2 and SiO_2-TiO_2 using the HD simulant, 2-chloroethyl ethyl sulfide (2-CEES), by FTIR show both partial and full oxidation products in addition to aldehydic, carboxylate, and carbonate products that are surface bound [48, 49]. Multiple reactive sites are used by TiO_2, while Si–OH sites hydrogen bond to 2-CEES through the sulfur and chlorine moieties. Similar studies with diethyl sulfide (DES) found other specific roles for the chlorine.

Studies with nano MgO and 2-CEES also found that liquid phase reactions are greatly affected by the solvent used. The most rapid results were observed from use of inert organic solvents with some addition of water to enhance the surface hydroxyl group numbers [50]. Nanomaterials that are prepared such that they are intimately intermingled with two or more nanomaterials can show synergistic performance compared to those prepared singularly. Studies with 2-CEES and mixed metal oxides found that the surface hydroxyls and the Lewis acid sites were responsible for conversion to the surface bound alkoxy species and vinyl products and swift reaction time [51, 52].

Carnes et al. explored the destructive adsorption reaction with an aluminum oxide (Al_2O_3) and MgO mixture for performance against diethyl 4-nitrophenyl phosphate (paraoxon) [53]. UV–vis and FTIR studies showed that very little paraoxon is released back to the environment after washing with solvent. Similarly, nano zinc oxide (ZnO), copper oxide (CuO), and nickel oxide (NiO) have been found to have increased reactivity and capacity towards paraoxon over commercially available forms [54, 55].

In a series of studies, Wagner et al. have explored the room temperature reaction of nanocrystalline MgO, calcium oxide (CaO), Al_2O_3, and nanotubular TiO_2 with real agents (VX, GD, and HD) by ^{1}H, ^{31}P, ^{27}Al, and ^{13}C NMR [42, 56–58]. All of these nanomaterials hydrolyze VX and GD to surface bound non-toxic compounds. With HD, both hydrolysis and elimination were observed by destructive adsorption modes. The many layers of nanotubular TiO_2 and the surfaces and pores of MgO, CaO, and Al_2O_3 trap water allowing for rapid hydrolysis, whereas the presence of water on CaO with HD creates a $CaCl_2$ coated surface, catalyzing fast elimination of HCl after an initial induction period. Rates of some of these reactions were found to approach the speed of liquid decontamination solutions.

Koper et al. have also studied the destruction of various chlorinated hydrocarbons over high surface area CaO (conventionally prepared—CP and aerogel prepared—AP) and compared to commercially available non-nano or micro material (CM) in a bed at different temperatures with and without water [59, 60]. Comparison of these materials of differing surface areas (CM ~10 m^2/g; CP ~100 m^2/g; AP ~120 m^2/g) led to several key conclusions. The material preparation method is critical, as it affects the surface area. Higher surface area generally results in faster reaction time, lower reaction temperature, and higher reaction capacity; however, this is highly compound dependent. While CP performs better with certain compounds, the reactions are slower with a higher capacity when compared to their AP prepared counterparts. Temperatures must be fine tuned for high efficiency and no graphite poisoning of the bed.

Nanocrystalline zinc oxide (ZnO), copper oxide (CuO), and nickel oxide (NiO) have also been found to have increased reactivity and capacity towards carbon tetrachloride (CCl_4) over the corresponding non-nano metal oxides [54, 55].

High surface area MgO and TiO_2 reactions with the halocarbons CF_2Cl_2 and $CFCl_3$ were found to have increased efficiency at lower reaction temperatures compared to their non-nano commercial forms. Interestingly carbon coating of these high surface area metal oxides also appear to improve the reactivity of the nano forms [61]. Addition of Fe_2O_3 or other transition metal oxides to high surface area nanocrystalline CaO was found to improve destructive adsorption activity against CCl_4 and C_2Cl_4 to near stoichiometric quantities at elevated temperatures [62]. Several polar organic compounds are also destructively adsorbed by nanocrystalline MgO with transition metal oxide shells such as vanadium oxide; however, powder forms perform better than pelletized forms [63].

Decker et al. and Khaleel et al. explored this application with Al_2O_3/MgO and other coated formulations against SO_2 (an acid industrial gas) [62, 63]. This intermingling of metal oxides shows that acid gas adsorption is also improved over its singularly prepared individual components. Vacuum desorption and FTIR studies show well bound monodentate SO_2 species on the surface of the powder. Similarly, nano ZnO, CuO, and NiO were found to have increased reactivity and capacity towards SO_2 over the corresponding non-nano forms [54, 55].

Room temperature FTIR studies of nanocrystalline MgO, CaO, and Al_2O_3 with volatile organics, such as aldehydes, ketones, amines, and alcohols, found high capacity adsorption of these materials through a multi-layer dissociative process at functional groups [64]. Similarly, acetaldehyde is adsorbed by nano TiO_2 and as temperature is slowly raised surface intermediates are observed by aldol condensation [65].

Reactions of H_2S with sol- or aero-gel prepared CaO, ZnO, Al_2O_3, strontium oxide (SrO_2), and MgO proceed with increased capacity over their non-nano formulations due to their higher surface areas as well as higher inherent reactivities. Nano ZnO and CaO work with high efficiency and the observed efficacy is attributed to nanocrystallites and unique morphology of the materials allowing for deeper reaction into the material [66]. High surface area SrO_2 was found to be a unique material which had greatly enhanced efficiency (stoichiometric) at RT, while at elevated temperatures exhibited reduced efficiency due to temperature induced crystal growth [67].

Studies of nanocrystalline MgO, CaO, Al_2O_3, TiO_2, and cerium oxide (CeO_2) with and without halogen (Cl_2, Br_2, and I_2) or interhalogen (ICl, IBr, and ICl_3) treatment have been carried out against *Escherichia coli*, *Bacillus cereus*, *Bacillus globibii*, aflatoxins, and MS2 bacteriophage [68, 69]. Vegetative cell studies found appreciable reduction of *E. coli* and *B. cereus* with halogenated MgO and CaO within minutes. Spore studies show some reduction of *B. cereus* by both halogenated and non-halogenated nano MgO and CaO within several hours. Colorimetric studies with aflatoxin show a neutralizing effect and MS2 studies found no plaque forming units after 5 min with some halogenated formulations. Atomic force microscopy (AFM) and TEM imagery illustrate significant changes in the cells or spores [70]. These biocidal properties are thought to be promoted by the abrasiveness, basic character, electrostatic attraction, and oxidation power of the wet slurries or dry powders of nanomaterials [71]. Other known antimicrobial materials may be incorporated into these metal oxides to provide additional biocidal properties. Photocatalytic and biological testing of nanostructured novel Ag–C–S–TiO_2 formulations shows both photodegradation of acetaldehyde and inactivation of *E. coli* and *Bacillus subtilis* spores without light activation compared to P25–TiO_2 [72].

Commercial Uses

The commercialization of nanotechnology using destructive adsorption has, so far, been limited to studies developed by one company, which has been particularly successful with this technology. NanoScale Corporation, located in Manhattan, KS, has been using nanocrystalline metal oxides for neutralizing toxins since 1995 for customers worldwide. The technology was originally developed at Kansas State University under Dr. Kenneth J. Klabunde. After years of continuous progress, FAST-ACT® (First Applied Sorbent Against Chemical Threats) was developed for the US Army for chemical warfare decontamination. FAST-ACT can be used on a variety of chemical releases to adsorb and eliminate chemical threats. The same technology was adapted to fit industrial and institutional needs for chemical hazards in the ChemKlenz® line of products. NanoScale offers products based on destructive adsorption of odors as well [73].

Destructive adsorption provides a superior technology to many alternatives since it does not release or mask contaminants. Guild Associates also uses adsorption technology in their material provided to the US Army [74]. Guild provides chemical and biological decontamination in various products. They have developed several types of mats and blankets, a protective barrier used for transport, wipes, laundry systems, and masks, among others.

Nanostructured Materials in Catalysis

Nanotechnology in the field of catalysis is ahead of other nanotechnology segments and is referred as “the engine that powers the world at the nanometer length scale” [75]. The exceptional morphology of nanostructured materials is

Table 9.2 Estimated impact of nanocatalysts [80]

Industry	Cost saving ($billion/year)	Energy savings (trillion BTU/year)
Chemicals	2.5–4.0	200–400
Petroleum	0.2–0.8	80–200
Automobile	0.2–1.1	NL

NL not listed
Source: chemical industry vision 2020 technology partnership

responsible for their enhanced catalytic properties [76]. Catalysts, by definition, accelerate chemical reactions by transforming reactants to products without being consumed during the process [77]. In the presence of an ideal catalyst, the preferred reaction can proceed at a faster rate, resulting in desired products and improved yields. The speed of the reaction is guided by catalytic activity, while the preferred reaction is a reflection of catalytic selectivity [78]. Catalysts at the same phase as reactants are known as homogenous and contribute to 10–15% of the catalyst market, whereas catalysts in a different phase are referred to as heterogeneous catalysts. In this topic we will focus on heterogeneous catalysts due to their extensive applications. Approximately 90% of newly developed chemical processes utilize catalysts, and improving catalyst efficiency via nanotechnology would have a significant impact on resources and the environment [79]. More specifically, improved process efficiency through use of enhanced nanocatalysts will allow for greener chemistry. Breakthroughs in energy and environmental applications with nanoscale catalysts have established great commercial success. For example, petroleum processing and the catalytic converter for automobiles utilize catalysts at nanoscale and are considered to be significant contributors [14]. Table 9.2 indicates cost and energy saving in the US per annum from exploiting nanoscale catalysts [80].

Refinery Industry

The refinery industry has successfully applied nanotechnology to improve the previously inefficient refining process by exploiting catalysis. Naphtha reforming and cracking are examples of unit processes where nanocatalysts have made major contributions [81]. Naphtha is used as a feedstock for the production of high octane gasoline by reforming and restructuring the hydrocarbons of low octane. Currently, bimetallic nanocatalysts based on Pt-Re/Pt-Sn on an acidic alumina substrate are commercially in use in the reforming units. Pt-Re on acidic alumina allows for longer catalytic activity with improved stability and Pt-Sn on acidic alumina has enhanced selectivity at low pressure. Metal moieties are the catalytic sites where dehydrogenation reactions occur, while acidic sites are required for isomerization, cyclization, and hydrocracking reactions [82, 83]. The improvements over the Pt-alumina monometallic catalyst introduced in 1950 have led to higher octane levels and lower usage of the precious metals. Additionally, previously used octane enhancers like lead and benzene are being phased out. Techniques to produce oil at a lower cost by minimizing the amount of catalyst by reducing the particles to nano-size have been widely successful in the petroleum industry. Over the last six decades, advancement in technology has resulted in the particle size decreasing from 100 nm to less than 2 nm. Materials with improved catalysis efficiency are obtained by synthesizing materials with uniform active sites, size, and shapes [84].

In the petroleum refinery, cracking is the largest volume process where crude oil of high molecular weight (high boiling point) is cracked to low molecular weight hydrocarbons like gasoline, diesel, or olefinic gases by the process of fluid catalytic cracking (FCC) [85]. Currently, nano-size catalysts based on zeolites are commonly used for catalytic cracking. Zeolite structures can be manipulated to 1 nm and provide an adjustable cage-like structure that allows specific molecular interactions for enhanced selectivity. AkzoNobel, ChevronTexaco, Engelhard, Exxon Mobil, and CRI International are major producers of zeolite catalysts [84].

Environmental Related

Air Purification

Nanocatalysts play a crucial role in air filtration applications related for odor removal and chemical remediation. Combustion engines, coal burning facilities, and chemical plants are all sources of toxic nitrogen, sulfur, and other volatile organic compounds (VOC) [86]. Nanocatalysts are designed for the specific polluting source for VOC removal. Metal oxides based on heterogeneous photocatalytic oxidation (HPO) in the presence of sunlight have been proven to degrade toxic organic pollutants to less toxic by-products by semiconductor photocatalysts in the presence of an energetic radiation source and an oxidizing agent [87–89]. For example, heterogeneous photocatalysts based on TiO_2 and ZnO are useful for capturing toxic VOCs by UV activation, similar to the water purification described in the next section. The energy of solar electromagnetic radiation should be larger than the band gap of the nanomaterials in order to activate the photocatalytic reaction. One such photocatalyst that can be activated by electromagnetic radiation from the sun is crystalline phase anatase TiO_2 which has a band gap of 3.2 eV at wavelength shorter than 387 nm [90, 91]. However, only 2.7% of the solar energy is at this region. To harvest more of the electromagnetic radiation from the sun, photocatalysts exploiting the visible light region are being explored [92]. Doping

Table 9.3 Cost analysis of chemical production utilizing nanocatalysts [80]

Chemical	Old selectivity (% of theoretical)	New selectivity (% of theoretical)	Base feedstock losses from theoretical (M lb/yr)	Feedstock losses with new selectivity (M lb/yr)	Feed stock cost ($/lb)	Feedstock cost saving ($M/y)	Base energy loss from theoretical (Q/yr)	Energy loss with new selectivity (Q/yr)	Net energy saving (Q/yr)
Sulfuric acid	99.5	99.75	476	320	0.115	18	0.0008	0.0006	0.0002
Ammonia	99	99.5	625	601	0	0	0.2944	0.2919	0.0025
Propylene	95	97.5	0	0	0	0	0.143	0.126	0.017
Ethylene dichloride	99	99.5	261	237	0.315	8	0.0203	0.0188	0.0015
Ethyl benzene	99	99.5	29	15	0.315	8	0.0203	0.0188	0.0015
Styrene	90	95	693	160	0.33	176	0.02	0.0173	0.002
Methanol	99	95.5	736	351	0.068	26	0.0061	0.0029	0.0032
Formaldehyde	91	95.5	736	351	0.068	26	0.0061	0.0029	0.0032
p-Xylene	70	85	841	0	0.33	278	0.0936	0.0648	0.0288
Terephthalic acid	90	95	660	660	0.35	0	0.0078	0.0061	0.0017
Butadiene	90	95	1,588	1,331	0.34	87	0.0806	0.0704	0.0102

Source: chemical industry vision 2020 technology partnership

with transition metal ions such as V, Mn, Cr, and Fe, or with nonmetals such as N, C, F, and P, semiconductor coupling with WO_3, CdS, or In_2O_3, and dye incorporated photocatalysts are a few methods to utilize a broader range of electromagnetic radiation from the sun [93–97]. For example, Lee et al. [92] reported carbon-doped titania with a pore size of 5–17 nm had a narrow band gap and was able to be activated at a wavelength of less than 550 nm. Similarly, Sathish et al. [98] reported a nano-photocatalyst of N-doped TiO_2 with activity in the visible region of the electromagnetic spectrum. These nanophotocatalysts are incorporated in curtains, blinds, or glass, where natural available UV sources are used to activate and purify air, as well as destroy pathogenic microorganisms [96, 99, 100].

Nanocatalysts in the catalytic converter and exhaust system of a chemical or coal burning facility have a different approach to neutralizing the exhaust air. These exhaust air streams are rich in NO_x, CO, and unburned hydrocarbons (UHC), which are greenhouse gases and are regulated by environmental agencies. A conventional catalytic converter utilizes precious metals such as Pt, Rh, and Pd to purify the exhaust stream. These metals tend to agglomerate into larger clumps upon exposure to hot exhaust gases, reducing the catalytic activity due to decreased surface area and thus requiring more of the expensive metal particles. Mazda Motor Corporation and Nissan Motor Company have successfully utilized metal nanoparticles in their catalytic converter systems for purifying exhaust emissions. Although their techniques are proprietary, the Mazda improved catalytic converter requires 70% less of the precious metal nanoparticles. Their nanosized particles are embedded in the substrate preventing agglomeration and enabling improved oxygen absorption and release rates for enhanced emission cleaning [101].

Chemical Industry

Nanocatalysts in the chemical industry are credited for maximizing synthesis efficiency, minimizing by-product formation, and saving energy by improving selectivity and activity. This is achieved by manipulating and controlling the size, shape, spatial distribution, surface composition, electronic structures, and stability of the catalyst materials. The detailed synthesis and mechanism of various nanocatalysts used in the chemical industry is beyond the scope of this report; however, economic benefits due to the use of nanocatalysts is briefly discussed. According to catalyst experts in Los Alamos National Laboratory, a $4 billion per year cost saving in production cost is possible with nanocatalysts. Table 9.2 depicts savings in production cost and energy savings utilizing catalysts at nanoscale in a common chemical manufacturing process. The data reported in Table 9.3 is based on the REMI model widely used in nonprofit institutions, universities, and US state governments [80].

Catalysts are widely used in all industries; however, the value of improved catalysts such as nanoscale catalyst is in exponential increase as the high energy demand, stringent pollution regulations, high process efficiencies, reduced supply of material, and low carbon footprint requirement are drivers for superior catalysts.

Commercialization

According to Global Industry Analysts, Inc., the global nanocatalyst market is estimated to reach six billion USD by 2015. A large part (about 743 million USD) is contributed to the refinery industry, but the environmental applications will be the fastest growing. Concerns about air pollution and depleted energy sources are drivers for the environmental applications. Other major markets include the food processing industry and the chemical industry, where catalysts at the nano size offer unique and valuable benefits [102].

Green Millennium, Inc., Corona, CA, is using nanotechnology in their photocatalytic air purifying technology. Their TiO_2, when exposed to light and water vapor, produces hydroxyl radicals and a superoxide anion which allow the oxidation of airborne VOCs and toxic organic matter into carbon dioxide and water. This reaction provides deodorization, air and water purification, and sterilization providing a cheap and low energy consumption technology for air purification. With this technology, windows and walls can become self-cleaning and anti-soiling. Green Millennium offers three TiO_2 coatings for use across industries such as transportation, environmental, medical, food, construction, manufacturing, and more [103].

NanoStellar, Inc., Redwood City, CA, develops materials containing precious metals for the power industry. They have created more efficient nano-engineered catalyst materials that reduce exhaust and increase catalyst effectiveness of precious metals by 25–30%. They stress the use of computation materials science, novel synthesis, and chemical engineering for their success in development.

NanoStellar's answer to the increasingly ambitious goals of countries to reduce emissions and pollutants are three products. NanoStellar promoted Platinum is available for low CO oxidation light-off temperatures and tunable NO oxidation characteristic applications. NanoStellar Pt:Pd 2:1 offers 25–30% better performance over traditional platinum only materials and can be used for light duty diesel vehicles. NanoStellar Gold™ improves hydrocarbon oxidation by 20% over Nanostellar Pt:Pd and is suitable for heavy-duty trucks and high efficiency diesel engines. NanoStellar products offer improved thermal stability, greater resistance to poisoning, and better light-off temperature [104].

Altimate Envirocare, Singapore, has created a PhotoCatatSmart Coating known as EnviroCare TiO_2 Photo-Compound Range. This is a spray coating that binds to many surfaces to provide a protective coating. Their TiO_2 Photo-Compounds, in reaction with light, produces "Super-Oxides" to destroy micro-organic substances through vaporization. This provides buildings, vehicles, textiles, and other surfaces with continuously clean surfaces. EnviroCare TiO_2 can be utilized in medical facilities to destroy infectious diseases like MRSA, *E. coli,* and others [105].

Nanostructured Materials in Environmental Remediation

Environmental pollution is a global concern. In the United States alone, there are approximately 1,244 US EPA Superfund sites—sites where the soil, air, and/or water are so polluted as to be an imminent health threat to anyone in the vicinity [106]. Environmental pollution may arise from innumerable sources and processes, both natural and anthropogenic. Wildfires and volcanoes pollute the air, as do many forms of transportation and industrial processes. Torrential rainfall washes topsoil and agricultural chemicals into watersheds. Seasonal winds blow pollutants far from their source into forests and grasslands. As pollution does not tend to stay confined to one area because of the natural patterns of air and water movement, pollutants released in one region may have dramatic impacts on another. Research by both the Commonwealth Scientific and Industrial Research Organization Marine and Atmospheric Research in Australia (CMAR) and the National Aeronautics and Space Administration (NASA) in the US has confirmed a link between air pollution and severe drought; a link which became more obvious after wildfires in the western US led to a drought in sub-Saharan Africa [107, 108]. More recently, there has been a great deal of debate about the process of "fracking," where high pressure fluids are injected into hydrocarbon-containing bedrock in order to liberate stores of oil and natural gas. In April 2010, the Cabot Oil and Gas Corporation was banned from further oil production in the state of Pennsylvania after an investigation showed that its operation in Dimock Township had contaminated the drinking water wells of 14 homes there [109]. April 2010 also saw the BP oil disaster, where 4.9 million barrels of oil were released into the Gulf of Mexico when the Deepwater Horizon drilling rig exploded [110]. Contaminants from such assorted sources as chemical spills, pharmaceuticals, fertilizer and pesticide runoff, abandoned mining and industrial sites, and airborne gaseous and particulate matter from various exhausts pollute the environment we live in and demand our attention for an efficient prevention strategy. Sustaining and enhancing air, soil, and water quality represent some of the most difficult challenges facing the global society in the twenty-first century.

Arsenic contamination of water is found worldwide, with incidents occurring in Thailand, mainland China, Taiwan, Bangladesh, India, Nepal, Argentina, Chile, and many sections of the U.S [111]. Arsenic is a toxic element that occurs naturally in soils, rocks, and groundwater. It enters drinking water supplies from natural deposits in the earth or from agricultural and industrial practices. Long-term drinking water exposure can cause serious health problems such as skin, lung, bladder, and kidney cancer. Many different nanostructured materials have been used to remediate As(III) contamination, including Fe-based materials (particularly nanoscale zero valent iron, or NZVI), high surface area alumina [112], and photocatalytic reactions with nanocrystalline TiO_2 [113].

Because of these diverse sources, causes, and types of pollution, there is a constant need for environmental remediation. Effective remediation strategies can render polluted air breathable, enable polluted water to support aquatic life

Fig. 9.5 Surface waters with heavy metal contamination. *Left*: uranium contamination; *right*: acid mine drainage and lead contamination (Images courtesy of the Center for Hazardous Substance Research, Kansas State University)

once more, and restore polluted soils to a more pristine state. Nanostructured materials are expected to play a role in these strategies, and some successes have already been reported in the literature. Because of their high surface area to volume ratio, nanostructured materials can be significantly more chemically reactive than their bulk counterparts. Tunable nanostructured materials designed to target specific pollutants in a specific setting (water, air, or soils) are also a possibility.

Water Remediation

Water pollution is a very common environmental problem. While surface water pollution is more prevalent, groundwater pollution is also widespread. A number of different nanostructured materials, many of them based on Fe(0) nanoparticles, have been shown to be effective against common water pollutants including halogenated organic compounds (HOCs) and heavy metals. Previously, it has been shown that granular iron can degrade many chlorinated compounds, including chlorinated aliphatics, chlorinated aromatics, and polychlorinated biphenyls (PCBs), as well as nitroaromatic compounds. Early studies by Gillham and O'Hannesin examined the utility of zero valent iron [Fe(0)] in the degradation of 14 chlorinated aliphatic hydrocarbons [114]. Fe(0) nanoparticles have been shown to actively degrade PCBs and trichloroethene [115]. Additionally, bimetallic nanoparticles of Pd/Fe, Pd/Ag, Ni/Fe, and Pd/Zn have been shown to destructively dechlorinate hexachlorobenzene [116–118].

Nanostructured Fe-based materials have also shown some capacity for removing inorganic metal contaminants, such as Cr(VI) [119], As(III) [120], and Pb(II) [121–123], from polluted water. Chromium contamination is a common result of the tanning industry and lead contamination results from activities such as mining and wildfowl hunting, while arsenic contamination is usually natural in origin. As ingestion of these metals may cause both acute and long-term chronic health issues, the ability to remove them from the water supply is sorely needed. Unfortunately, the removal of Cr (VI) proceeds poorly in the presence of bicarbonate, magnesium, or calcium ions. One study showed a 55–77% drop in removal capacity when these ions were present [124]. The removal of As(III) has also been shown to be impacted by the presence of certain ionic species [125, 126]. Figure 9.5 shows the typical appearance of surface waters with heavy metal contamination.

Probably the most common cause of water pollution is microbial contamination. A multitude of microorganisms commonly reside in water, both in surface water and in groundwater. While the water supply in developed countries is highly regulated and protected from sources of contamination, this is not the case in developing countries where people commonly obtain water from communal wells or surface bodies of water. In these cases microbial contamination from human and animal waste, as well as soil bacteria, is widespread. Contaminated water can spread cholera, polio, typhus, amoebic dysentery, hookworms, elephantiasis, and many other diseases. Fortunately, nanostructured materials have been shown to have some activity against many common waterborne diseases.

It has been reported that single or mixed metal oxide nanoparticles such as zinc oxide, copper oxide, aluminum oxide, or titanium oxide, incorporated into a filtration media containing a binder matrix, can destroy bacteria [127, 128]. These approaches are believed to offer advantages over the current process of drinking water purification, which relies upon the addition of disinfectants and sanitizers to kill microbes plus flocculants to settle out larger particles. In one study, the metal oxide nanocrystals were included in amounts ranging from approximately 0.1% up to about 10%

by weight, based on the entire filtration media. In another series of studies, it was shown that MgO nanoparticles are very effective biocides against Gram-positive and Gram-negative bacteria (*Escherichia coli* and *Bacillus megaterium*) and bacterial spores (*B. subtilis*) [70]. Other studies have shown that Ag-based materials are effective against viruses in drinking water [129].

Soil Remediation

Soil is essentially an agglomeration of mineral and organic matter, with pore spaces containing air, water, and nutrient solutions [130]. Specific soils develop under the influence of five different factors: Parent material, climate, living organisms, topography, and time [131]. The mineral portion may include both clays and mineral crystallites, while the organic portion consists of plant and animal residues and secondary products formed by bacterial action on the organic residues. Soil pollution, like water pollution, is relatively widespread and may be the result of both natural and anthropogenic practices.

Common soil contaminants include heavy metals such as arsenic and mercury; halogenated compounds such as PCBs; anthropogenic compounds such as estrones; and radioactive materials such as thorium [132]. The elemental contaminants, such as the radioactive materials and the heavy metals, may originate either from mining and manufacturing or may occur naturally as the result of weathering or leaching of minerals rich in those materials. For example, the minerals orpiment, lorandite, and smithite are rich in arsenic, and their weathering may release arsenic into the soil and water [133]. Historically, many metal-containing minerals have been valued by humans. Mining these minerals exposes even more of the naturally occurring mineral deposit to weathering and leaching thus accelerating soil contamination. The area of the Almadén del Azoque, Ciudad Real, Spain, where cinnabar has been mined since ancient times, is still heavily contaminated by mercury [134–136].

Non-elemental contaminants originate almost exclusively from human activity. Estrones, steroid hormones which are endocrine disrupting chemicals (EDCs), find their way into the environment as metabolites from birth control pills and fertility medications. They have a strong affinity for humic acid which leads them to bind to dissolved organic matter and be deposited when the matter drops out of solution [137]. PCBs were widely used as coolants and lubricants in capacitors, transformers, and other electrical equipment because of their nonflammable and insulating properties. While PCBs are no longer manufactured in the US, they are still present in old transformers and capacitors and are pollutants in soils and sludge in more than 400 US sites [138].

Many of the same nanostructured materials which are useful for water remediation also have some utility for soil remediation. Free iron and iron-based nanoparticles have been demonstrated to be highly effective at removing heavy metals, chlorinated organic solvents, polyaromatic hydrocarbons, and PCBs; however, they tend to aggregate rapidly and demonstrate poor mobility in porous media [139, 140]. Stabilized Fe(0) nanoparticles, which can be produced by a variety of methods including by surfactant modification or in colloidal solution with activated carbon or colloidal poly(acrylic acid) remain unagglomerated in solution for extended periods of time, which increases their mobility in contaminated soils, and are as reactive as free nanoparticles [141, 142]. There have also been efforts to incorporate these materials into membranes which can be placed in the path of contaminant plumes [143]. One study of removal of As(III) by NZVI found that the reaction resulted in the rapid formation of the minerals magnetite and lepidocrocite, which arose from corrosion products of the NZVI and adsorbed As(III). At pH values ranging from 4 to 10, between 88.6 and 99.9% of the aqueous As(III) was removed from solution. The presence of NO_3^-, SO_4^-, or HCO_3^- negatively affected the ability of the NZVI to adsorb As(III), as did the presence of H_4SiO_4 and $H_2PO_4^{2-}$[125, 126].

Pd(0) nanoparticles are also used for soil remediation as both free nanoparticles and as supported nanoparticles, where the supporting substrate may be various solid foams. Free Pd(0) nanoparticles produced by the bacteria *Shewanella oneidensis* were shown to reductively dechlorinate PCBs in solution, and supported Pd-Fe nanoparticles have been shown to assist in reductively dechlorinating 1,2,4-trichlorobenzene. In one such experiment, PCB 21 (2,3,4-chloro biphenyl) solubilized in M9 microbiological media at a concentration of 1 mg/L was reduced to undetectable levels within 1 h when 500 mg/L of palladized *S. oneidensis* (bioPd) was added to the mixture and incubated at 28°C. For the most part, Pd(0) nanoparticles appear to resist the agglomeration and oxidation problems which plague Fe (0) nanoparticles [144].

Nanostructured TiO_2 may be used to remove radioactive Th from soils. Experiments conducted in soils in the presence or absence of soil humic acid and fulvic acid demonstrated that both the fulvic and humic acids increased the sorption of Th(IV) to TiO_2 nanoparticles at acidic pH. Bare TiO_2 nanoparticles by themselves were able to form surface complexes with 94% of the available Th(IV); this percentage increased to 97–98% in the presence of fulvic or humic acids and remained stable with increasing pH [145].

Air Purification

Many VOCs, nitrogen oxides (NO_x), and sulfur oxides (SO_x) in air contribute to smog and high ozone levels, which harm human health [146]. Some of these, such as SO_x, may be produced both naturally and anthropogenically, while others are primarily the result of human activity.

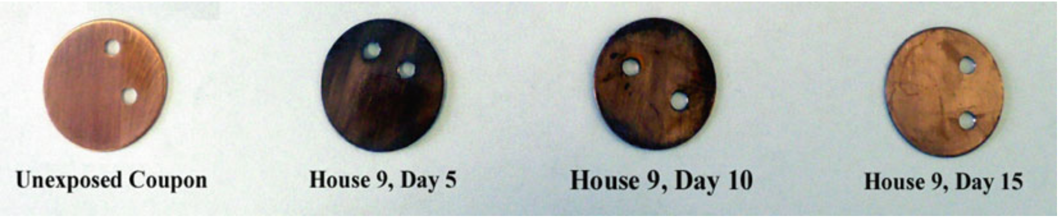


Fig. 9.6 Copper coupon corrosion (unexposed coupon = negative control, day 5 = baseline prior to treatment, day 10 = after first 5-day treatment, day 15 = after second 5-day treatment)

Nanostructured materials which are being used to purify the air include catalysts, which are currently in use and constantly being improved upon [147, 148], and nanostructured membranes, which are under development [149, 150].

Historically, carbon-based adsorbents or destructive oxidation have been used for air purification and particularly for VOC removal. More recently, nanoparticles of metal oxides (MgO and CaO) and core/shell binary oxides (Fe_2O_3/MgO or V_2O_3/MgO) have been tested against typical air pollutants such as acetaldehyde, propionaldehyde, perfluoropropene, and a number of other polar organic compounds [63]. MgO or CaO having a monolayer coating of Fe_2O_3 show enhanced reactivity for the destruction of chlorocarbons, organophosphates, and acid gases [152]. This destructive adsorption, where toxic gases are broken down into non-toxic or less-toxic products, is more desirable than the simple adsorption capability offered by carbon filters.

TiO_2 is widely used for photocatalysis of pollutants, as it generates hydroxyl radicals under UV radiation and the generated hydroxyl radicals then oxidize pollutants on the catalyst surface to form less harmful products. For example, oxidation of NO_x forms NO_3^- on the catalyst surface [153], meaning that TiO_2 has the potential for removing NO_x from polluted air [154]. In another study, it was found that mesoporous γ-MnO_2 displayed a good performance in the removal of NO_x (72 mg/g) and SO_x (700 mg/g) [155]. Nanocrystalline CaO was shown to react with SO_2 at a relatively low temperature to generate a mixture of calcium sulfite, calcium sulfate, and calcium sulfide [40]. The presence of a small amount of Fe_2O_3 on the surface of CaO enhances the ability of the CaO to act as a destructive adsorbent for SO_2.

Commercial Uses

According to a Bharat Book report, environmental remediation accounted for the largest end-user market, capturing 56% in 2007. The projection, however, is that environmental applications will decline through 2013 [156]. However, some technologies have already entered industry use.

To date, a number of businesses have begun utilizing nanotechnology for environmental remediation. Golder Associates, founded in Toronto, Canada, has been in the remediation industry for almost 30 years. They utilize NZVI to treat organic contaminants. The large surface area and high reactivity provide a cost efficient solution for remediating deeply contaminated sites. By injecting the NZVI into the subsurface, the chlorinated organics can be broken down. Other potential applications include treating sludge from various industries and polluted soils in situ instead of excavating them [157].

NanoH_2O, Inc., Los Angeles, CA, uses nanotechnology to desalinate and reuse water. Its Thin Film Nanocomposite membrane has specific characteristics including the membrane roughness, hydrophilicity, and surface charge. This inhibits the adhesion of bacterial cells leading to less biofilm coverage. The minimization of biofilm formation decreases the energy consumption and chemicals used making the reverse osmosis process more productive [158].

NanoStellar, Inc. is helping automakers bring diesel emissions levels down by using a unique nanotechnology. NanoStellar uses the surface chemistry of nanomaterials to engineer precious metals such as platinum, palladium, and gold on a nano-scale level in order to make more effective catalysts for the chemical reactions in emissions control. It is also a key supplier in the diesel industry and is expecting a growth in business due to the worldwide need for reduced diesel emissions [104].

NanoScale Corporation produces and markets nanomaterial-containing filters for improving indoor air quality in houses affected by contaminated drywall (which emits corrosive hydrogen sulfide and other air pollutants). Use of the filters has been shown to dramatically reduce copper corrosion and odor issues in affected houses. Figure 9.6 shows the copper corrosion in affected houses before, during, and after treatment.

Finally, the Dow Chemical Company offers a water treatment media called ADSORBSIA. This treatment media is used by many municipalities to remove dangerous levels of arsenic from water. It offers a number of advantages over conventional water treatment media including handling and disposal advantages [159].

Environmental Concerns Related to Nanomaterial Usage

While nanomaterials offer a great deal of promise in environmental remediation, their usage also comes with some environmental concerns. The impact of nanomaterials on the microbial populations of bodies of water, including surface water and groundwater, is not well studied. The impact of nanomaterials on the microbial populations of different soils has also not yet been thoroughly explored. Finally, questions have arisen about the potential health effects of nanomaterial exposure on wildlife and plants. Preliminary studies indicate that the microbial populations of aquifers change rapidly in response to nanomaterial exposure, but revert to baseline shortly thereafter [160]. NZVI particles also do not appear to have adverse effects on the germination and survival of certain plant species [161]. However, they do appear to have adverse effects on fish, particularly on the gills and intestines [162]. Their effect on larger animals and environmentally sensitive animals, such as frogs and salamanders, is currently unknown.

In conclusion, the environmental remediation possibilities offered by nanomaterials are extensive. However, as nanomaterials themselves may cause environmental damage in certain situations, it is imperative that the decision to use nanomaterials in the environment be made responsibly and with consideration to potential harmful effects. Companies such as NanoScale Corporation, located in Manhattan KS, have undertaken toxicology testing in response to these conditions; however, many nanomaterial producers have yet to follow suit.

Nanostructured Materials in Textiles

Nanotechnology in textiles is a powerful combination; integrating materials manipulated at atomic to molecular levels results in textiles with superior properties. Previously, the functions of textiles were traditionally limited mainly to clothing, comfort, decoration, and protection from temperature extremes. Advancement in nanotechnology, particularly in the field of textiles, led to novel functions and superiority over former materials. These novel functions include anti-static, water repellency, wrinkle resistance, UV protection, self-cleaning, protection against toxic chemical, biological and other pollutants, impact resistance, and protection against fire [163]. Textiles using nanotechnology have also advanced to provide superior protection against extreme weather and improved comfort in terms of weight, moisture transport and permeability, plus superior tear and puncture resistance [164]. The market for nanotechnology in textiles will reach $115 billion in 2012 [165]. In the following sections, highly specialized applications of nanotechnology in textiles are presented.

Chemical Protection

A chemical and biological (CB) protective suit is a necessary safety requirement for individuals who may come in contact with toxic chemicals and bioactive agents. Farmers, chemical plant employees, firefighters, soldiers, and all individuals who will likely come in contact with toxic chemicals or biological agents require a safety barrier. In the contemporary world, no market driver is as important as human life or safety. Exposure to hazardous chemicals and bioactive agents often leads to grave expenses, fines, and negative public relations. The unfortunate outcome of the 2001 anthrax attack in the US, the 1995 sarin gas attack on the Tokyo subway, the Bhopal disaster in 1984 due to the methyl isocyanate (MIC) gas leak, and 1980s Iran–Iraq chemical warfare are some of the more recent ill-fated events that have cost human life and suffering [166, 167]. Currently, two main categories of CB protective suits are available: impermeable polymer ensemble and permeable sorbent-based gear. Although impermeable polymer provides the most protection against CB threats, they cause a great amount of physiological strain. Productive time for the wearer is limited during strenuous activity due to heat stress [168]. The lack of moisture exchange can be detrimental to health causing various illnesses and in some cases death [169]. In sorbent-based technology, sorbent is sandwiched between the textile layers to create a composite laminate. This shields the wearer from the ill effects of toxic contaminants. Figure 9.7 shows a schematic of the concept behind sorbent-based CB protective clothing.

Activated carbon spheres have been the traditional sorbent incorporated between the textile layers to provide chemical protection. The carbon-based technology predominantly involves physical entrapment of toxins or chemical pollutants [170]. The toxin's proximity to the wearer, the potential of preferential adsorption of water by carbons, and a change in temperature resulting in off-gassing of adsorbed toxins are the major disadvantages of carbon-based chemical protective clothing (CPC). In sharp contrast, highly reactive nanocrystalline metal oxide sorbents integrated with textile matrices are expected to offer significant improvement in terms of true protection [171]. These high surface area sorbents with neutralizing capability decontaminate the toxic chemicals to non-toxic by-products.

NanoScale Corporation, as a commercial supplier of nanocrystalline metal oxides, markets its products under the NanoActive® brand name. Their nanocrystalline metal oxides are fine powders, with particles in the 1–10 μm range, similar to ordinary metal oxides, such as talcum powder. However, closer examination shows that they consist of very small crystals (usually 2–10 nm in size) that form large clusters. Figure 9.8 shows a photomicrograph of NanoActive TiO_2.

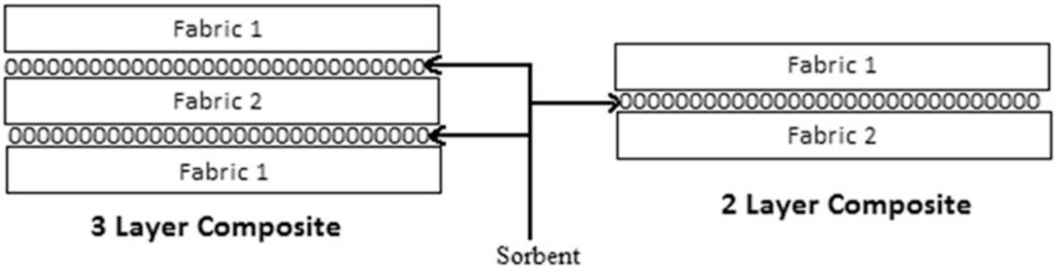


Fig. 9.7 Schematic of CB protective clothing liner

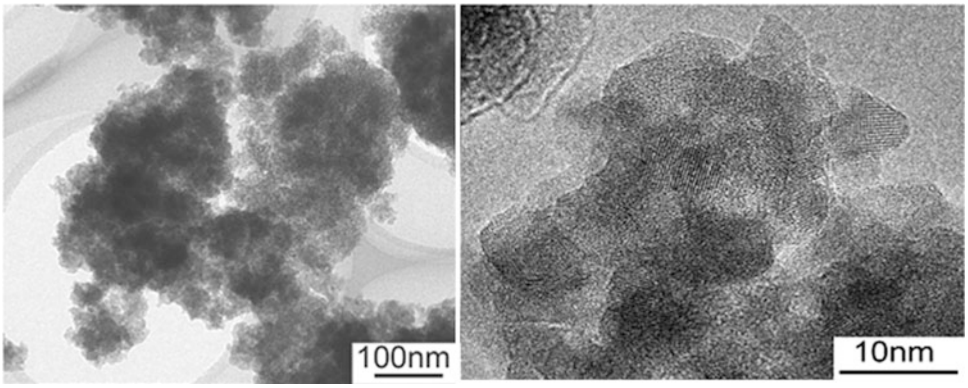


Fig. 9.8 Photomicrographs of NanoActive TiO_2. High-resolution photo (*right*) shows amorphous and nanocrystalline regions

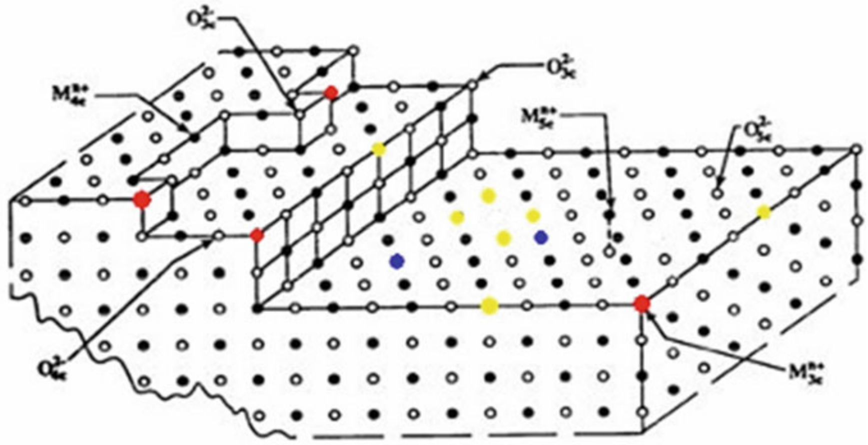


Fig. 9.9 Conceptual sketch, showing coordinatively unsaturated sites, as well as a vacancy defect site

NanoActive metal oxides are highly porous, with high surface areas (for example, NanoActive TiO_2 typically has a surface area exceeding 500 m^2/g). However, for reactivity, even more important than particle size and surface area is the presence of edges and corners, in which ions have a reduced coordination number, as conceptually illustrated in Fig. 9.9 (blue dots are examples of pentavalent surface ions, yellow dots are tetravalent edge ions, and red dots are trivalent corner ions). Such unusual morphology results in enhanced chemical reactivity and suggests a two-step decomposition mechanism on nanocrystalline metal oxides (the first step being adsorption of toxic agent on the surface by means of physisorption, followed by a second step, chemical decomposition).

The fate of chemical warfare agent simulants 2-CEES and DMMP was studied after permeation testing of nanomaterial incorporated textile swatches by ASTM F739. The performance of prototype reactive liners incorporated with nanocrystalline metal oxides was compared against carbon laminate. After the permeation testing, fabric laminates were extracted with ethyl ether and analyzed by gas chromatography-flame photometric detector (GC-FPD). Table 9.4 shows the outcome of permeation testing of 2-CEES exposed liners. As seen from the data, production of 2-hydroxyethyl ethyl sulfide (HEES) confirms destructive adsorption by nanocrystalline sorbent incorporated reactive liners.

Table 9.4 Gas chromatography-flame photometric detector (GC-FPD) analysis of textile swatches after permeation testing with 2-CEES

	% Extracted			2-CEES, (% permeated)
Sample	2-CEES	EVS	HEES	
Carbon laminate	92	0	0	23
NanoScale laminate 1	1.4	<1	18	2.8

Table 9.5 GC-FPD analysis of textile swatches from permeation testing of dimethyl methylphosphonate (DMMP)

Sample	DMMP (% extracted)	MPA	DMMP (% permeated)
Carbon laminate	85	Absent	6.2
NanoScale laminate	24	Present	0.60

Multiple analyses were performed to determine the fate of DMMP from the reactive liners. Solvent extraction of the liners retrieved from the permeation testing yielded unreacted DMMP. As seen from the data in Table 9.5, there was less extractable DMMP from the NanoScale prototype liner compared to the carbon control. The evidence for destructive adsorption of DMMP was acquired by TGA of sorbent material after permeation testing, which confirmed the presence of methyl phosphonic acid (MPA).

Figure 9.10 displays the TGA coupled with a 3D FTIR spectrum for DMMP adsorbed on nanocrystalline metal oxide. The band at 3039–2985 cm^{-1} due to off-gassing at 483–600°C is assigned to CH$_3$-P asymmetric stretch, a characteristic band for MPA seen with NanoScale sorbent, but not with carbon-based sorbent (Fig. 9.11). This band was confirmed by analyzing the neat MPA TGA profile (Fig. 9.12). From the comparison of Fig. 9.10 against Fig. 9.12, it is clear that the DMMP challenged nanocrystalline formulation generates MPA.

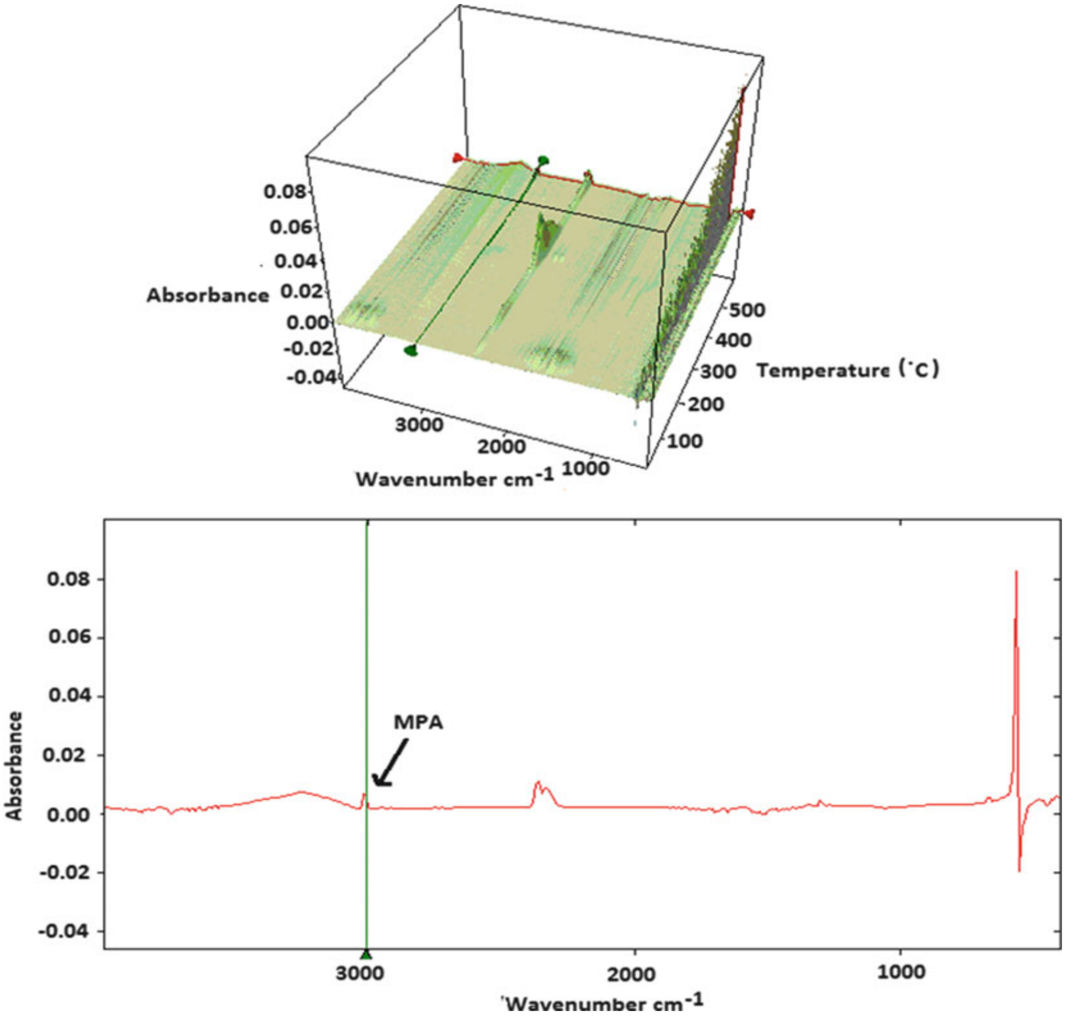


Fig. 9.10 Thermo gravimetric analysis (TGA) of dimethyl methylphosphonate (DMMP) adsorbed on nanoscale sorbent

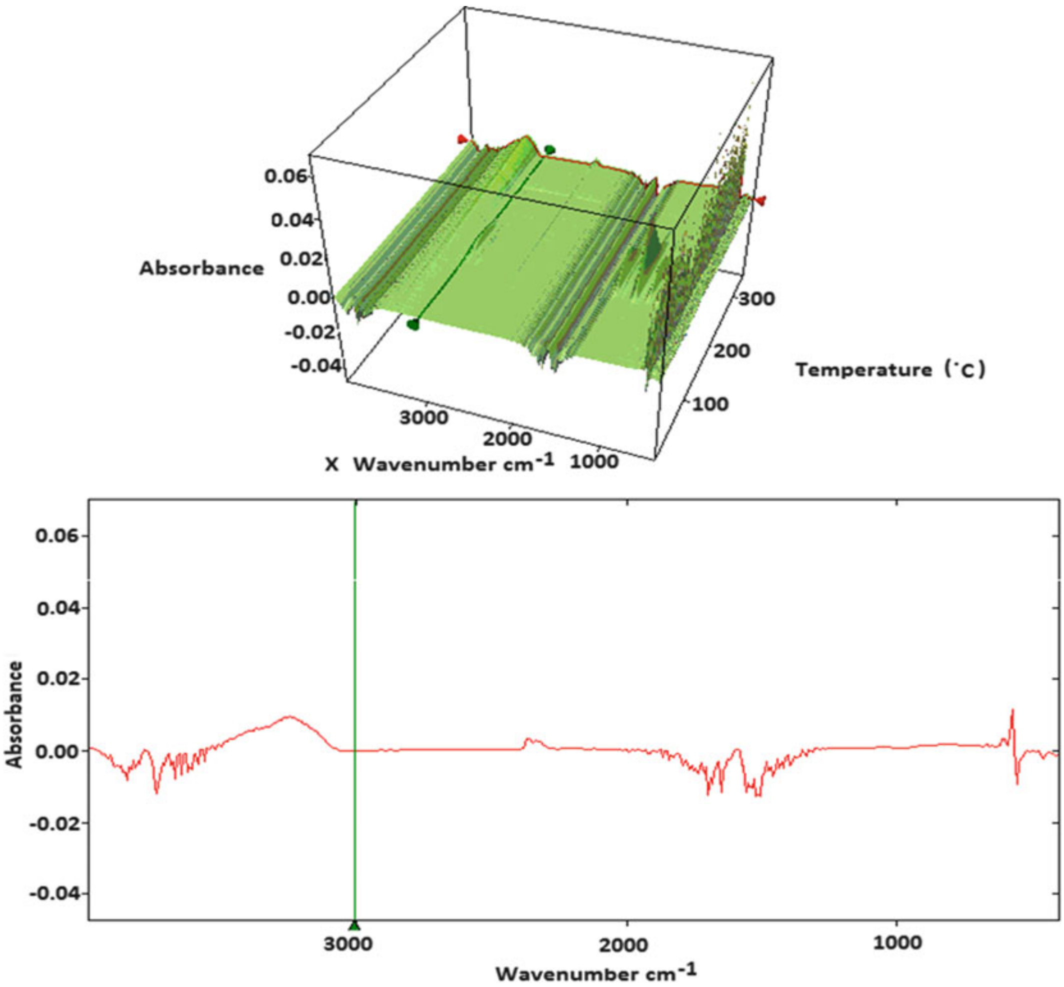


Fig. 9.11 TGA analysis of DMMP adsorbed on carbon

Biological Protection

Nanocrystalline metal oxides have also shown activity against biological agents. Table 9.4 displays the biological activity of modified nanocrystalline metal oxides against *B. subtilis* spores and vegetative cells of *Burkholderia gladiolii*. Sporicidal testing was conducted according to ASTM 2414-05 "Standard test method for quantitative sporicidal three-step method (TSM) to determine sporicidal efficacy of liquids, liquids sprays, and vapor or gases on contaminated carrier surfaces." Biocidal activity was determined as per ASTM E 2149-01 "Standard test method for determining the antimicrobial activity of immobilized antimicrobial agents under dynamic contact conditions" [172].

Textile industries in the field of medicine have made significant progress by incorporating nanotechnology in the area of antibacterial textiles, antimicrobial wound dressing, and anti-adhesive wound dressings [173]. Antibacterial textiles have broad applications in the medical field; patient's clothing, hospital bedding, facial masks, and hospital furniture are a few examples. Wound dressings exploiting nanotechnology have also become commonplace. Commercial wound dressings with incorporated nanocrystalline silver release ionized silver, facilitating wound healing. They have been clinically tested on burn wounds [174], ulcers, and other wounds [175]. Acticoat®, marketed by Smith & Nephew, Hull, UK, is an example of this class of wound dressings. Wound dressings based on a silica nanosol derived from long chain alkyltrialkoxysilane function as anti-adhesive bandages with improved moisture permeability. These cover the wound without causing further irritation, unlike the traditional wound dressing, which adheres to the wound. The anti-adhesive dressings accelerate the healing process and provide additional comfort.

Self-Cleaning Fabrics

Self-cleaning fabric, a property made possible by exploiting nanotechnology, is useful for consumers ranging from

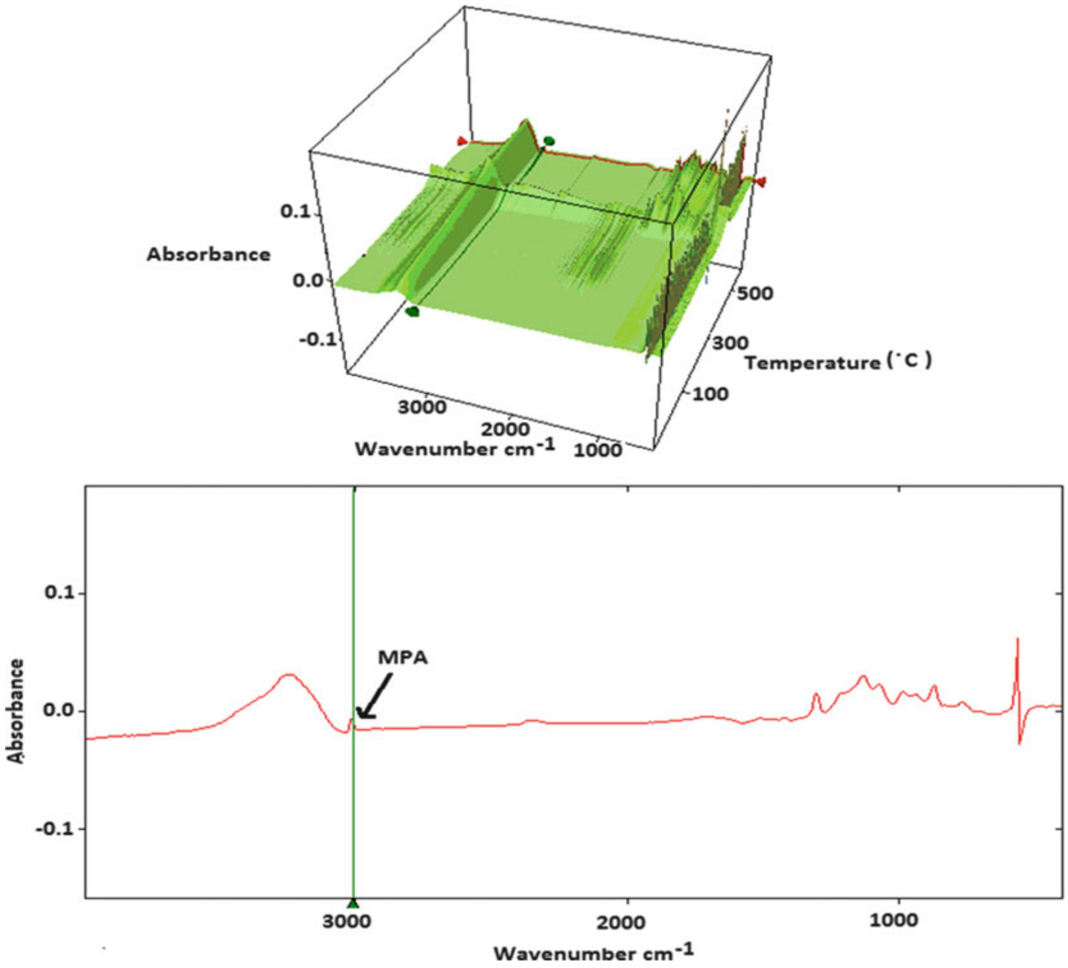


Fig. 9.12 TGA analysis of neat methyl phosphonic acid (MPA)

soldiers in the battlefield to toddlers at home. In the US, approximately two billion kg of surfactant are used annually. The laundry process requires significant energy, water, and detergent, causing greater energy usages and contamination of water streams from the use of detergent [176]. Nanotechnology is expected to make a promising contribution to this area, in the production of self-cleaning fabric via green chemistry. These fabrics are an improvement not only because they have a positive impact on the environment, but also because their usage saves time and effort. Self-cleaning fabrics can be further classified into those which act via photocatalytic activity and those which act via the Lotus effect.

The photoactivity of nanomaterials, in particular metal oxides with high surface area, is useful in decontaminating toxins and neutralizing odors. In the presence of sunlight or UV, electrons jump from their valance band (lower energy state) to their conduction band (higher energy state), leaving a strong oxidizing site in the valance band and creating a reduction site in the conduction band. These sites are reactive to adsorbed species. These reactive metal oxides, in particular TiO_2 coated on fabric, exist as Ti(IV) with an oxygen atom. When charged with UV radiation, the Ti(IV) cation is converted to Ti(III) by electron transfer and simultaneously reacts with the adsorbed species using the oxygen. This electron hole weakens the titania and oxygen bond, resulting in detachment of oxygen containing the adsorbed species. This site is further hydroxylated, and later oxidized back to Ti(IV) in a slow process in the absence of UV radiation. Thus, the fiber coated with TiO_2 acts as a catalyst [177].

The lotus effect is a phenomenon utilized by the Lotus plant (*Nelumbo nucifera*) to self-clean its leaves by exploiting the property of contact angle and sliding angle. The contact angle of a lotus leaf can reach 160 °, which is classified as a superhydrophobic surface. Lotus leaves have a rough surface due to structures called micropapillae which are 5–9 μm, and nanostructures within the micropapillae. When a water droplet contacts the leaf surfaces, it sits on the nanostructure as the air bubbles occupy the crevices in between nanostructures, thus minimizing drag for the droplet allowing the droplet to roll off the leaf [178].

Table 9.6 Biological activity of modified nanocrystalline metal oxides[a]

Modified nanocrystalline metal oxide formulations	Spore log reduction	Bacterial log reduction
[Ag]NA–TiO_2	3.8	8.3
[Ag]NA–TiO_2+ [Ag]NA–CaO	3.3	4.2
[Ag]FS–TiO_2+ NA–CaO+ NA–TiO_2	>3.0	2.6

NA nanoactive; *FS* fisher scientific
[a]NanoScale Corporation, internal research

This droplet removes dirt or particulate matter on the surface during the roll off making the surface clean. Mimicking this natural phenomenon, advancements in the area of superhydrophobic fabrics with high surface roughness have been made by using coatings with nanoparticles [179]. Carbon nanotubes on cotton fabric [180], fluorinated carbon [181], zinc oxide nanorod film [182], titiania [183], and silica [184] on polymeric substrates are a few examples of materials utilized to enhance superhydrophobicity.

Smart Textiles

Sensors using nanoscale systems have gained momentum in recent years. Smart textiles with sensors have become a complex tool monitoring the environment, health of the wearer, and gauging the life of the fabric. The unique characteristics of nanoscale materials are ideal for sensing physical, chemical, and biological signals. Sensors are widely used in present technology, and advanced sensors based on nanostructured metal oxides enable sensing of much lower signals at room temperature [185, 186]. Electrical conductivity, catalytic activity, high crystallinity, and high surface area of these metal oxides are properties required for sensing toxic gases in low levels [187]. Examples of metal oxides include: SnO_2, ZnO, TiO_2, WO_3, In_2O_3, Fe_2O_3, CuO, NiO, Ga_2O_3, and V_2O_5. SnO_2 is sensitive to all gaseous species, which makes it an excellent substrate for surface doping. This functionalized material enhances gas selection and sensitivity. Sensors based on nanostructured metal oxides are thermally activated; however, textile incorporation requires sensors operating at room temperature which limits the selection of metal oxides. SnO_2-based sensors for sensing NO_2, H_2, and CO function at room temperature. Similarly, ZnO-based sensors are reactive to H_2, H_2S, and NH_3 at room temperature [188]. Lim et al. outlines the mechanism of hydrogen gas sensors based on ZnO nanorods [189]. Oxygen molecules adsorb to the surface of ZnO nanorods and withdraw electrons from the conduction band to form the electron rich oxygen species O_2^-, O^-, and O^{2-}, which chemisorb at room temperature. This results in an electron poor space at the ZnO nanorod surface and creates resistance; however, when exposed to hydrogen gas, adsorbed oxygens are reduced to release electrons. This results in an electron rich space at the surface allowing decreased resistance. The sensitivity of this ZnO nanorod sensor can be improved greatly by doping with Pt or Pd metal which enhances the catalytic dissociation of H_2 to atomic hydrogen and eases the reduction of chemisorbed oxygen. These nanostructured metal oxides, when incorporated into textile matrices, are excellent indicators for exposure to toxic gases. Physical sensors based on nanoscale systems integrated into textiles which convert mechanical strain to electricity have multiple functions [190]. For example, CNT-based fabric sensors monitor respiration signals by measuring electrical resistance from the strain caused by respiration movements. Similarly, piezoelectric materials based on CNT-based fabric sensors generate electrical energy by sensing strain caused by movement [191]. Carboxylic acid and lactate oxidase functionalized single carbon nanotube (SCNT) are used to detect pathological disorders by sensing the pH and lactate present in excreted sweat. These biological sensors are excellent diagnostic tools [192]. Although incorporation of these physical, chemical, and biological sensors are at the initial stage of development, these smart textile-based nanoscale sensors are bound to change textile industry in the near future.

Textiles exploiting nanotechnology have long strides to make before their full introduction to the market place; these improvements include the durability to be washed and worn repeatedly without compromising functionality, large scale development needs, cost, and safety concerns. However, many textile products utilizing nanotechnology are commercially available and a few examples are listed below [193].

Commercial Uses

Nano-Tex®, based in Oakland, CA, uses nanotechnology to transform the molecular structures of fibers to produce fabrics that resist stains, moisture, odors, and static. According to Nano-Tex, more than 80 textile mills worldwide use Nano-Tex treatments in products sold by more than 100 leading apparel and commercial interior brands. Nano-Tex claims their treatment is permanently attached to fibers to improve the function of the garment without limiting the life [194].

NanoHorizons Inc., in Bellefonte, PA, uses technology based on silver nanoparticles less than 15 nm in size, incorporated into fabrics. A patented technology uses the antimicrobial properties of silver and their nanotechnology capabilities to combat microbial growth. NanoHorizons is able to permanently chemically bind molecules to fiber, foam, plastic, or coatings. NanoHorizons markets their technology under the brand SmartSilver® [195].

Although great advancements are being made in the textile industry, some of the most promising technology is still being developed. There is a substantial benefit to not having smelly, wrinkled clothing, but nanotechnology can do so much more. Sensatex Inc., Bethesda, MD, is developing what they have called the "SmartShirt."™ The SmartShirt has nanoscale wires interwoven into the fabric to track vital signs and control temperature and can communicate results through a PDA to a base station where they can be monitored [196]. Garments like these can be used on soldiers and law enforcement to promote safe missions and provide immediate notice if care is needed.

Nanostructured Materials in Electronics

When one thinks of nanostructured materials in electronics, several topics come to mind: batteries, devices, and components. In reality these three topics could be broken down to numerous subtopics. We use some form of nanostructured materials everyday in our cars, homes, schools, hospitals, and our personal belongings. However, a closer look reveals that two basic formats exist for preparing these materials, top-down and bottom-up production. Historically, the top-down approach has been used to make smaller, faster, improved devices and components, starting from bulk materials and carving or etching away unwanted material to create nanosystems, devices, and materials [197]. Much electronic circuitry is prepared in this manner. The bottom-up approach is where a desired object is built up from smaller building blocks such as atoms or molecules. Both methods have promise to improve current electronics, but much of the research focused on improving bottom-up methods is more closely related to traditional nanomaterials synthesis.

Nanostructures are of great interest in electronics due to the shape, size, and observed composition properties, as compared to their larger scale counterparts. With these smaller dimensions, electron relaxation rates and interfacial crossing rates of electrons and holes in composite structure become more favorable [198]. Controlling these parameters during growth can be done by template synthesis (colloidal synthesis using various ratios of precursor and capping material) and surface impurities. Similarly, Wei and Zamborini have produced a synthetic vessel which allows real time monitoring of nanostructure growth through AFM [199].

Several materials have found use as lithium-ion batteries. It is thought that these materials can potentially be the answer for high power and energy demands of small mobile electronics and electric cars; however, cost, safety concerns, and various deficiencies must be overcome. Cathode materials from $LiNi_xMn_yCo_{1-x-y}O_2$, $LiMn_2O_4$, and $LiFePO_4$ are considered the most desirable while the most used anode material is still graphite. The lithium materials are prepared through many methods including solid-state, hydrothermal, and sol–gel processes. The small particles of the base cathode material are a good fit for high power applications due to the large surface area, internal structure, short diffusion length, and faster kinetics. Nanoparticles are also more accommodating to volume changes with less risk of cracking and minimal diffusion path lengths. However, the large surface area allows for side reactions on the electrode, coating the material and reducing the life-cycle of the cell, affecting the capacity retention [200, 201].

Several devices use nanomaterials for their electronic components. These include carbon nanotubes used in flat panel color displays, silicon nanowires for electrical conduits in electric and optical devices (photovoltaics and LEDs), and other materials for chemical and biological sensors [202, 203]. Caminade and Majoral have investigated the implementation of phosphorous-based dendrimer macromolecules as a method of coating and modifying surfaces, imparting nanostructures to various components [204].

Woehrle et al. has found a method to organize gold nanoparticles (1.5 nm) into linear chains through biomolecular nanolithography. These gold chains of DNA templates result in angstrom-level precision for controlled interparticle spacing which is a key requirement for electrical and optical applications [205].

Hannon et al. have continued the search for methods to selectively place and orient carbon nanotubes for use in circuitry by micro contact printing and conventional lithography of alkylphosphonic acids onto metal oxide surfaces. This method could improve large-scale integration of superior performance of the carbon nanotube field-effect transistor over the conventional silicon device [206].

Low resistive and good corrosion-resistant noble metal, transition metal, or silicon thin films have been used as contacts and conductors in electronics. Two deposition methods are typically used: vapor-phase or liquid-based growth. Díaz et al. used a different method for thin film deposition, namely, pyrolysis of metallophosphazene in air at 800°C [207]. The resulting thin film's morphology is dependent on the polymeric or oligomeric nature of the precursor, preparation method used, and the surface to be coated. This could create a cheaper more controlled method for clean array formation with versatile film application compared to currently restrictive gold films through chemical methods.

Commercialization

In 2005, the National Center for Manufacturing Sciences (NCMS) surveyed 600 manufacturing companies to scope the future of nanotechnology. These companies included automotive, semiconductor, aerospace, energy, utility, and information technology sectors. The response? Sixty percent expected to include nanotechnology in their products by 2009, just a few years later [208].

Last year the NNI created a multi-agency collaboration, led by NSF, to promote nanoelectronics. “Nanoelectronics for 2020 and Beyond” pushes NSF, DOD, NIST, DOE, and IC to promote five areas of research and development.

- Exploring new or alternative “state variables” for computing
- Merging nanophotonics with nanoelectronics
- Exploring carbon-based nanoelectronics
- Exploiting nanoscale processes and phenomena for quantum information science
- Creating a National Nanoelectronics Research and Manufacturing Infrastructure Network

The goal of this research is to reduce power consumption and heat production in small electronics through carbon-based solutions, gain significant growth over other nations’ electronic capabilities, create high quality jobs, and lastly, create a foundation to foster other nanoelectronic successes [209].

NanoMarkets predicted the nanoelectronics market to reach $82.5 billion this year. The leading push on this market is expected to come from the demand of high-performance, non-volatile memories. Nanosensors are expected to capture a significant portion of the market in medical, homeland security, and aerospace applications [210].

Nanostructured Materials in Medicine

One of the most exciting potential uses of nanostructured materials is in medicine. Research into the mechanisms underlying different medical pathologies, particularly cancers, has led to a better understanding of the causes, progression, outcomes, and potential treatments for these conditions. Unfortunately, a greater understanding of a medical condition does not necessarily lead to an immediate advance in the therapy for the condition. The human body is an extremely complicated and very intricate structure, with the result that a therapy must be carefully designed to provide the maximum benefit with a minimum of associated side effects or toxicity. Nanostructured materials, with their unique capacity to be directed to specific individual cells, will provide safer and more effective treatments for a large number of medical conditions.

Nanomaterials and Cancer Therapies

One area where nanostructured materials are being investigated is in the diagnosis and treatment of cancer. Cancer is one of the most feared medical diagnoses, even though cancerous conditions are relatively common and most cancers are well understood. According to the CDC, the most common cancers are prostate, breast, lung, colon, uterine, urinary bladder, non-Hodgkin’s lymphoma, melanoma, kidney, and ovarian [211]. Cancer incidence is strongly associated with age, and its incidence increases dramatically after age 40 [212]. In the 6-year period from 1999 to 2005, more than nine million people in the US were diagnosed with cancer, which works out to approximately 1.5 million people in the US who are diagnosed annually, or approximately 1 in 200 [213].

Why is this relatively common and well-understood type of disease so feared? One reason for the anxiety surrounding this disease is the way it is currently treated. Most cancers are treated with a combination of surgery (to remove as much of the tumor as possible), radiation (to destroy any cancerous cells which could not be removed by surgery), and chemotherapy (to destroy any invasive or metastatic cancerous cells). Most cancer patients find the course of radiation and chemotherapy to be extremely unpleasant, which is to be expected when the side effects include fatigue, skin problems, loss of appetite, hair loss, and lowered blood count (from radiation therapy) [214] and nausea, muscle and nerve problems, fatigue, anemia, and difficulty breathing (from chemotherapy) [215]. Worse yet, more than one round of treatment may be required to eliminate all the cancerous cells from the body. Cancers which have evaded detection until they have progressed to advanced stages may be incurable, and the patient may undergo treatment only to prolong life.

Nanostructured materials are expected to provide substantial improvements in the diagnosis and treatment of cancer. Most types of cancer are very treatable in the early stages, with good chances of cure or lasting remission. However, many cancer types are difficult to detect in these early, highly treatable stages. Soft tissue cancers, such as ovarian, esophageal, and pancreatic cancer, are especially difficult to detect. These cancers commonly cause vague, poorly defined symptoms such as fatigue and digestive disturbances. As a consequence, these cancers are often not diagnosed until they are very advanced. The median survival time after a diagnosis of pancreatic cancer is currently—3–6 months, with fewer than 4% of those diagnosed still alive 5 years post-diagnosis [216].

Fortunately, improved diagnostic tools for these cancers are in development. In one experiment, functionalized semiconductor quantum rods were used to optically image pancreatic cancer in mice [217]. Other nanomaterials, including gold nanorods [218] and CdSe quantum dots, [219] offer enhancement of existing imaging techniques by acting as contrast agents. Improved imaging capabilities can mean substantial improvements in patient quality of life and treatment outcomes via improved staging of cancers and improved ability to gauge tumor invasiveness. In situ imaging, which discloses the exact location and size of the tumor, is particularly useful. However, nanomaterial-based assays which demonstrate improved detection sensitivity are also desirable.

One such assay, in which a nanoparticle immunoassay measured the degree of protein aggregation of known cancer biomarkers, found that the level of protein aggregation or complexation was a positive indicator of certain cancer types [220]. Another research group designed a detection system capable of detecting as few as 25 target cancer biomarker protein complexes. This system is based on fluorescent-conjugated polymers on Ag/Au nanorods [221].

Many developmental nanomaterials have shown promise as potential treatments for cancer. Iron-based nanoparticles, including iron oxide nanoparticles, are being widely investigated for both their imaging properties and their potential use in anti-tumor therapies. These nanomaterials may be used for drug delivery [222, 223] or may be used for imaging and hyperthermia treatments [224, 225]. Other nanomaterials which show potential as therapies include carbon nanotubes [226–228], copper sulfide nanoparticles [229], and silica-based nanomaterials [230, 231].

Drug Targeting

In addition to providing improved detection and treatment of cancer, nanostructured materials may be designed to enhance existing treatments. The drug Cisplatin is widely used to treat solid malignancies, both alone and in combination with other chemotherapy medications. Unfortunately, the drug also has a wide array of dangerous side effects, including fatigue, nausea, vomiting, diarrhea, nerve damage, immune system suppression, kidney damage, hearing loss, and even the formation of secondary cancers resulting from treatment. The kidney damage resulting from administration of this drug can be severe, so much so that it is often dose-limiting [232].

In an effort to improve the drug by reducing its nephrotoxicity, a team from the Department of Medicine at Brigham and Women's Hospital in Cambridge, MA, engineered a nanostructured version of this drug. Their nanostructured version consists of a polymer which self-assembles in the presence of the appropriate amount of platinum to produce a nanoparticle. This nanoparticle releases Cisplatin in a pH-dependent manner and is rapidly targeted into the lysosomes of cancer cells, where it releases the drug and kills the cells. Compared to the free Cisplatin, the nanoparticle demonstrated a reduced distribution in the system as a whole, which led to reduced toxicity [233].

A number of medicines have now been prepared by coprecipitating drugs and polymers to produce polymeric nanoparticles, including Cyclosporine A [234], Tamoxifen [235], and Daunorubicin [236]. In general, this approach is reported to result in improved tumor targeting, greater efficacy, and reduced toxicity.

The improved tissue targeting capability of nanoparticulate preparations of drugs is not limited to anti-cancer agents. Other areas of the body, most notably the central nervous system (CNS), are difficult to reach for most medications. In one study, poly(DL-lactic-co-glycolic acid) (PLGA) nanoparticles with modified surfaces (chitosan or Polysorbate 80) were found to be capable of passing the blood–brain barrier [237]. Other therapeutic nanomaterial preparations have been designed to deliver to the eye [238], across the mucosa [239], through the dermis [240], and to bone [241].

In contrast, organs involved in detoxification and excretion, such as the liver and kidneys, are frequent recipients of medications whether or not the medication is supposed to reach them [242]. This leads to toxicity and side effects from the medication, which can limit the dosage or usage. However, medical nanomaterials have been shown to greatly improve the targeting of these drugs to the tissue in need of treatment and away from the liver and kidneys [233]. This improved targeting not only allows the drugs to be more effective therapeutically, but also lessens any toxicity caused by detoxification and excretion.

Treatments for Lifestyle Illnesses

Medical nanomaterials also show promise in the treatment of "lifestyle illnesses" such as obesity, heart disease, and type 2 diabetes. As the population in the industrialized world grows fatter, more and more people will suffer the effects of being overweight. As a result, the demand for improved medical therapies for these conditions is growing. Patients want less invasive therapies for diabetes (no needle sticks for insulin administration or blood glucose analysis), heart disease (no shunts or coronary by-passes), and obesity (no gastric surgery and the ability to eat what they want, when they want). While it is improbable that medicine will advance to the point where patients can "have their cake and eat it too," any advances in treatment will be welcomed.

Diabetes currently affects 7.8% of the total US population. Of those 60 and older, the prevalence rises to 23.5% [243]. The CDC predicts that by 2050 this incidence will double or triple [244]. Clearly, the need for less invasive treatments will continue to increase for the foreseeable future.

The current method of insulin delivery by injection has disadvantages including the potential for infection and the stress associated with self-injection. Nanoencapsulation of insulin in zirconium phosphate for oral delivery has been explored in vitro and is reported to result in the stable release of insulin from the nanocapsules [245]. Supercritical antisolvent (SAS) micronization of insulin has also been investigated for transdermal insulin delivery [246], and solid lipid nanoparticles have been investigated as carriers of insulin for pulmonary delivery [247]. If one or more of these insulin delivery methods proves to be as safe and effective as insulin injections, daily injections will soon be a thing of the past.

Fewer nanomaterial options are available for the treatment of obesity. Currently, one of the few efforts reported in the literature is the production of nanosized particles of orlistat, a gastrointestinal lipase inhibitor [248]. This allows the drug to be handled and administered more easily, plus the nanosized particles are better absorbed than their conventional counterparts. Another is the development of PEGylated all-trans retinoic acid nanoparticles, which have been shown in vitro to prevent adipocyte development and differentiation [249]. As the incidence of obesity increases, obesity research involving nanomaterials is likely to become more urgent.

The options for treatment of cardiovascular disease are more varied. Approaches currently under study include the use of nanomaterials to produce synthetic High-Density Lipoproteins (HDL) [250], injection of peptide nanofibers to enhance stem cell therapy [251], the use of nanoparticles to prevent atherosclerosis [252], and making nanoformulations of growth factors for therapeutic angiogenesis [253]. These approaches encompass both the prevention of cardiovascular disease via the prevention of atherosclerosis and treatment of diagnosed disease via cell therapy to re-grow the damaged tissues and improve organ function.

Other Medical Nanomaterial Uses

Nanomaterials also have potential usages in gene therapy. The use of nanomaterials has been shown to enhance the transfection of different target cells, including cancer cells [254] and skeletal muscle cells [255]. Transfecting regulatory genes into cancer cells encourages the cells to down-regulate their growth, up-regulate their differentiation, and can cause them to recognize their abnormal condition and trigger their self-destruction pathway. The ability to transfect skeletal muscle cells will, it is hoped, provide a cure for several inherited degenerative muscle diseases which are currently fatal. Nanomaterials also offer a great deal of hope for various opthalmic conditions which are currently difficult to treat or for which there is no treatment, such as retinal degenerative disease [256]. Because of the unique structure, composition, and function of the eye, many conditions are currently difficult or impossible to treat. However, current research in medical nanomaterials may eventually change this.

Finally, nanomaterials may provide an improved way to prevent infection. Some potential uses for nanomaterials in this area include oral biofilm prevention [257], control of airborne indoor bacteria [258], and the production of titanium alloys which discourage infection [259].

Commercial Uses

The impact of nanotechnology on healthcare is likely to be the largest among nanotech applications. North America has the largest market, worth $4.75 billion in 2009 with Europe following at $3.65 billion. The drug delivery market alone is expected to grow to $16 billion by 2014. Other markets such as biocompatible implants and coatings and diagnostics are expected to grow 42 and 21.8%, respectively, from 2009 to 2014 [260].

Elan Drug Technologies, headquartered in Dublin, Ireland, has found a solution to a problem that could lower the dosage of the medicines you take. Their NanoCrystal® technology reduces the size of drug particles to less than 2,000 nm, exposing a larger surface area. Many drugs exhibit poor solubility and the NanoCrystal technology is a commercially available answer. This technology has produced five products and brought in $1.8+ billion annually in market sales [261].

NanoBio Corporation, Ann Arbor, MI, has technologies in development for topical anti-infective treatments and mucosal vaccines. This technology, NanoStat™, uses emulsions manufactured at the range of 150–400 nm. NanoStat treatments are toxic to microbes but non-irritating to skin and mucous membranes. This technology has shown effectiveness against certain bacteria, viruses, fungi, and spores. Since the mechanism involves a physical and not a chemical process, the potential for the development of drug resistance is substantially lowered. Vaccines for influenza and Hepatitis B are in current development [262].

There are numerous companies working on the delivery of drugs, typically in cancer applications. For example, the nanoscale development of gold-based drug compounds can be used to target tumors, limiting the impact on surrounding organs and tissues. Current issues with drug delivery are in the efficiency of loading and controlling the drug [263]. Future successes in nanotech drug delivery will depend greatly on the ability to scale up the technology, meaning a close scientist–engineer relationship to create an efficient, effective product.

In conclusion, the medical uses for nanomaterials are numerous and varied. Nanomaterials may eventually be used to provide routine treatments for conditions which are currently untreatable, such as muscular dystrophy and retinal degeneration, in addition to providing improved therapies for conditions such as cancer, cardiovascular disease, and diabetes.

Nanostructured Materials in Coatings

Another potential usage of nanomaterials is in coatings technology. There are innumerable different types of coatings, all designed for improving the characteristics of the coated material. A short list of coating types includes fabric coatings for improved stain resistance [264], antimicrobial coatings for infection prevention [265] and product preservation [266], conductive coatings for solar cells and other electronic applications [267], scratch-resistant coatings for optics [268], and antifouling coatings for marine applications [269]. Nanomaterials offer the opportunity for improvements in nearly all coating fields and applications.

Edible Nanomaterial Coatings

A number of different industries are currently using or are expected to use nanomaterial coatings. The food processing industry has been researching nanocellulose-based edible coatings for improved shelf life of fresh fruits, vegetables, and meat products [270, 271]. The medical industry uses nanosilver coatings on catheter surfaces and textiles [272], and marine equipment manufacturers use antifouling nanomaterial coatings [273]. Multiple applications for nanomaterial coatings exist in the electronics industry [274], most notably for applications such as solar cells [275].

Food safety is a significant public health concern. The CDC estimates that 48 million Americans are sickened annually by foodborne illnesses [276]. Because many food products are now mass produced, contaminated foodstuffs may easily produce nationwide outbreaks of foodborne illness. Additionally, fresh foodstuffs, such as produce, are highly perishable and have short shelf lives even under ideal conditions. Spoiled food represents a loss to everyone involved in food production and consumption: the consumer, who sees higher prices and lower availability thanks to spoilage; the grocer, who must pay to dispose of the spoiled food; and the producers and distributors, who don’t get paid for producing and transporting the food.

One high-tech solution which has been proposed to lower the incidence of contamination and prevent spoilage is the use of edible coatings. These coatings, which would predominantly be used on fresh or raw foods, could be designed with antioxidant capabilities (to prevent spoilage) as well as antimicrobial capabilities (to prevent contamination). Most commonly, the coatings consist of macro materials such as chitosan [277] or alginate [278]. However, a small number of researchers have reported that nanomaterials such as cellulose nanofibers [279] may be used to enhance the properties of these coatings. While only a few instances of edible nanomaterial coatings are currently known, the use of such coatings can be expected to increase as their safety is established and their organoleptic properties (their taste and smell) are improved.

Medical Nanomaterial Coatings

As was mentioned in depth in the Sect. 8, nanomaterials have a large number of uses in the medical field. However, here we will focus solely on nanomaterial coatings with medical uses.

One use for medical nanomaterial coatings is for infection prevention. The use of catheters and ports leaves patients more vulnerable to infection, as these indwelling medical devices lack the anti-infective properties of natural body orifices and are readily colonized by infectious organisms [280]. The most common antimicrobial nanomaterial coating currently in use is nanosilver, which is commonly encountered on medical instruments [281], in catheters [282], and in wound dressings [272]. The antimicrobial activity of nanosilver, when incorporated into these materials, offers the patient some protection against nosocomial infection and microbial colonization. These products are widely commercially available from a variety of suppliers and are widely used in hospitals and clinics.

Other medical uses of nanomaterial coatings include drug delivery [283, 284] and tissue scaffolding to enhance medical implant integration, particularly into bone [285, 286]. Naturally, prosthetic implants such as dental and joint implants are more successful if the patient’s healing process will integrate the implant into the existing bone, thereby securing it and allowing for more natural functioning. Hydroxyapatite nanoparticles (which can be incorporated by the body into bone structures) are especially useful when combined with the silica-polymerizing enzyme silicatein α (which encourages the process). Hydroxyapaptite nanoparticles are commercially available from suppliers such as Sigma Aldrich for use as a bone injectable substrate and for bone grafting.

Additionally, researchers are working on a nanomaterial coating that can safely eradicate *Staphylococcus aureus* (MRSA). A coating of this nature would be applied to a variety of surfaces including medical instruments and surgical masks. The research uses a naturally occurring enzyme which kills the target bacteria, attached to a carbon nanotube for increased mobility [287].

Industrial Nanomaterial Coatings

In addition to applications in food processing and medicine, nanomaterial coatings also have a great many industrial uses. Antifouling coatings are highly useful for the shipping industry, as smooth hulls and propellers create much less drag, which makes fuel usage more efficient [288]. Likewise, bioreactors and oil pipelines are expected to benefit from this technology [289], as are heat exchangers involved in food [290] or energy production [291]. The use of anti-adhesive or antifouling coatings on heat exchangers in power plants is expected to dramatically increase efficiency, as a 1 mm thick biofilm accumulation on the wall of a low carbon steel heat exchanger results in a resistance to heat transfer equivalent to an 80 mm increase in tube wall thickness [292]. There are environmental and safety concerns with some of this technology, as biocide-releasing coatings can adversely affect marine organisms or can lead to the presence of biocides in foodstuffs.

Other types of nanomaterial coatings, including self-cleaning coatings [293], are also expected to be extremely useful. Self-cleaning coatings on items such as solar cells could decrease maintenance costs and increase energy output. Currently, solar cell output can be severely affected by the presence of dust, with even tiny amounts such as 1/7th of an ounce per square yard decreasing power conversion by 40% [294]. Since the best locations for solar cells are in areas which receive large amounts of sunshine (and conversely, very little precipitation), dust is a serious problem. Self-cleaning coatings would go a long way towards remedying the effects of dust. Additionally, many optical instruments have parts which are difficult or time-consuming to clean, so self-cleaning coatings would decrease maintenance needs in these cases as well. Other less obvious uses for self-cleaning coatings include exterior coatings for building materials such as wood and stucco, which soil easily and can be colonized by algae and other microorganisms [295].

Scratch-resistant nanomaterial coatings have some obvious uses, such as in abrasion-resistant coatings for automobiles, industrial parts, furniture, and flooring [296]. Less obvious uses include abrasion-resistant coatings for use on teeth and bones [297]. Alumina nanoparticles are most commonly used for these coatings, although silica nanoparticles may also be used [298].

Finally, conductive nanomaterial coatings offer improved performance characteristics over currently used conductive coatings like indium tin oxide. Indium tin oxide coatings, while widely used, are brittle, expensive, and chemically unstable. Nanomaterials such as single-walled carbon nanotubes [299], conductive polyamide fibers [300], and single-walled carbon nanotubes with silver nanowire films [301] all offer advantages such as flexibility and improved chemical stability. Additionally, some of these coatings can be applied to inexpensive substrates to produce inexpensive energy storage devices.

Commercial Uses

Nanomaterial coatings are already a large portion of the commercial nanotechnology industry, as a whole. In 2008, commercial nanomaterial coatings generated revenues exceeding 600 million USD. Targets of these technologies include construction, healthcare, transportation, and defense markets. The commercial nanomaterial coatings market is expected to reach over five billion USD in revenues in 2013, according to a Research and Markets report [302].

Companies like Xurex, Inc., a manufacturer of nanomaterial coatings located in Albuquerque, NM, could make significant contributions to conserving building materials and lowering costs across many fields. Xurex offers coatings to combat the corrosion and degradation of pipes, roads, bridges, ships, etc., which could potentially save millions of dollars [303]. Other applications in technology include self-cleaning surfaces and more durable materials, such as concrete. These energy efficient improvements could reduce the 41% of all energy use that is lost to commercial and residential buildings [304].

There is a real commercial need for alternative coatings in relation to the aerospace and defense industries. Coatings for thermal barrier, anti-ice, and surface protection are a necessity. Funding from governments is pushing these applications as nanocoatings can typically offer additional properties, including heat resistance, as well as offering an environmentally friendly alternative to standard coatings [305].

Environment, Health, Safety Issues Related to Nanostructured Materials

Nanostructured materials have come into play in the consumer market after a relatively short development period. As they are part of such a new science, relatively little is known about them except for those that have been studied as desirable characteristics for developing products with distinct performance advantages. The greatest concern surrounding the use of nanotechnology is the possibility of toxicity or other damaging interactions with living organisms. Naturally, while toxic interactions with pathogens are desirable for some classes of nanomaterials, toxic interactions with organisms in general are not. The situation is further complicated by the fact that chemically identical nanomaterials with different sizes and shapes can and will show different toxicological profiles [306], different tissues and organisms used to assess toxicity will exhibit greater or lesser sensitivities, and approved, standardized methods for evaluating nanomaterial toxicity are scarce. Worse, the use of in vitro assays as a means of assessing toxicity may lead to erroneous conclusions regarding the toxicity of the nanomaterial in whole organisms.

For example, it is known that some mammalian organs and organ systems are more sensitive to nanomaterial toxicity than are others. The lungs, in particular, are known to be highly sensitive to nanomaterials [307]. However, toxicity assays which are conducted on lung tissues in vitro and in vivo can show very different outcomes. A study conducted on the toxicity of both nanoscale and fine zinc oxide particles in rat lungs found several confounding factors, including the failure of commercial samples to match the manufacturer's specifications and the development of “metal fume fever” in animals exposed to levels of zinc oxide which caused no demonstrable cytotoxicity in vitro [308]. Indeed, at exposure levels which were pathological to the animal, the in vitro assays failed to show evidence of any inflammatory responses including increased chemokine production. And what are we to make of another study, which found that the toxicity of zinc oxide nanoparticles could be lessened by doping them with iron [309], although iron oxide nanoparticles have themselves been found to be toxic? [310].

Some nanoparticles have demonstrated similar toxicity towards plants and bacteria. Phytotoxicity is usually measured by assessing seed germination and root elongation in seeds and seedlings exposed to the test material. One study which measured the effects of multiwalled carbon nanotubes and nanoparticulate forms of metallic copper, metallic silver, zinc oxide, or silicon on zucchini plants found that while exposure to these materials did not reduce germination, root elongation and plant biomass production was negatively affected [311]. Biomass in particular was greatly affected, with exposed plants showing biomass reductions of up to 75% of controls. As has been found in other studies, the negative effects were significantly more pronounced after exposure to nanoparticulate forms as compared to bulk forms. Nanomaterials have also been shown to be toxic to cyanobacteria and microalgae [312] and to microbial soil communities collected from California grassland [313]. In the case of the soil bacteria, nanomaterial exposure reduced the overall biomass of the community and altered the composition and diversity of the community. The effect of these changes on the chemical processes which occur in the soil, and on other denizens of the grassland, is not known.

How are analysts to proceed when confronted by these types of toxicological results? Conducting subsequent experiments to better understand the underlying mechanisms is one obvious answer, but various interested parties must still determine how to reasonably proceed safely in the meantime. The approach to nanomaterial safety currently taken by the US government is to make a coordinated effort to (1) monitor both expected and unexpected consequences of nanotechnology and (2) identify and prioritize nanomaterial health and safety research needs [314]. These actions were taken as part of the NNI, which was established in 2000. The findings from research undertaken by this group will be used to guide public policies at regulatory agencies such as the EPA and FDA. Currently, the EPA can’t ask a chemical producer to provide data regarding the risk assessment of a chemical, unless the EPA already has data showing that the chemical presents an “unreasonable risk” to human health or the environment [315]. Regulatory policies are greatly needed, as nanomaterial production and usage becomes more widespread. As a group of Swiss analysts found, many companies in industrial sectors other than those specifically identified as “nanotechnology” are already using nanomaterials [316].

Another approach, taken by members of the insurance industry, was to perform a relative risk assessment based on the nanomaterial manufacturing process. This risk assessment took into account the normal operations risk, incident risk, and latent contamination risk [317]. At the conclusion of the assessment, the risks of fabricating five different nanomaterial types (single-walled carbon nanotubes, buckyballs, zinc selenide quantum dots, alumoxane nanoparticles, and titanium dioxide nanoparticles) were determined to be comparatively low in relation to other manufacturing processes such as battery production.

There also is a movement towards understanding what safety steps other industrial nanomaterial producers are taking. Surveys of health and safety measures taken by industrial producers have found that most producers have given the issue some thought, with many specifying the use of PPE and engineering exposure controls [318]. Undertaking exposure and hazard assessment procedures are common in industry [319].

Waste management is another area impacted by nanotechnology. The introduction of nanomaterial containing products has instigated use of various assessment tools (i.e., life-cycle assessment, LCA) to establish the impact of products on global warming/climate change, stratospheric ozone depletion, human toxicity, ecotoxicity, and others. Nanocomponents are assessed for aspects that include material selection, manufacturing, application, and disposal/recycle. Physical properties must be well documented and investigated because they influence the toxicological impact through transport and interaction of the materials when released to the environment. Many products are currently using nanocomponents containing titanium, carbon, silver, iron oxide, zinc, gold, and/or silicon/silica. These include appliances, automotives, electronics and computers, food and beverage, children’s goods, health, personal care, sporting goods, and home and garden products. Most products are assessed for toxicity by producers on a cradle-to-gate basis, vs. cradle-to-grave, leaving the onus for safe disposal on the consumer and the public waste facilities [320].

Walser et al. performed a LCA for nanosilver t-shirts and compared the data to both triclosan biocidal t-shirts and non-biocidal t-shirts. The cradle-to-gate study found the flame spray pyrolysis method of producing nanoparticles had a climate footprint of 2.7 kg of CO_2-equivalents; the plasma polymerization with silver co-sputtering had 7.67–166 kg of CO_2-equivalents (depending on maturity). In contrast, regular and triclosan t-shirts had 2.55 kg of CO_2-equivalents. Based on this study, it is clear that the consumers’ wash cycle habits and eventual disposal of garments have minor effects on climate footprint. On the other hand, mining operations for raw materials have a greater effect due to toxic silver emissions [321].

One of the beneficial uses of TiO_2 is as a sunscreen due to its effective UV reflection and adsorption capacities. The nano-TiO_2 has higher transparency and efficiency making it even more desirable. Sunscreens undergo extensive testing in direct exposure to the skin scenarios but not further in its life-cycle. Botta et al. have investigated environmental release of these nano-TiO_2 containing sunscreens to water (fresh and salt) after artificial aging of the solutions. Their results found that in salt water, a major part of the nano-TiO_2 containing residue will aggregate and sediment, further affecting bottom dwelling sea life. In contrast, the fresh water released portion will remain stable and accessible to the entire water column [322].

Another application in use is nanoparticle CeO_2 as a diesel fuel additive. This additive has been designed to increase mileage by performing as a combustion efficiency aid. However, some CeO_2 <100 nm has been found in the expelled particulate matter. This is then released into the environment in relatively low concentrations. Several studies have been performed to determine toxicity of these particles, but no testing has been done on the diesel additive exhaust system emitted particles, or on low concentrations of ceria nanoparticles [323].

Currently, limited studies are available for nanomaterials that enter waste facilities. A newly emerging term, "nanowaste," is defined as any waste stream(s) that contain nanomaterials or synthetic by-products of nanoscale dimensions, generated either during production, storage, and distribution, or waste stream(s) resulting from the end of lifespan of formerly nanotechnology-enabled materials and products, or items contaminated by nanomaterials such as pipes, PPE, etc. Additionally, BSI British Standards Guide PD 6699-2 defines nanowastes in four forms: pure nanomaterials at point of production, materials and surfaces that have been contaminated with nanomaterials (containers, PPE, etc.), liquid suspensions containing nanomaterials, and solid matrices containing nanomaterials. Concerns that have been raised dealing with these nanowastes include increased uncontrolled releases of nanomaterials to the environment, the technological and legislative challenges nanowastes pose to waste management systems, establishing a nanowaste classification system, and ways to enhance current and ongoing management of nanowastes. The very nanostructures that are prized for their applications are now a concern due to their potential ability to aid faster bonding with pollutants, facilitating faster transfer and greater transport of hazardous materials through air, soil, and water. Traditional wastewater treatment by-products (biosolids from settling tanks) can become too concentrated with pollutants such as Ag to be reused as compost or fertilizer, flocculation methods can be thwarted by surface coatings or other functionalities present on the nanomaterial, or nanomaterials can just simply pass though without any interaction at all [324].

Some concerns with nanoparticulate materials are their capability for dispersion into air and water, then further into the environment. Drugs, personal care products, quantum dots, and environmental remediation projects all have high potential for uncontrolled environmental release and their incidence is expected to dramatically increase. Nanostructure characteristics that are a determining factor include particle size, surface area, chemical nature, and charge, while reactive oxygen species at the surface may have an effect. As human data is rather limited, animal studies have been performed. Major hazards include inhalation of particles <100 nm as they deposit in the pulmonary region causing inflammation of the lungs. Nanoparticulate to ultra fine particles (including PTFE, carbon [general, fullerenes, and nanotubes], and TiO_2) have been found to impair lung function, increase respiratory syncytial virus infection, impair phagocytosis, encourage tissue thickening or fibrosis, increase coughing and sputum production, give rise to higher risk of chronic bronchitis, encourage tumors, and cause lung cancer. Nanoparticles may penetrate lung tissue to enter the blood stream (and on to the heart, liver, and brain). Studies have found that SiO_2, TiO_2, ZnO, and Al_2O_3 can penetrate skin layers. Ingested TiO_2 particles (200 nm) are immunologically active and can further penetrate the intestinal barrier (onto the lymphatic and blood systems) [325].

NIOSH has partnered with several organizations throughout the world in government, academia, and business, creating the Nanotechnology Research Center (NTRC) to assess safe nanotechnology in the workplace. They have been investigating critical topic areas; namely, toxicity and internal dose, measurement methods, exposure assessment, epidemiology and surveillance, risk assessment, engineering controls and PPE, fire and explosion safety, recommendations and guidance, communication and information, and applications. NIOSH NTRC is collaborating with NNI to pursue these mutual goals [326, 327]. The investigations have been instrumental in providing guidance for those who produce and work with nanomaterials and nanomaterial-enhanced products. As work is continuing in each of these critical areas, information has been found to make the nanomaterial workplace a safer one. General guidelines have been determined for working with nanomaterials. These include minimizing worker exposure through engineering controls to reduce aerosol exposure, even up to source enclosure and isolation. Implementation of a risk management program should include: (1) evaluation of hazard posed by nanomaterials based on its available physical and chemical property data, toxicology, or health effects; (2) worker task exposure assessment; (3) good work practice education and training; (4) establishing criteria and procedures for installing and evaluating engineering controls at exposure locations; (5) procedures for determination of need and selection of PPE; and (6) systematic evaluation of exposures to ensure control measures are working properly and workers are being provided appropriate PPE. Filtration studies to date have indicated that high-efficiency particulate air (HEPA) filter applications (exhaust ventilation and vacuum pickup) should effectively remove nanomaterials from the air. When engineering controls are not enough, preliminary data suggests that a NIOSH-certified respirator will be useful; however, no guidelines have been set for airborne engineered nanoparticles [328]. Several tools have been developed for guidance in the form of documents for medical screening and hazard surveillance (DNNH NIOSH Publication No. 2009-116), managing health and safety concerns (DNNH NIOSH Publication No. 2008-112), and a searchable on-line library for nanoparticles, etc. [329].

From these results, it is obvious that more research into the environmental effects and fate of nanomaterials, and their known and potential effects on human health and the environment, is needed. Additionally, it is clear that an improved regulatory framework is also needed, in order to facilitate the collection, interpretation, and dissemination of the experimental findings.

Nanostructured Materials in Other Consumer Products

Nanotechnology has been finding its way into more and more consumer products over the past decade. As the realization of potential applications for nanotechnology grows, the knowledge will continue to spread. The bottom line for nanotechnology in consumer products will be “Is there a demand?” The consumer has the final say. Skeptics and scientists alike will be susceptible to the consumer’s last word, or rather, last dollar. The final say in 2009 was a resounding 1.545 billion USD spent on golf clubs, electronics, skin cream, and a plethora of other products that utilize nanotechnology. The skeptics lose out to the optimistic forecasts of Research & Markets report “The World Market for Nanotechnology and Nanomaterials in Consumer Products, 2010–2015,” which expects that number to triple by 2015. In 2015, 5.335 billion USD will be spent on innovative products that stand out from the rest because of something 100,000 times smaller than a human hair [330].

Tracking Commercial Nanomaterial Products

In April 2005, the Project on Emerging Nanotechnologies (PEN) was initiated. PEN contains the largest online inventory of commercial products with nanotechnology along with other nanotech learning resources. New products are added as often as they are found, which in 2008, was at a rate of 3–4 per week. PEN is a joint effort between the PEW Charitable Trusts and Woodrow Wilson International Center for Scholars. As of August 2009, the list of commercially available products contained over 1,000 products. The list realized a 379% growth since March 2006 [331].

Many groups worldwide, like PEN, are trying to identify the number of products available that use nanotechnology. Challenges exist which prevent a complete list from being formed. Currently, there are no requirements for nanotechnology products that would identify or separate them from their non-nano competition. Some products which claim to contain nanomaterials do not, and some products that do contain nanomaterials do not outwardly claim this. Nanotech products do not have special regulations, as of yet. Debates are raging on this topic and governments are being pressured to look further into the proposed risks. Until a decision is reached and the technology is tracked, nano products may be sold under the label of the manufacturers’ choosing.

Commercial Uses

Amid all of the debates and controversies, innovative and earth friendly products are being developed by companies large and small. What most people do not know is that some of the most widely known products, like the iPhone by Apple®, are products of nanotechnology. The iPhone features a memory chip which is a result of nanotechnology [332]. Additionally, many everyday products are being improved through the use of nano-sized materials. According to “The Nanotechnology Consumer Products Inventory,” the most common material mentioned in the product description was silver (259 products) [333]. Carbon was the second most referenced (82 products including fullerenes and nanotubes), followed by titanium dioxide (31), zinc oxide (24), silica (15), and cerium oxide (1). Among potential environmental applications of nano materials, remediation of contaminated ground water with nanoscale iron is one of the most important examples [10]. Nanoparticles of titanium dioxide and zinc oxide are included in personal care products such as beauty products, sun screens, toothpaste [334], and textiles [335]. Silver nanoparticles are increasingly used as antimicrobial components in detergents, food packaging, and textiles such as socks and underwear [333]. Other potential uses for silver nanomaterials include portable filters for water purification which could be used in remote or underdeveloped areas lacking water treatment infrastructure [336].

In addition to the use of nanotechnology to improve these classic consumer products, entirely new uses which take advantage of the novel physics of nanomaterials are being developed. Nano-emulsions, a variation on the “oil and water” mixture used for salad dressings and cosmetics, also hold great promise for applications as different as future food technology, pharmaceutical development, disinfection, and pesticide application [337].

Large companies, like Samsung and IBM, are in the midst of nanotechnology use and development and have products on the market. Samsung makes nanopowders and has prototype field emission displays (FED) [338]. IBM has a science department devoted to nanoscale research in carbon nanotubes and nanometer-scale local oxidation [339].

In addition to the large companies making the technology known through their products and research, smaller companies are also working to get their name and products on the market. The odor elimination products marketed as OdorKlenz® by NanoScale Corporation use destructive adsorption to neutralize odors. This technology stems from NanoScale’s flagship product, FAST-ACT®. The OdorKlenz product line includes a laundry additive, a surface treatment, a skunk treatment for pets, and a vomit absorbent. OdorKlenz also features an air filtration cartridge available to the disaster restoration market for odors after fires, floods, or other disasters. The OdorKlenz branded products were introduced to the market in 2007 [73].

Regulation

Decisions made in the near future may alter the currently optimistic outlook of nanotechnology. Governments have difficult decisions to make on the risk of nanotechnology in products from skin creams to foods to socks. One author, a member of the European Commission, stated “Further research is needed on the toxicological and ecotoxicological properties of nanoparticles, their uptake in the body, accumulation in the tissues and organs, transport characteristics, exposure and dose–response data, and their distribution and persistence in the environment” [340]. This uncertainty about regulatory issues has a discouraging effect on the commercial development of nanomaterials, as manufacturers are reluctant to develop products without a suitable regulatory framework and guidance on safety requirements. However, at the present time the promise of nanotechnology appears to outweigh any potential risks, and thus R&D efforts continue. There is no doubt that, one way or another; nanotechnology will become an integral part of our daily lives.

Concluding Remarks and Future Outlook

Previous chapter sections have reviewed a large number of different nanotechnology applications in various areas, including medicine, textiles, catalysis, electronics, and others. Nanomaterials are already in commercial use in many of these areas, with new technologies in development, as well. Consequently, nanomaterials are now widely encountered in daily life, although many consumers remain unaware of their presence in the products they use.

Despite the enormous promise offered by nanomaterials, there are still multiple barriers to their development and use. One substantial consideration for those seeking to commercialize a nanotechnology or include it in their commercial product is the regulatory framework which exists for nanotechnology. Consumers demand safe products, so product manufacturers and distributors put a great deal of time and effort (and money) into ensuring that their products comply with the appropriate regulations—and this is the primary hurdle to nanomaterial usage. Presently, many regulatory agencies such as the US EPA have few guidelines on how to regulate nanomaterials [341]. Without a suitable regulatory framework, manufacturers and distributors have difficulty assuring themselves and their buyers of the safety of their product.

How did this situation come about? The same unique qualities that enable nanomaterials to have their capabilities also make them difficult to regulate. Bulk materials, such as carbon, have set physical properties and capabilities. However, carbon nanomaterials have an entirely different set of physical properties and capabilities. Worse (from a regulatory standpoint), these physical properties and capabilities change with the size, shape, and structure of the carbon nanomaterial. Whereas bulk carbon does not conduct electricity very well, carbon nanotubes do [342]. Bulk carbon has few detrimental effects on aquatic life, but carbon nanotubes are known to be toxic [343], and a slightly different form, fullerenes, may accentuate the toxicity of other chemicals [344]. The quantum mechanics which regulate nanomaterials’ interactions with other substances also make their toxicological behavior difficult to predict.

A number of organizations, including federal bureaus such as the EPA and testing standards developers such as ASTM, have begun developing methods to assess nanomaterial properties. ASTM currently has seven standards related to nanotechnology: One defining the standard terminology [345], two defining methods for determining some of the physical characteristics of nanomaterials [346, 347], and four standards relating to health and safety [348–351]. Once sufficient research has been performed, trends in the toxicology of different nanomaterial types should become apparent, and a regulatory framework with standard assay methods can be established.

Previous chapter sections have given multiple estimates on the future size of different nanomaterial markets. If nanomaterials are not found to have serious adverse health or environmental effects, then the optimistic predictions of their worth will likely not fall short of the mark. However, if nanomaterials are found to have serious environmental effects or detrimental effects on human health, their commercial future is considerably less rosy.

In conclusion, much more research is needed into the properties and capabilities of nanomaterials. Nanomaterials offer great improvements in many different areas and may have very positive impacts on human health and the environment, but they may also have negative impacts which are not yet understood. As research and development efforts progress, both the promise of and the problems with nanomaterials will become apparent. Only after both of these aspects are recognized can the full potential of nanomaterial technologies be realized.

Questions

1. Why do nanomaterials behave differently from their macro counterparts?
   Answer: Because of the small size of nanomaterials, a large fraction of atoms are located close to the surface of the material instead of in the bulk, which increases their reactivity. Additionally, at such a small scale, quantum effects become increasingly important in governing the behavior of the material.
2. What are the advantages and disadvantages of a government guided research initiative like that of Germany as opposed to more independent research other countries use?
   Answer: The main advantage of government guided research is that a more focused approach will be implemented. With more researchers looking into the same set of problems, an outcome could be more likely. The disadvantage would be the limitations that are set on exploring new discoveries.
3. Should companies manufacturing nanotechnology containing products be mandated to add this information to their product labels?
   Answer: Yes, companies should be mandated to add the phrase “nanotechnology containing products” to their product labels should safety concerns arise later. All products should be labeled to include this information.—OR—No, the manufacturer should maintain the right to market/label as they see necessary unless risks become known. Restricting labels may inhibit the advancements of the nanotech industry.
4. What are some benefits of nanomaterials for use in cancer treatments?
   Answer: The benefits of using nanomaterials in cancer treatments include more accurate dosing, reduced side effects, reduced toxicity, accurate and non-invasive diagnostics.
5. Describe the significance of Richard Feynman’s speech and his approach to nanotechnology.
   Answer: Richard Feynman’s speech in 1959 opened discussion of the bottom-up approach as opposed to the top-down approach to nanotechnology and revolutionized the idea of nanotechnology.
6. What benefits do nanostructured materials bring to environmental remediation?
   Answer: Increased surface area means increased reactivity and adsorption capacity, therefore nanomaterials are more efficient than their non-nano counterparts in environmental remediation. For example, Fe(0) nanoparticles do not have mobility issues as larger Fe(0) particles.
7. Why is there so much regulatory uncertainty surrounding nanomaterials?
   Answer: One of the reasons there is so much regulatory uncertainty surrounding nanomaterials is because the properties of nanomaterials are dramatically different from those of bulk materials. Furthermore, because of quantum effects, the physical configuration (shape, structure) of a given nanomaterial can cause it to behave differently than another chemically identical nanomaterial with a slightly different physical structure. For example, carbon nanotubes will behave differently from carbon nanospheres, even though they may be chemically identical, even to the extent of containing the same number of carbon atoms.
8. Describe the two mechanisms of action of nanocrystalline high surface area TiO_2 on the chemical warfare agent simulant DMMP.
   Answer: Nanocrystalline high surface area TiO_2 can adsorb DMMP via hydrogen bonding to exposed hydroxyl groups (molecular adsorption) and via its titanium sites (reactive adsorption).
9. Why is the protection offered by textiles containing nanomaterials potentially preferable to that offered by textiles containing activated carbon spheres?
   Answer: Activated carbon spheres entrap chemical pollutants or toxins, but do not destroy them as nanomaterials do. Additionally, activated carbon spheres can off-gas the absorbed chemical during temperature changes and preferentially absorb humidity in the air rather than the toxin or pollutant.
10. You are a graduate student supported on an EPA fellowship research grant whose focus is removing Cr(VI)

contamination from groundwater in a limestone (calcium carbonate) aquifer. One of your research committee members has suggested Fe-based nanomaterials. Is this approach likely to work?

Answer: No, using Fe-based nanomaterials for removing Cr(VI) contamination from groundwater in a limestone (calcium carbonate) aquifer is not likely to work. Research has shown that, while Fe-based nanomaterials can remove Cr(VI) from contaminated waters, removal capacity drops to 55–77% in the presence of magnesium, calcium, or carbonate ions. As limestone is composed of calcium carbonate, the groundwater would almost certainly contain calcium and bicarbonate ions.

11. Which country has the largest number of nanotechnology companies?

    Answer: The United States of America has the largest number of nanotechnology companies.
12. How will self-cleaning and conductive coatings promote greener energy production?

    Answer: Both self-cleaning and conductive coatings are expected to improve the performance of solar cells promoting greener energy production. Self-cleaning surfaces will prevent dust build-up, while conductive coatings can provide both more efficient energy capture and potentially cheaper production of the cells and energy storage devices.

Acknowledgment The authors would like to thank Ms. Krista Hill-Combs for her assistance in preparing this manuscript.

References

1. Nouailhat A (2008) An introduction to nanoscience and nanotechnology. Wiley, New York
2. Iseli A, Kwen H, Rajagopalan S (2009) Nanomaterials for environmental remediation. In: Klabunde KJ, Richards RM (eds) Nanoscale materials in chemistry. Wiley, New York
3. Hornyak GL, Dutta J, Tibbals HF, Rao AK (2008) Introduction to nanoscience. CRC Press, Boca Raton
4. Wilson M, Kannangara K, Smith G, Simmons M, Raguse B (2002) Nanotechnology: basic science and emerging technologies. Chapman & Hall/CRC, Boca Raton
5. Shong CW, Haur SC, Wee ATS (2010) Science at the nanoscale, an introductory textbook. Pan Stanford, Singapore
6. National Nanotechnology Initiative (2011) FAQs: nanotechnology. Online at http://www.nano.gov/. Accessed 4 Feb 2011
7. Feynman R (1960) There's plenty of room at the bottom. Caltech Eng Sci 23:22–36
8. Maynard A, Michelson E (2006) First nanotechnology consumer products inventory Available to public. Woodrow Wilson International Center for Scholars. Online at http://www.nanotechproject.org/inventories/consumer/. Accessed 25 May 2011
9. Zhang W (2003) Nanoscale iron particles for environmental remediation: an overview. J Nanopart Res 2:323–332
10. Tratnyek PG, Johnson RI (2006) Nanotechnologies for environmental cleanup. Nano Today 1:44–48
11. Serpone N, Dondi D, Albini A (2007) Inorganic and organic UV filters: their role and efficacy in sunscreens and suncare products. Inorg Chim Acta 360:794–802
12. Yuranova T, Laub D, Kiwi J (2007) Synthesis, activity, and characterization of textile showing self-cleaning activity under daylight irradiation. Catal Today 122:109–117
13. RNCOS (2010) Nanotechnology market forecast to 2013. Online at http://www.emailwire.com/release/37447-RNCOS-Releases-a-New-Report-Nanotechnology-Market-Forecast-to-2013.html. Accessed 25 May 2011
14. Report to the President and Congress of the Third Assessment of the National Nanotechnology Initiative (2010) President's Council of Advisors on Science and Technology, Executive Office of the President, Washington, DC
15. Cientifica Ltd. (2009) Nanotechnology takes a deep breath... and prepares to save the world! Global Nanotechnology Funding in 2009. Online at http://www.cientifica.com/WhitePapers/054_Nanotechnology%20Takes%20a%20Deep%20Breath.pdf. Accessed 14 September 2012
16. Organisation for Economic Co-operation and Development (2009) Inventory of National Science, Technology and Innovation Policies for Nanotechnology 2008. Working Party on Nanotechnology, 17 July 2009
17. The National Nanotechnology Initiative (2009) Research and development leading to a revolution in technology and industry, supplement to the president's FY 2010 Budget. Online at http://www.whitehouse.gov/sites/default/files/microsites/ostp/nni_2010_budget_supplement.pdf. Accessed 14 Sep 2012
18. National Institute for Nanotechnology (2011) online at http://nint-innt.nrc-cnrc.gc.ca/, 14 May 2010. Accessed 5 May 2011
19. CNPq (2011) online at http://www.cnpq.br/english/cnpq/index.htm, 14 May 2010. Accessed 19 May 2011
20. Kay L, Shapira P (2009) Developing nanotechnology in Latin America. J Nanopart Res 11:259–278
21. Nanoforum EULA report (2008) Fact finding mission nanotechnology. In: Brazil. Sixth Framework Programme for Research and Technological Development (September 2008). Online at http://www.utwente.nl/mesaplus/nanoforumeula/nanotechnology_in_brasil/nanoforumeula_final_report_ffm.pdf. Accessed 3 May 2011
22. Technology Strategy Board, Nanoscale Technologies Strategies Report (2009). Online at http://www.innovateuk.org/_assets/pdf/Corporate-Publications/NanoscaleTechnologiesStrategy.pdf. Accessed 4 May 2011
23. Technology Strategy Board (2011) online at http://www.innovateuk.org/aboutus.ashx. Accessed 4 May 2011
24. Federal Ministry of Education and Research (2004) Nanotechnology conquers markets, German innovation initiative for nanotechnology, 2004. Online at http://www.bmbf.de/pub/nanotechnology_conquers_markets.pdf. Accessed 3 May 2011
25. Federal Ministry of Education and Research (9 Nov 2009) online at http://www.bmbf.de/en/nanotechnologie.php. Accessed 14 May 2010
26. Vahtra P (2010) The Rusnano Corporation and internationalisation of Russia's nanotech industry. Electronic Publication of Pan-European Institute. Online at http://www.tse.fi/FI/yksikot/erillislaitokset/pei/Documents/Julkaisut/Vahtra2.pdf. Accessed 3 May 2011
27. Rusnano (2009) Transforming technology into business. Russian Corporation of Nanotechnologies Annual Report. http://www.rusnano.com/Section.aspx/Show/26528. Accessed 3 May 2011
28. Yamaguchi Y, Komiyama H (2001) Structuring knowledge project in nanotechnology material program launched in Japan. J Nanopart Res 3:105–110

29. Xie S (2011) Funding and network for nanotechnology in China. National Center for Nanotechnology in China, Institute of Physics. Chinese Academy of Science. http://mri63.adtech.icair.org/pls/portal/docs/PAGE/IVIPG/f_info/report/Xie-ICAS-Chinar.pdf. Accessed 3 May 2011
30. Italian Trade Commission (2009) Market Report on China Biotechnology and Nanotechnology Industries, Market Report. http://www.ice.gov.it/paesi/asia/cina/upload/174/Market%20Report%20on%20China%20Biotechnology%20and%20Nanotechnology%20Industries.pdf. Accessed 3 May 2011
31. Wieczorek I (2007) Nanotechnology in Korea—actors and innovative potential, Physica, Heidelberg, ISBN 978-3-7908-1973-7
32. Department of Science and Technology (2010) Nano Mission. http://nanomission.gov.in/. Accessed 12 Oct 2010.
33. Varadarajan N (2008) Taking baby steps into nano world, Business Today, 12. http://businesstoday.intoday.in/index.php?option=com_content&task=view&id=5150. Accessed 4 May 2011
34. The Energy and Resource Institute (2009) The nanotechnology development in India- a status report. http://www.teriin.org/nano-uploads/D5_NT_Development_in_India_Apri_2010.pdf. Accessed May 3, 2011.
35. Australian Academy of Science (2009) Nanotechnology in Australia: Trends, application and collaborative opportunities. http://www.science.org.au/reports/documents/nanotechnology09.pdf. Accessed 3 May 2011
36. Australian Office of Nanotechnology (2011) National Nanotechnology Strategy (NNS) Annual Report (2007–08). http://www.innovation.gov.au/Industry/Nanotechnology/NationalEnabling-TechnologiesStrategy/Documents/NNSAnnualReport_2007-08.pdf. Accessed 3 May 2011
37. Organization for Economic Co-operation and Development (2009). An overview based on indicators and statistics. Directorate for Science, Technology and Industry. http://www.oecd.org/dataoecd/59/9/43179651.pdf. Accessed 4 May 2011
38. Ellison DH (2007) Handbook of chemical and biological warfare agents, vol 2. CRC Press, Boca Raton
39. Li Y-X, Koper O, Atteya M, Klabunde KJ (1992) Adsorption and decomposition of organophosphorus compounds on nanoscale metal oxide particles. In situ GC-MS studies of pulsed microreactions over magnesium oxide. Chem Mater 4:323–330
40. Klabunde K, Stark J, Mohs C, Park D, Decker S, Jiang Y, Lagadic I, Zhang D (1996) Nanocrystals as stoichiometric reagents with unique surface chemistry. J Phys Chem 100:12142–12153
41. Stark JV, Park DG, Lagadic I, Klabunde KJ (1996) Nanoscale metal oxide particles/clusters as chemical reagents. Unique surface chemistry on magnesium oxide as shown by enhanced adsorption of acid gases (sulfur dioxide and carbon dioxide) and pressure dependence. Chem Mater 9:1904–1912
42. Wagner GW, Bartram PW, Koper O, Klabunde KJ (1999) Reactions of VX, GD and HD with nanosize MgO. J Phys Chem B 103:3225–3228
43. Winter M, Hamal D, Yang X, Kwen H, Jones D, Rajagopalan S, Klabunde KJ (2009) Defining reactivity of solid sorbents: what is the most appropriate metric? Chem Mater 21:2367–2374
44. Trubitsyn DA, Vorontsov AV (2005) Experimental study of dimethyl methylphsphonate decomposition over anatase TiO_2. J Phys Chem B 109:21884–21892
45. Panayotov DA, Morris JR (2009) Thermal decomposition of a chemical warfare agent simulant (DMMP) on TiO_2: adsorbate reaction with lattice oxygen as studied by infrared spectroscopy. J Phys Chem C 113:15684–15691
46. Panayotov DA, Morris JR (2009) Uptake of a chemical warfare agent simulant (DMMP) on TiO_2: reactive adsorption and active site poisoning. Langmuir 25:3652–3658
47. Quenneville J, Taylor RS (2010) Reactive molecular dynamics studies of DMMP adsorption and reactive on amorphous silica surfaces. J Phys Chem C 114:18894–18902
48. Panayotov DA, Paul DK, Yates JT Jr (2003) Photocatalytic oxidation of 2-chloroethyl ethyl sulfide on TiO_2-SiO_2 powders. J Phys Chem B 107:10571–10575
49. Thompson TL, Panayotov DA, Yates JT Jr, Martyanov I, Klabunde KJ (2004) Photodecomposition of adsorbed 2-chloroethyl ethyl sulfide on TiO_2: involvement of lattice oxygen. J Phys Chem B 108:17857–17865
50. Narske RM, Klabunde KJK, Fultz S (2002) Solvent effects on the heterogeneous adsorption and reactions of (2-chloroethyl) ethyl sulfide on nanocrystalline magnesium oxide. Langmuir 18:4819–4825
51. Martin ME, Narske RM, Klabunde KJ (2005) Mesoporous metal oxides formed by aggregation of nanocrystals. Behavior of aluminum oxide and mixtures with magnesium oxide in destructive adsorption of the chemical warfare surrogate 2-chloroethylethyl sulfide. Micropor Mesopor Mat 83:47–50
52. Kim S, Byl O, Liu JC, Johnson JK, Yates JT Jr (2006) Spectroscopic measurement of diffusion kinetics through subnanometer and larger Al_2O_3 particles by a new method: the interaction of 2-chloroethylethyl sulfide with γ-Al_2O_3. J Phys Chem B 110:9204–9210
53. Carnes CL, Kapoor PN, Klabunde KJ (2002) Synthesis, characterization, adsorption studies of nanocrystalline aluminum oxide and a bimetallic nanocrystalline aluminum oxide/magnesium oxide. Chem Mater 14:2922–2929
54. Carnes CL, Klabunde KJ (2000) Synthesis, isolation, and chemical reactivity studies of nanocrystalline zinc oxide. Langmuir 16:3764–3772
55. Carnes CL, Stipp J, Klabunde KJ (2002) Synthesis, characterization, and adsorption studies of nanocrystalline copper oxide and nickel oxide. Langmuir 18:1352–1359
56. Wagner GW, Koper OB, Lucas E, Decker S, Klabunde KJ (2000) Reactions of VX, GD, and HD with nanosize CaO: autocatalytic dehydrohalogenation of HD. J Phys Chem B 104:5118–5123
57. Wagner GW, Procell LR, O'Conner RJ, Munavalli S, Carnes CL, Kapoor PN, Klabunde KJ (2001) Reactions of VX, GB, GD, and HD with nanosize Al_2O_3 formation of aluminophosphonates. J Am Chem Soc 123:1636–1644
58. Wagner GW, Chen Q, Wu Y (2008) Reactions of VX, GD, and HD with nanotubular Titania. J Phys Chem C 112:11901–11906
59. Koper O, Lagadic I, Klabunde KJ (1997) Destructive adsorption of chlorinated hydrocarbons on ultrafine (nanoscale) particles of calcium oxide. 2. Chem Mater 9:838–848
60. Koper O, Klabunde KJ (1997) Destructive adsorption of chlorinated hydrocarbons on ultrafine (nanoscale) particles of calcium oxide. 3. Chloroform, trichloroethene, and tetrachloroethene. Chem Mater 9:2481–2485
61. Volodin AM, Bedilo AF, Heroux DS, Zaikovskii VI, Mishakov IV, Cheshokov VV, Klabunde KJ (2006) Nanoscale oxides as destructive sorbents for halogenated hydrocarbon. In: Blitz J, Gun'ko V (eds) Surface chemistry in biomedical and environmental science. Springer, New York, pp 403–412
62. Decker SP, Klabunde JS, Khaleel A, Klabunde KJ (2002) Catalyzed destructive adsorption of environmental toxins with nanocrystalline metal oxides. Fluoro-, chloro-, bromocarbons, sulfur, and organophosphorous compounds. Environ Sci Technol 36:762–768
63. Khaleel A, Lucas E, Pates S, Koper OB, Klabunde KJ (1999) Nanocrystals as adsorbents for chemical agents and air pollutants. Proc ERDEC Sci Conf Chem Bio Def Res, 323–329
64. Khaleel A, Kapoor PN, Klabunde KJ (1999) Nanocrystalline metal oxides as new adsorbents for air purification. Nanostruct Mater 11:459–468

65. Singh M, Zhou N, Paul DK, Klabunde KJ (2008) IR spectral evidence of aldol condensation: acetaldehyde adsorption over TiO_2 surface. J Catal 260:371–379
66. Klabunde KJ, Carnes CL (2002) Unique chemical reactivities of nanocrystalline metal oxide toward hydrogen sulfide. Chem Mater 14:1806–1811
67. Medine GM, Klabunde KJ, Zaikovskii V (2002) Unusual behavior of nanocrystalline strontium oxide toward hydrogen sulfide. J Nanopart Res 4:357–366
68. Koper OB, Klabunde JS, Marchin GL, Stoimenov P, Bohra L (2002) Nanoscale powders and formulations with biocidal activity toward spores and vegetative cells of bacillus species, viruses and toxins. Curr Microbiol 44:49–55
69. Haggstrom JA, Stoimenov PK, Klabunde KJ (2008) Synthesis and characterization of nanosized halogenated and interhalogenated metal oxide adducts. Chem Mater 20:3174–3183
70. Stoimenov P, Klinger RL, Marchin GL, Klabunde KJ (2002) Metal oxide nanoparticles as bactericidal agents. Langmuir 18:6679–6686
71. Stoimenov P, Klabunde KJ (2005) Nanotechnology in biological agent decontamination (Chapter 14). In: Kumar CSSR, Hormes J, Leuschner C (eds) Nanofabrication towards biomedical applications. Wiley, Weinheim
72. Hamal DB, Haggstrom JA, Marchin GL, Ikenberry MA, Hohn K, Klabunde KJ (2009) Multifunctional biocide/sporocide and photocatalyst based on titanium dioxide (TiO_2) codoped with silver, carbon, and sulfur. Langmuir 26:2805–2810
73. NanoScale Corporation (2011) www.nanoscalecorp.com. Accessed 27 Jan 2011
74. Guild Associates (2011) http://www.guildassociates.com/. Accessed 27 Jan 2011
75. National Science Foundation (2003) Future directions in catalysis: structure that function on the nanoscale, NSF Workshop Report, 19 June 2003
76. Semagina N, Kiwi-Minsker L (2009) Recent advances in the liquid-phase synthesis of metal nanostructures with controlled shape and size for catalysis. Cataly Rev 51:147–217
77. World Technology Evaluation Center (2009) International assessment of research and development in catalysis by nanostructured materials. NSF, DOE, AFOSR, DTRA
78. Li Y, Somorjai G (2010) Nanoscale advances in catalysis and energy applications. Nano Lett 10:2289–2295
79. Hagen J (2006) Industrial catalysis: a practical approach. Wiley, Weinheim
80. U.S. Department of Energy (2006) Estimated energy saving and financial impacts of nanomaterials by design on selected applications in the chemical industry, industrial technologies program. http://www.chemicalvision2020.org/pdfs/LANL_Nano_Rept_Final_4-17-06.pdf. Accessed 5 May 2011
81. Davis B (2009) Nanocatalysis for fuel production (Chapter 11). In: Garcia-Martinez J (ed) Nanotechnology for the energy challenge. Wiley, Weinheim
82. Prestvik R, Moljord K, Grande K, Holmen A (2004) Compositional analysis of naphtha and reformate (Chapter 1). In: Antos GJ, Aitani AM (eds) Catalytic naphtha reforming. Marcel Dekker, Inc., New York
83. Gary J, Handwerk G (2001) Petroleum refining technology and economics. Marcel Dekker, New York
84. Zhou B, Balee R, Groenedaal R (2005) Nanoparticle and nanostructure catalysis: technologies and markets. Nanotechnology Law & Business (NLB) 2:222–229
85. Wikipedia (2011) Fluid catalytic cracking. http://en.wikipedia.org/wiki/Fluid_catalytic_cracking. Accessed 22 Apr 2011
86. Ashby M, Ferreira P, Schodek D (2009) Nanomaterials and nanotechnologies in health and the environment (Chapter 11). In: Nanomaterials, nanotechnologies and design. Elsevier, New York
87. Ahmed S, Rasul M, Martens WN, Brown R, Hashib MA (2010) Advances in heterogeneous photocatalytic degradation of phenols and dyes in wastewater: a review. Water Air Soil Poll 215:3–29
88. Garcia-Ripoll A, Arques A, Vicente R, Domenech A, Amat AM (2008) Treatment of aqueous solutions containing four commercial pesticides by means of TiO_2 solar photocatalysis. J Sol Energ Eng 130:041011–041016
89. Garcia-Ripoll A, Amat A, Arques A, Vicente R, Lopez MF, Oller I, Maldonado MI, Gernjak W (2007) Increased biodegradability of Ultracid in aqueous solutions with solar TiO_2 photocatalysis. Chemosphere 68:293–300
90. Akurati K (eds) (2008) Hetrogeneous photocatalysis (Chapter 2). In: Synthesis of TiO_2 based nanoparticles for photocatalytic application. Cuvillier, Gottingen
91. Veronovski N, Rudolf A, Smole MS, Kreze T, Gersak J (2009) Self-cleaning and handle properties of TiO_2-modified textiles. Fiber Polym 10:551–556
92. Lee S, Yun C, Hahn M, Lee J, Yi J (2008) Synthesis and characterization of carbon-doped titania as visible-light-sensitive photocatalyst. Korean J Chem Eng 25:892–896
93. Yamashita H, Harada M, Misaka J, Takeuchi M, Neppolian B, Anpo M (2003) Photocatalytic degradation of organic compounds diluted in water using visible light-responsive metal ion-implanted TiO_2 catalysts: Fe ion-implanted TiO_2. Catal Today 84:191–196
94. Asahi R, Morikawa T, Ohwaki T, Aoki K, Taga Y (2001) Visible-light photocatalysis in nitrogen-doped titanium oxides. Science 293:269–271
95. Lliev V, Tomova D, Bilyarska L (2003) Phthanlocyanine modified TiO_2 or WO_3- catalysts for photooxidation of sulfide and thiosulfate ions upon irradiation with visible light. J Photochem Photobiol A 159:281–287
96. Chengyu W, Huamei S, Ying T (2003) Properties and morphology of CdS compounded TiO_2 visible-light photocatalytic nanofilms coated on glass surface. Sep Purif Techno 32:357–362
97. Shchukin D, Poznyak S, Kulak A, Pichat P (2004) TiO_2-In_2O_3 photocatalysts: preparation, characterizations and activity for 2-chlorophenol degradation in water. J Photochem Photobiol A 162:423–430
98. Sathish M, Viswanathan B, Viswanath RP, Gopinath C (2005) Synthesis, characterization, electronic structure, and photocatalytic activity of nitrogen-doped TiO_2, nanocatalyst. Chem Mater 17:6349–6353
99. Chen Y, Dionysiou D (2008) Sol–gel synthesis of nanostructured TiO_2 films for water purification (Chapter 4). In: Innocecenzi P, Zub YL, Kesslor VG (eds) Sol-gel methods for materials processing. Springer Science + Business BV, Dordrecht
100. Sasaki M (2001) Development of photocatalytic TiO_2 coated curtain. R&D Review of Toyota CRDL 36:1–2
101. Pagliaro M (2010) Nano-age: how nanotechnology changes our future (Chapter 4). Wiley, Weinheim
102. Nanotechnology Now Website (2011) Global nanocatalysts market to reach $6.0 Billion by 2015, According to new report by Global Industry Analysts, Inc. http://www.nanotech-now.com/news.cgi?story_id=32763. Accessed 22 Apr 2011
103. Green Millennium Photocatalyst (2011) http://www.greenmillennium.com/. Accessed 22 Apr 2011
104. Nanostellar Website (2011) http://www.nanostellar.com/products. Accessed 22 Apr 2011
105. Altimate EnviroCare website (2011) http://www.altimateenvirocare.com/. Accessed 22 Apr 2011
106. U.S. EPA (2004) Cleaning up the nation's waste sites: markets and technology trends. EPA 542-R-04-015 (2004).

http://www.clu-in.org/download/market/2004market.pdf. Accessed 5 May 2011
107. CSIRO (2006) Climate change and the risk to water supply. CSIRO press release 06/191. http://www.csiro.au/news/ps2cd.html. Accessed 5 May 2011
108. NASA (2003) NASA ties El Nino-induced drought to air pollution from fires. NASA press release 03-128, Washington, DC
109. Rubinkam M (2010) Pennsylvania regulators shut down Cabot drilling (April 15, 2010). www.pressconnects.com. Accessed 12 Oct 2010
110. Wikipedia (2010) Deepwater horizon oil spill. http://en.wikipedia.org/wiki/Deepwater_Horizon_oil_spill. Accessed 31 Oct 2010
111. Mohan D, Pittmann C Jr (2007) Arsenic removal from water/wastewater using adsorbents—a critical review. J Hazard Mater 142:1–53
112. Hathway SW, Rubel JF (1987) Removing arsenic from drinking water. J Am Water Words Assoc 79:61–65
113. Pena ME, Korfiatis GP, Patel M, Lippincott L, Meng X (2005) Adsorption of As(V) and As (III) by nanocrystalline titanium dioxide. Water Res 39:2327–2337
114. Gillham RW, O'Hannesin SF (1994) Enhanced degradation of halogenated aliphatics by zero-valent iron. Ground Water 32:958–967
115. Wang C-B, Zhang W-X (1997) Synthesizing nanoscale iron particles for rapid and complete dechlorination of TCE and PCBs. Environ Sci Technol 31:2154–2156
116. Zhang W-X, Wang C-B, Lien H-L (1998) Treatment of chlorinated organic contaminants with nanoscale bimetallic particles. Catal Today 40:387–395
117. Lien H-L, Zhang W-X (2001) Nanoscale iron particles for complete reduction of chlorinated ethenes. Colloid Surface 191:97–105
118. Schrick B, Blough L, Jones AD, Mallouk TE (2002) Hydrochlorination of trichloroethylene to hydrocarbons using bimetallic nickel-iron nanoparticles. Chem Mater 14:5140–5147
119. Lai KCK, Lo IMC (2008) Removal of Cr(VI) by acid-washed zero-valent iron under various groundwater geochemistry conditions. Environ Sci Technol 42:1238–1244
120. Kanel SR, Manning B, Charlet L, Choi H (2005) Removal of arsenic(III) from groundwater by nanoscale zero-valent iron. Environ Sci Technol 39:1291–1298
121. Powell RM, Puls RM, Hightower SK, Sabatini DA (1995) Coupled iron corrosion and chromate reduction-mechanisms for subsurface remediation. Environ Sci Technol 29:1913–1922
122. Blowes DW, Ptacek CJ, Jambor JL (1997) In-situ remediation of chromate contaminated groundwater using permeable reactive walls: laboratory studies. Environ Sci Technol 31:3348–3357
123. Smith EH (1996) Uptake of heavy metal in batch system by a recycled iron-bearing materials. Water Res 30:2424–2434
124. Lai KCK, Lo IMC (2008) Removal of chromium (IV) by acid-washed zero-valent iron under various groundwater geochemistry conditions. Environ Sci Technol 42:1238–1244
125. Kanel SR, Nepal D, Manning B, Choi H (2007) Transport of surface-modified iron nanoparticle in porous media and application to arsenic(III) remediation. J Nanopart Res 9:725–735
126. Zhang W-X (2003) Nanoscale iron particles for environmental remediation. J Nanopart Res 5:323–332
127. Levy E (2007) Nanocrystal-containing filtration media. U. S. Patent 7,264,726 (4 Sept 2007)
128. Levy E (2007) Transition metal oxide-aluminosilicate purification media. U. S. Patent 7,288,498 (30 Oct 2007)
129. Gusseme BD, Sintubin L, Baert L, Thibo E, Hennebel T, Vermeulen G, Uyttendaele M, Verstraete W, Boon N (2010) Biogenic silver for disinfection of water contaminated with viruses. Appl Environ Microbiol 76:1082–1087
130. Soil Composition (2010) http://student.britannica.com/comptons/article-208229/soil. Accessed 4 Feb 2010
131. Soils (2010) http://esa21.kennesaw.edu/activities/soil/soilcomposition.pdf. Accessed 4 Feb 2010
132. Soil Contaminants (2010) http://www.epa.gov/ebtpages/pollsoilcontaminants.html. Accessed 4 Feb 2010
133. Minceva-Sukarova B, Jovanovski G, Makreski P, Soptrajanov B, Willis R, Grzetic IG (2003) Vibrational spectra of $M^IM^{III}S2$ synthetic materials (M^I=Ti or Ag, M^{III}=As or Sb). J Mol Struct 651–653:181–189
134. Nevado JJB, Bermejo LFG, Martin-Dolmeadios RCR (2003) Distribution of mercury in the aquatic environment at Almadén, Spain. Environ Pollut 122:261–271
135. Ferrara R, Maserti BE, Anderson M, Edner H, Ragnarson P, Svanberg S, Hernandez A (1998) Atmospheric mercury concentration and fluxes in the Almaden district. Atmos Environ 32:3297–3904
136. Gray JE (2003) Geologic studies of mercury by the U.S. Geological Survey, U.S. Geological Survey Circular, 1248. USGS, Reston
137. Jin X, Hu J, Ong S (2007) Influence of dissolved organic matter on estrone removal by NF membranes and the role of their structures. Water Res 41:3077–3088
138. Varanasi P, Fullana A, Sidhu S (2007) Remediation of PCB contaminated soils using iron nanoparticles. Chemosphere 66:1031–1038
139. Kanel SR, Choi H (2007) Transport characteristics of surface-modified nanoscale zero-valent iron in porous media. Water Sci Technol 55:157–162
140. Xu Y, Zhau D (2007) Reductive immobilization of chromate in soils and groundwater by stabilized zero-valent iron nanoparticles. Water Res 41:2101–2108
141. Mackenzie K, Schierz A, Georgi A, Kopinke F-D (2008) Colloidal activated carbon and carbon-iron: novel materials for in-situ ground water treatment. Global Nest J 10:54–61
142. Schrick B, Hydutsky W, Blough JL, Mallouk TE (2004) Delivery vehicles for zerovalent metal nanoparticles in soil and groundwater. Chem Mater 16:2187–2193
143. Nuxoll EE, Shimotori T, Arnold WA, Cussler EL (2003) Iron nanoparticles in reactive environmental barriers. AIChE Annual Meeting, Minneapolis
144. De Windt W, Aelterman P, Verstraete W (2005) Bioreductive deposition of palladium (0) nanoparticles on Shewanella oneidensis with catalytic activity towards reductive dechlorination of polychlorinated biphenyls. Environ Microbiol 7:314–325
145. Tan X, Wang X, Chen C, Sun A (2006) Effect of soil humic and fulvic acids, pH and ionic strength on Th(IV) sorption to TiO_2 nanoparticles. Appl Radiat Isotopes 65:375–381
146. Wei D, Radakrisnan R, Venderspurt TH (2004) Bifunctional manganese oxide/titanium dioxide photocatalyst/thermocatalyst for improving indoor air quality. U.S. Patent Application # 2004/0258581 (23 Dec 2004). USEPA, Washington, DC
147. Suzuki K, Sinha A (2007) Monodisperse, bimodal mesoporous ceria catalysts and adsorbents for air purification. J Mater Chem 17:2547–2551
148. Sinha AK, Suzuki K (2005) Preparation and characterization of novel mesoporous ceria-titania. J Phys Chem B 109:1708–1714
149. Gauthier-Manuel B, Pichonat T (2005) Nanostructured membranes: a new class of protonic conductor for miniature fuel cells. J Nanotechnol Online. http://www.azonano.com/article.aspx?ArticleID=1303. Accessed 25 May 2011
150. Zorne S, David A (2007) Open electric circuits optimized in supercritical fluids that coexist with non supercritical fluid thin films to synthesis nano scale products and energy production. International Patent Application # WO/2007/117274, 18 Oct 2007
151. Khaleel A, Lucas E, Pates S, Koper OB, Klubunde KJ (1999) Nanocrystals as adsorbents for chemical agents and air pollutants. Proc ERDEC Sci Conf Chem Biol Def Res:323–329

152. Decker S, Klabunde KJ (1996) Enhancing effect of Fe_2O_3 on the ability of nanocrystalline calcium oxide to adsorb SO_2. J Am Chem Soc 118:12465–12466
153. U.S. EPA Deactivation and regeneration of environmentally exposed Titanium Dioxide (TiO_2) based products (2003). Testing report of the environmental protection department, HKSAR. http://www.epd.gov.hk/epd/english/resources_pub/publications/pub_reports_epr.html
154. Akbari H, Berdahl P (2008) PIER final project report. www.energy.ca.gov/2007publications/CEC-500-2007-112/CEC-500-2007-112.PDF
155. Sinha AK, Suzuki K, Takahara M, Azuma H, Nonaka T, Fukumoto K (2007) Mesostructured manganese oxide/gold nanoparticle composites for extensive air purification. Angew Chem Int Ed 46:2891–2894
156. Nanotechnology Now website (2011) Global market for nanotechnology slated for high growth through 2013. http://www.nanotech-now.com/news.cgi?story_id=39492. Accessed 4 Feb 2011
157. Mace C (2011) Controlling groundwater VOCs: do nanoscale ZVI particles have any advantages over microscale ZVI or BNP? (May 31, 2006). http://www.golder.com/us/en/modules.php?name=Publication&sp_id=190&page_id=212. Accessed 4 Feb 2011
158. NanoH$_2$O (2011) Nanotechnology advances reverse osmosis membrane performance. http://www.nanoh2o.com/Technology.php5?category=Nanotechnology. Accessed 4 Feb 2011
159. PBR (2011) Dow Technology's ABSORBIA reduces arsenic. http://specialitychemicals.pharmaceutical-business-review.com/news/dow_technologys_adsorbsia_reduces_arsenic_090818. Accessed 4 Feb 2011
160. Kirschling TL, Gregory KB, Minkley EG Jr, Lowry GV, Tilton RD (2010) Impact of nanoscale zero valent iron on geochemistry and microbial populations in trichloroethylene contaminated aquifer materials. Environ Sci Technol 44:3474–3480
161. Tilston EL, Cullen L, Collins CD, Shaw LJ (2009) Deployment of nZVI to soil for polychlorinated biphenyl reduction: impacts on soil microbial communities: abstract for OECD Conference on the Potential environmental benefits of nanotechnology. http://www.oecd.org/document/27/0,3343,en_2649_37015404_43278299_1_1_1_1,00.html#deployment_of_nZVI. Accessed 4 Feb 2011
162. Keane E (2009) Fate, transport, and toxicity of nanoscale zero-valent iron (nZVI) used during Superfund remediation, Fellowship Report for U.S. EPA. http://www.clu-in.org/download/studentpapers/Zero-Valen-Iron-Keane.pdf. Accessed 4 Feb 2011
163. Ahmed N, El-Shihtawy R (2010) The use of new technologies in coloration of textile fibers. J Mater Sci 45:1143–1153
164. Ronacone K (2004) Nanotechnology: what next-generation warriors will wear. JOM-J Min Met Mat S 56:31–33
165. Chow C (2011) Nanotechnology in textile industry and advantages, 2008 Second international conference on integration and commercialization of micro and nanosystems (MICRONANO2008), 3–5 June 2008. Clear Water Bay, Kowloon, Hong Kong Sponsor: Nanotechnology Institute. http://scitation.aip.org/getabs/servlet/GetabsServlet?prog=normal&id=ASMECP002008042940000761000001&idtype=cvips&gifs=yes&ref=no. Accessed 25 May 2011
166. Frykberg E (2007) Disaster and mass casualty management (Chapter 16). In: Britt LD, Trunkey DD, Feliciano DV (eds) Acute Care Surgery Principles and Practice. Springer Science + Business Media, New York
167. Wikipedia (2011) Chemical warfare. http://en.wikipedia.org/wiki/Chemical_Warfare. Accessed on 6 May 2011
168. Cheung SS, McLellan TM (1997) Influence of hydration status and fluid replacement on heat tolerance while wearing NBC protective clothing. Eur J Appl Physiol O 77:139–148
169. Nunneley S (1989) Heat stress in protective clothing. Scand J Work Environ Health 15:52–57
170. Sundarrajan S, Ramakrishna S (2007) Fabrication of nanocomposite membranes from nanofibers and nanoparticles for protection against chemical warfare stimulants. J Mater Sci 42:8400–8407
171. Rajagopalan S (2010) From nanoparticles to novel protective garments, phase II SBIR final report funded by CDC/NIOSH 2 R44OH007963-02A2, NanoScale Corporation, Manhattan, KS
172. U.S. RDECOM (2010) Enhanced contaminated human remains pouch based on exceptionally responsive advanced materials, U.S. Army/DTRA sponsored project
173. Tredget E, Shankowsky H, Greoneveld A, Burrel R (1998) A matched-pair randomized study evaluating the efficacy and safety of Acticat silver-coated dressing for treatment of burn wounds. J Burn Care Rehabil 19:531–537
174. Sibbald R, Browns A, Coutts P, Queen D (2001) Screening evaluation of an ionized nanocrystalline silver dressing in chronic wound care. Ostomy Wound Manag 47:38–43
175. ObservatoryNano (2010) Medical textile, sport/outdoor textiles. Report on nanotechnology and textile. http://www.observatorynano.eu/project/filesystem/files/FocusReport2010_MedicalTextiles_Sport_OutdoorTextilesv2.pdf. Accessed at 6 May 2011
176. Warne M, Schifko A (1999) Toxicity of laundry detergent components to a freshwater cladoceran and their contribution to detergent toxicity. Ecotox Environ Safe 44:196–206
177. Veronovski N, Rudolf A, Smole M, Kreze T, Gersak J (2009) Self-cleaning and handle properties of TiO_2-modified textiles. Fibers Polym 10:551–556
178. Bhushan B, Jung Y, Nosonovsky M (2010) Lotus effect: surfaces with roughness-induced superhydrophobicity, self-cleaning, and low adhesion. Springer Handbook of Nanotechnology Part F, 1437–1524
179. Su C (2012) A simple and cost-effective method for fabricating lotus-effect composite coatings. J Coat Technol Res 9(2):135–141
180. Liu Y, Tang J, Wang R, Lu H, Li L, Kong Y, Qi K, Xin J (2007) Artificial lotus leaf structure from assembling carbon nanotubes and their application in hydrophobic textiles. J Mater Chem 17:1071–1078
181. Hsieh C, Chen W, Wu F (2008) Fabrication and superhydrophobicity of fluorinated carbon with micro/nanoscaled two-tier roughness. Carbon 46:1218–1224
182. Xu B, Cai Z (2008) Fabrication of a superhydrphobic ZnO nanorod array film on cotton fabric via a wet chemical route and hydrophobic modification. Appl Surf Sci 254:5899–5904
183. Daoud W, Xin J (2004) Nucleation and growth of anatase crystallites on cotton fabric at low temperature. J Am Ceram Soc 87:953–955
184. Li Z, Xing Y, Dai J (2008) Superhydrophobic surface prepared from glass and non-fluorinated alkyl saline on cotton substrate. Appl Surf Sci 254:2131–2135
185. Nanotechnology Enabled Sensing (2009) Report of the National nanotechnology initiative workshop (2009). http://www.nano.gov/sites/default/files/NNI-Nanosensors-stdres.pdf?q=NNI-Nanosensors-stdres.pdf. Accessed 6 May 2011
186. Choi K, Jang H (2010) One-dimensional oxide nanostructures as gas-sensing materials: review and issues. Sensors 10:4083–4099
187. Sundarrajan S, Chandrasekara A, Ramakrishna S (2010) An update on nanomaterials-based textiles for protection and decontamination. J Am Ceram Soc 93:3955–3975
188. Zhou J, Xu N, Wang Z (2006) Dissolving behavior and stability of ZnO wires in biofluids: a study on biodegradability of ZnO nanostructures. Adv Mater 18:2432–2435
189. Lim Z, Chai Z, Kevin M, Wong A, Ho G (2010) A facile approach toward ZnO nanorods conductive textile for room temperature multifunctional sensors. Sensor Actuat B-Chem 151:121–126
190. Cho G (2010) Smart clothing: technology and applications. Taylor & Francis, Boca Raton

191. Laxminarayana K, Jalili N (2005) Functional nanotube-based textiles: pathway to next generation fabric with enhanced sensing capabilities. Text Res J 75:670–680
192. Weber J, Kumar A, Kumar A, Bhansali S (2006) Novel lactate and pH biosensor for skin and sweat analysis based on single walled carbon nanotubes. Sensor Actuat B-Chem 117:308–313
193. Thilagavathi G, Raja A, Kannaian T (2008) Nanotechnology and protective clothing for defense personnel. Defense Sci J 58:451–459
194. NanoTex website (2011) www.nanotex.com. Accessed on 4 May 2011
195. NanoHorizons (2011) www.nanohorizons.com. Accessed on 4 May 2011
196. Sensatex (2010) www.sensatex.com. Accessed on 16 Aug 2010
197. Pradeep T (2007) NANO: the essentials. understanding nanoscience and nanotechnology. Tata McGraw-Hill, New Delhi
198. El-Sayed MA (2004) Small is different: shape-, size-, and composition-dependent properties of some colloidal semiconductor nanocrystals. Acc Chem Res 37:326–333
199. Wei Z, Zamborini FP (2004) Directly monitoring the growth of gold nanoparticle seeds into gold nanorods. Langmuir 20:11301–11304
200. Li J, Daniel C, Wood D (2011) Materials processing for lithium ion batteries. J Power Sources 196:2452–2460
201. Long JW, Dunn B, Rolison DR, White HS (2004) Three-dimensional battery architectures. Chem Rev 104:4463–4492
202. Teo BK, Sun XH (2007) Silicon-based low-dimensional nanomaterials and nanodevices. Chem Rev 107:1454–1532
203. Talapin DV, Lee J-S, Kovalenko MV, Shevchenko EV (2010) Prospects of colloidal nanocrystals for electric and optoelectronic applications. Chem Rev 110:389–458
204. Caminade A-M, Majoral J-P (2004) Nanomaterials based on phosphorous dendrimers. Acc Chem Res 37:341–348
205. Woehrle GH, Warner MG, Hutchison JE (2004) Molecular-level control of feature separation in one-dimensional nanostructure assemblies formed by biomolecular nanolithography. Langmuir 20:5982–5988
206. Hannon JB, Afzali A, Klinke C, Avouris P (2005) Selective placement of carbon nanotubes on metal-oxide surfaces. Langmuir 21:8569–8571
207. Díaz C, Valenzuela ML, Laguna A, Lavayen V, Jiménez J, Power LA, O'Dwyer C (2010) Metallophosphazene precursor routes to the solid-state deposition of metallic and dielectric microstructures and nanostructures on Si and SiO_2. Langmuir 26:10223–10233
208. Allan R (2006) This little nano went to market…eventually. http://electronicdesign.com/article/components/this-little-nano-went-to-market-eventually12817/2.aspx. Accessed 26 Jan 2011
209. NNI Signature Initiative (2010) Nanoelectronics for 2020 and beyond. http://www.nano.gov/NNI_SigInit_Nanoelectronics_Feb_2010.pdf. Accessed 26 Jan 2011
210. (2004) Emerging nanoelectronics markets: the next opportunity for the semiconductor and advanced materials industries. http://www.azonano.com/news.asp?newsID=392. Accessed 26 Jan 2011
211. U.S. Cancer Statistics Working Group (2011) United States Cancer Statistics: 1999–2006 incidence and mortality web-based report. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; 2010, Atlanta. www.cdc.gov/uscs. Accessed 4 Feb 2011
212. U.S. Cancer Statistics Working Group (2010) United States Cancer Statistics: 1999–2005 incidence, WONDER on-line database. United States Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; August 2008. http://wonder.cdc.gov/cancer-v2005.html. Accessed 2 Nov 2010
213. U.S. Census Bureau (2010) U.S. and World Population Clocks. http://www.census.gov/main/www/popclock.html. Accessed 2 Nov 2010
214. Fayed L (2010) Side effects of radiation therapy: what the side effects are and how to manage them. http://cancer.about.com/od/radiationtherapy/a/radiationeffect.htm. Accessed 5 Nov 2010
215. Fayed L (2010) Chemotherapy side effects: what are the side effects and how to prevent them. http://cancer.about.com/od/chemotherapy/a/chemoeffects.htm. Accessed 5 Nov 2010
216. American Cancer Society (2010) Pancreatic cancer: detailed guide, American Cancer Society, 2010. http://www.cancer.org/Cancer/PancreaticCancer/DetailedGuide/index. Accessed 5 Nov 2010
217. Yong K-T, Hu R, Roy I, Ding H, Vathy LA, Bergey EJ, Mizuma M, Anirban Maitra A, Prasad PN (2009) Tumor targeting and imaging in live animals with functionalized semiconductor quantum rods. ACS Appl Mater Interfaces 1:710–719
218. Wei A, Leonov AP, Wei Q (2010) Gold nanorods: multifunctional agents for cancer imaging and therapy. Methods Mol Biol 624:119–130
219. Wu W, Aiello M, Zhou T, Berliner A, Banerjee P, Zhou S (2010) In-situ immobilization of quantum dots in polysaccharide-based nanogels for integration of optical pH-sensing, tumor cell imaging, and drug delivery. Biomaterials 31:3023–3031
220. Huo Q (2010) Protein complexes/aggregates as potential cancer biomarkers revealed by a nanoparticle aggregation immunoassay. Colloid Surf B Interfaces 78:259–265
221. Zheng W, He L (2010) Multiplexed detection of protein cancer markers on Au/Ag-barcoded nanorods using fluorescent-conjugated polymers. Anal Bioanal Chem 397:2261–2270
222. Abdalla MO, Karna P, Sajja HK, Mao H, Yates C, Turner T, Aneja R (2011) Enhanced noscapine delivery using uPAR-targeted optical-imaging trackable nanoparticles for prostate cancer delivery. J Control Release 149:314–322
223. Yang X, Grailler JJ, Rowland IJ, Javadi A, Hurley SA, Matson VZ, Steeber DA, Gong S (2010) Multifunctional stable and ph-responsive polymer vesicles formed by heterofunctional triblock copolymer for targeted anticancer drug delivery and ultrasensitive MR imaging. ACS Nano 4:6805–6817
224. Balivada S, Rachakatla RS, Wang H, Samarakoon TN, Dani RK, Pyle M, Kroh FO, Walker B, Leaym X, Koper OB, Tamura M, Chikan V, Bossmann SH, Troyer DL (2010) A/C magnetic hyperthermia of melanoma mediated by iron(0)/iron oxide core/shell magnetic nanoparticles: a mouse study. BMC Cancer 10 (2010). http://www.biomedcentral.com/1471-2407/10/119. Accessed 8 May 2011
225. Du Y, Zhang D, Liu H, Lai R (2009) Chemotherapy effect of nanosized As_2O_3/Fe_3O_4 complex on experimental mouse tumors and its influence on the expression of CD44v6, VEGF-C, and MMP-9. BMC Biotechnol 9. http://www.biomedcentral.com/1472-6750/10/84. Accessed 8 May 2011
226. Thakare VS, Das M, Jain AK, Patil S, Jain S (2010) Carbon nanotubes in cancer theragnosis. Nanomedicine 5:1277–1301
227. Huang N, Wang H, Zhao J, Lui H, Korbelik M, Zeng H (2010) Single-wall carbon nanotubes assisted photothermal cancer therapy: animal study with a murine model of squamous cell carcinoma. Lasers Surg Med 42:638–648
228. Liu X, Tao H, Yang K, Zhang S, Lee ST, Liu Z (2010) Optimization of surface chemistry on single-walled carbon nanotubes for in vivo photothermal ablation of tumors. Biomaterials 32:144–151
229. Li Y, Lu W, Huang Q, Huang M, Li C, Chen W (2010) Copper sulfide nanoparticles for photothermal ablation of tumor cells. Nanomedicine (Lond) 5:1161–1171
230. Li L, Tang F, Liu H, Liu T, Hao N, Chen D, Teng X, He J (2010) In vivo delivery of silica nanorattle encapulated docetaxel for liver cancer therapy with low toxicity and high efficacy. ACS Nano 4:6874–6882

231. Lebret V, Raehm L, Durand JO, Smaïhi M, Werts MH, Blanchard-Desce M, Méthy-Gonnod D, Dubernet C (2010) Folic-acid targeted mesoporous silica nanoparticles for two-photon fluorescence. J Biomed Nanotechnol 6:176–180
232. Side effects of cisplatin (2010) http://www.cisplatin.org/effects.htm. Accessed 5 Nov 2010
233. Paraskar AS, Soni S, Chin KT, Chaudhuri P, Muto KW, Berkowitz J, Handlogten MW, Alves NJ, Bilgicer B, Dinulescu DM, Mashelkar RA, Sengupta S (2010) Harnessing structure-activity relationship to engineer a cisplatin nanoparticle for enhanced antitumor efficacy. Proc Natl Acad Sci U S A 107:12435–12440
234. Azzi J, Tang L, Moore R, Tong R, El Haddad N, Akiyoshi T, Mfarrej B, Yang S, Jurewicz M, Ichimura T, Lindeman N, Cheng J, Abdi R (2010) Polyactide-cyclosporin A nanoparticles for targeted immunosuppression. FASEB J 24:3927–3938
235. Jain AK, Swarnakar NK, Godugu C, Singh RP, Jain S (2010) The effect of the oral administration of polymeric nanoparticles on the efficacy and toxicity of tamoxifen. Biomaterials 32:503–515
236. Liu J, Qui Z, Wang S, Zhou L, Zhang S (2010) A modified double-emulsion method for the preparation of daunorubicin-loaded polymeric nanoparticle with enhanced in vitro anti-tumor activity. Biomed Mater 5:065002
237. Tahara K, Miyazaki Y, Kawashima Y, Kreuter J, Yamamoto H (2011) Brain targeting with surface-modified poly(DL-lactic-co-glycolic acid) nanoparticles delivered via carotid artery administration. Eur J Pharm Biopharm 77:84–88
238. Lui Z, Zhang X, Wu H, Li J, Shu L, Liu R, Li L, Li N (2010) Preparation and evaluation of solid lipid nanoparticles of baicalin for ocular drug delivery system in vitro and in vivo. Drug Dev Ind Pharm 37:475–481
239. Nagpal K, Singh SK, Mishra DN (2010) Chitosan nanoparticles: a promising system in drug delivery. Chem Pharm Bull 58:1423–1430
240. Liu D, Ge Y, Tang Y, Yuan Y, Zhang Q, Li R, Xu Q (2010) Solid lipid nanoparticles for transdermal delivery of diclofenac sodium: preparation, characterization, and in vitro studies. J Microencapsul 27:726–734
241. Lalatonne Y, Monteil M, Jouni H, Serfaty JM, Sainte-Catherine O, Lièvre N, Kusmia S, Weinmann P, Lecouvey M, Motte L (2010) Superparamagnetic bifunctional bisphosphonates nanoparticles: a potential MRI contrast agent for osteoporosis therapy and diagnostic. J Osteoporos. doi:10.4061/2010/747852
242. Treatment With Cisplatin (2010) http://www.cisplatin.org/treat.htm. Accessed 24 Nov 2010
243. National Institute of Diabetes and Digestive and Kidney Diseases (2008) National diabetes statistics, 2007, NIH Publication No. 08-3892
244. Centers for Disease Control (2010) Number of Americans with diabetes expected to double or triple by 2050. http://www.cdc.gov/media/pressrel/2010/r101022.html. Accessed on 24 Nov 2010
245. Diaz A, David A, Perez R, González ML, Báez A, Wark SE, Zhang P, Clearfield A, Colón JL (2010) Nanoencapsulation of insulin into zirconium phosphate nanoparticles for oral delivery applications. Biomacromolecules 11:2465–2470
246. Zhao X, Zu Y, Zu S, Wang D, Zhang Y, Zu B (2010) Insulin nanoparticles for transdermal delivery: preparation and physicochemical characterization and in vitro evaluation. Drug Dev Ind Pharm 36:1177–1185
247. Bi R, Shao W, Wang Q, Zhang N (2009) Solid lipid nanoparticles act as insulin inhalation carriers for enhanced pulmonary drug delivery. J Biomed Nanotechnol 5:84–92
248. Dolenc A, Govedarica B, Dreu R, Kocbek P, Srčič S, Kristl J (2010) Nanosized particles of orlistat with enhanced in vitro dissolution rate and lipase inhibition. Int J Pharm 396:149–155
249. Moon HS, Guo DD, Song HH, Kim IY, Jin HL, Kim YK, Chung CS, Choi YJ, Lee HK, Cho CS (2007) Regulation of adipocyte differentiation by PEGylated all-trans retinoic acid: reduced cytotoxicity and attenuated lipid accumulation. J Nutr Biochem 18:322–331
250. Luthi AJ, Patel PC, Ko CH, Mutharasan RK, Mirkin CA, Thaxton CS (2010) Nanotechnology for synthetic high-density lipoproteins. Trends Mol Med 16:553–560
251. Yin YD, Yeh ML, Yang YJ, Tsai D-C, Chu T-Y, Shih Y-Y, Chang M-Y, Liu Y-W, Tang ACL, Chen T-Y, Luo C-Y, Kung-Chao Chang K-C, Chen J-H, Wu H-L, Tin-Kan Hung T-K, Hsieh PCH (2010) Intramyocardial peptide nanofiber injection improves postinfarction ventricular remodeling and efficacy of bone marrow cell therapy in pigs. Circulation 122:S132–S141
252. Antoniardes C, Psarros C, Tousoulis D, Bakogiannis C, Shirodaria C, Stefanadis C (2010) Nanoparticles: a promising therapeutic approach in atherosclerosis. Curr Drug Deliv 7:303–311
253. Sinha RR, Soni S, Harfouche R, Vasudevan PR, Holmes O, de Jonge H, Rowe A, Paraskar A, Hentschel DM, Chirgadze D, Blundell TL, Gherardi E, Mashelkar RA, Sengupta S (2010) Coupling growth-factor engineering with nanotechnology for therapeutic angiogenesis. Proc Natl Acad Sci U S A 107:13608–13613
254. Hattori Y (2010) Development of non-viral vector for cancer gene therapy. Yakugaku Zasshi 130:917–924
255. Kodama T, Aoi A, Watanabe Y, Horie S, Kodama M, Li L, Chen R, Teramoto N, Morikawa H, Mori S, Fukumoto M (2010) Evaluation of transfection efficiency in skeletal muscle using nano/microbubbles and ultrasound. Ultrasound Med Biol 36:1196–1205
256. Zarbin MA, Montemagno C, Leary JF, Ritch R (2010) Nanotechnology in ophthalmology. Can J Opthalmol 45:457–476
257. Allaker RP (2010) The use of nanoparticles to control oral biofilm formation. J Dent Res 89:1175–1186
258. Zhao YK, Sung WP, Tsai TT, Wang HJ (2010) Application of silver-doped nanoscale titanium dioxide as photocatalyst for indoor airborne bacteria control: a feasibility study in medical nursing institutions. J Air Waste Manag Assoc 60:337–345
259. Cökeliler D, Göktaş H, Tosun PD, Mutlu S (2010) Infection free titanium alloys by stabile thiol based nanocoating. J Nanosci Nanotechnol 10:2583–2589
260. Business Insight Nanotechnology Industry (2010) Nanotechnology in healthcare: market outlook for applications, tools and materials, and 40 company profiles
261. Elan Drug Technologies website, NanoCrystal® Technology—the insoluble solved™ (2010) http://www.elandrugtechnologies.com/nav/14/. Accessed 3 Nov 2010
262. NanoBio Corporation website, NanoStat™ Platform (2010) http://www.nanobio.com/Platform-Technology/NanoStat-Platform.html. Accessed 12 Dec 2010
263. Park K (2007) Nanotechnology: what it can do for drug delivery. J Control Release 120:1–3. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1949907/. Accessed 12 Dec 2010
264. UnderstandingNano.com (2011) Nanotechnology in fabrics. http://www.understandingnano.com/fabrics.html. Accessed 31 Jan 2011
265. Cohen MS, Stern JM, Vanni AJ, Kelley RS, Baumgart E, Field D, Libertino JA, Summerhayes IC (2007) In vitro analysis of a nanocrystalline silver-coated surgical mesh. Surg Infect (Larchmt) 8:397–403
266. NanoPhos website (2011) http://www.nanophos.com/en/Products_en.html. Accessed 31 Jan 2011
267. IndustrialCoatingsWorld.com (2011) Glimpse of NanoMarkets report: conductive coatings markets for 2010 and beyond. http://industrialcoatingsworld.com/newsstory/glimpse-nanomarkets-report-conductive-coatings-markets-2010-and-beyond. Accessed 31 Jan 2011
268. Schmidt H (2010) Commercial success with nanomaterials. Nano Mag 12. http://www.nano.org.uk/articles/29/. Accessed 31 Jan 2011

269. Lines MG (2008) Nanomaterials for practical functional uses. J Alloys Compd 449:242–245
270. Azeredo HM, Mattoso LH, Avena-Bustillos RJ, Filho GC, Munford ML, Wood D, McHugh TH (2010) Nanocellulose chitosan composite films as effected by nanofiller loading and plasticizer content. J Food Sci 75:N1–N7
271. Rudra JS, Dave K, Haynie DT (2006) Antimicrobial polypeptide multilayer nanocoatings. J Biomater Sci Polym Ed 17:1301–1315
272. Chen X, Schluesener AJ (2008) Nanosilver: a nanoproduct in medical application. Toxicol Lett 176:1–12
273. Banerjee I, Pangule RC, Kane RS (2011) Antifouling coatings: recent developments in the design of surfaces that prevent fouling by proteins, bacteria, and marine organisms. Adv Mater 23:690–715
274. Green AA, Hersam MC (2008) Colored semitransparent conductive coatings consisting of monodisperse metallic single-walled carbon nanotubes. Nano Lett 8:1417–1422
275. Tettey KE, Yee MQ, Lee D (2010) Photocatalytic and conductive MWCNT/TiO_2 nanocomposite thin films. ACS Appl Mater Interfaces 2:2646–2652
276. Centers for Disease Control (2011) Estimates of foodborne illness in the United States. http://www.cdc.gov/foodborneburden/index.html. Accessed 17 Jan 2011
277. Souza BW, Cerqueira MA, Ruiz HA, Martins JT, Casariego A, Teixeira JA, Vicente AA (2010) Effect of chitosan-based coatings on the shelf life of salmon (Salmo salar). J Agric Food Chem 58:11456–11462
278. Juck G, Neetoo H, Chen H (2010) Application of an active alginate coating to control the growth of Listeria monocytogenes on poached and deli turkey products. Int J Food Microbiol 142:302–308
279. Azeredo HM, Mattoso LH, Wood D, Williams TG, Avena-Bustillos RJ, McHugh TH (2009) Nanocomposite edible films from mango puree reinforced with cellulose nanofibers. J Food Sci 74:N31–N35
280. Chaloupka K, Malam Y, Seifalian AM (2010) Nanosilver as a new generation of nanoproduct in biomedical applications. Trends Biotechnol 28:580–588
281. Eby DM, Luckarift HR, Johnson GR (2009) Hybrid antimicrobial enzyme and silver nanoparticle coatings for medical instruments. ACS Appl Mater Interfaces 1:1553–1560
282. Stevens KN, Crespo-Biel O, van der Bosch EE, Dias AA, Knetsch MLW, Aldenhoff YBJ, van der Veen FH, Maessen JG, Stobberingh EE, Koole LH (2009) The relationship between the antimicrobial effect of catheter coatings containing silver nanoparticles and the coagulation of contacting blood. Biomaterials 30:3682–3690
283. Dong J, Frethem C, Haustad G, Hoerr RA, Foley JD, Matuszewski MJ, Puskas JE (2009) Effect of the coating morphology on the drug release from engineered drug-polymer nanocomposites. Conf Proc IEEE Eng Med Biol Soc 2009:6006–6009
284. Lo CT, Van Tassel PR, Saltzman WM (2009) Simultaneous release of multiple molecules from poly(lactide-co-glycolide) nanoparticles assembled onto medical devices. Biomaterials 30:4889–4897
285. Junker R, Dimakis A, Thoneick M, Jansen JA (2009) Effects of implant surface coatings and composition on bone integration: a systematic review. Clin Oral Implants Res 20(s4):185–206
286. Natalio F, Link T, Müller WE, Schröder HC, Cui F-Z, Wang X, Wiens M (2010) Bioengineering of the silica-polymerizing enzyme silicaten-alpha for a targeted application to hydroxyapatite. Acta Biomater 6:3720–3728
287. Nanowerk News (2011) Nanocoating safely kills MRSA on contact. http://www.nanowerk.com/news/newsid=17659.php. Accessed 3 Feb 2011
288. Wouters M, Rentrop C, Willemsen P (2010) Surface structuring and coating performance: novel biocidefree nanocomposite coatings with anti-fouling and fouling-release properties. Prog Org Coat 68:4–11
289. Asuri P, Karajanagi SS, Kane RS, Dordick JS (2007) Polymer-nanotube-enzyme composites as active antifouling films. Small 3:50–53
290. Kananeh AB, Scharnbeck E, Hartmann D (2009) Application of antifouling surfaces in plate heat exchanger for food production. http://heatexchanger-fouling.com/papers/papers2009/21_Bani%20Kananeh_F.pdf. Accessed 21 Jan 2011
291. Nix G (2004) Selected energy conversion activities. http://www.geothermal.org/DOE_presentations/NIX_ENER.PPT#256,1. Accessed 21 Jan 2011
292. GE Water & Process Technologies (2008) Microbiological control—cooling system (Chapter 26). In: Handbook of industrial water treatment. General Electric Company (GE), GE Water & Process Technologies, Trevose
293. Adler-Abramovich L, Aronov D, Beker P, Yevnin M, Stempler S, Buzhansky L, Rosenman G, Gazit E (2009) Self-assembled arrays of peptide nanotubes by vapor deposition. Nature Nanotech 4:849–854
294. Czarneka M (2011) Self-cleaning solar panels could boost efficiency. http://techcrunch.com/2010/08/23/self-cleaning-solarpanels-could-boost-solar-efficiency/#. Accessed 23 Jan 2011
295. SurfaPore R™ (2011) http://www.nanophos.com/en/SurfaPoreR_en.html. Accessed 31 Jan 2011
296. Cayton RH (2011) Nanomaterials for abrasion-resistant coating applications. http://www.pcimag.com/Articles/Feature_Article/7f21fd89f66a7010VgnVCM100000f932a8c0____. Accessed 24 Jan 2011
297. Salata OV (2004) Applications of nanoparticles in biology and medicine. J Nanobiotechnol. http://www.jnanobiotechnology.com/content/pdf/1477-3155-2-3.pdf. Accessed 5 May 2011
298. Bauer F, Flyunt R, Czihal K, Buchmeiser MR, Langguth H, Mehnert R (2006) Nano/micro particle hybrid composites for scratch and abrasion resistant polyacrylate coatings. Macromol Mater Eng 5:493–498
299. Green AA, Hersam MC (2008) Colored semitransparent conductive coatings consisting of monodisperse metallic single-walled carbon nanotubes. Nano Lett 8:1417–1422
300. Daoud WA, Xin JH, Szeto YS (2005) Polyethylenedioxythiophene coatings for humidity, temperature and strain sensing polyamide fibers. Sens Act B Chem 109:329–333
301. Hu L, Choi JW, Yang Y, Jeong S, La Mantia F, Cui L-F, Cui Y (2009) Highly conductive paper for energy storage devices. Proc Natl Acad Sci U S A 106:21490–21494
302. Azonano (2011) Nanocoatings market expected to reach US$ 5 billion by 2013. http://www.azonano.com/news.asp?newsID=8950. Accessed 3 Feb 2011
303. Xurex (2011) Markets. http://www.xurex.com/markets. Accessed 3 Feb 2011
304. Rice University (2011) Nanomaterials poised for big impact in construction, ScienceDaily, 28 July 2010. http://www.sciencedaily.com/releases/2010/07/100728121337.htm. Accessed 3 Feb 2011
305. Research and Markets.com (2011) Nanocoatings for aerospace applications. http://www.researchandmarkets.com/reports/1442162/nanocoatings_for_aerospace_applications. Accessed 3 Feb 2011
306. Jones CF, Grainger DW (2009) In vitro assessments of nanomaterial toxicity. Adv Drug Delivery Rev 61:438–456
307. Liao CM, Chiang YH, Chio CP (2008) Model-based assessment for human inhalation exposure risk to airborne nano/fine titanium dioxide particles. Sci Total Environ 407:165–177
308. Warheit DB, Sayes CM, Reed KL (2009) Nanoscale and fine zinc oxide particles: can in vitro assays accurately forecast lung

hazards following inhalation exposures? Environ Sci Technol 43:7939–7945

309. Xia T, Zhao Y, Sager T, George S, Pokhrel S, Li N, Schoenfeld D, Meng H, Lin S, Wang X, Wang M, Ji Z, Zink JI, Mädler L, Castranova V, Lin S, Nel AE (2011) Decreased dissolution of ZnO by iron doping yields nanoparticles with reduced toxicity in the rodent lung and zebrafish embryos. ACS Nano 5:1223–1235
310. Voinov MA, Sosa Pagan JO, Morrison E, Smirnova TI, Smirnov AI (2011) Surface-mediated production of hydroxyl radicals as a mechanism of iron oxide nanoparticle biotoxicity. J Am Chem Soc 133:35–41
311. Stampoulis D, Sinha SK, White JC (2009) Assay-dependent phytotoxicity of nanoparticles to plants. Environ Sci Technol 43:9473–9479
312. Brayner R, Dahoumane SA, Yepremian C, Djediat C, Meyer M, Couté A, Fiévet F (2010) ZnO nanoparticles: synthesis, characterization, and ecotoxicological studies. Langmuir 26:6522–6528
313. Ge Y, Schimel JP, Holden PA (2011) Evidence for negative effects of TiO_2 and ZnO nanoparticles on soil bacterial communities. Environ Sci Technol 45:1659–1664
314. Roco MC (2005) Environmentally responsible development of nanotechnology. Environ Sci Technol 39:107A–112A
315. Choi JY, Ramachandran G, Kandlikar M (2009) The impact of toxicity testing costs on nanomaterial regulation. Environ Sci Technol 43:3030–3034
316. Schmid K, Riediker M (2008) Use of nanoparticles in Swiss industry: a targeted survey. Environ Sci Technol 42:2253–2260
317. Robichaud CO, Tanzil D, Weilenmann U, Wiesner MR (2005) Relative risk analysis of several manufactured nanomaterials: an insurance industry context. Environ Sci Technol 39:8985–8994
318. Conti JA, Killpack K, Gerritzen G, Huang L, Mircheva M, Delmas M, Harthorn BH, Appelbaum RP, Holden PA (2008) Health and safety practices in the nanomaterials workplace: results from an international survey. Environ Sci Technol 42:3155–3162
319. Helland A, Scheringer M, Siegrist M, Kastenholz HG, Wiek A, Scholz RW (2008) Risk assessment of engineered nanomaterials: a survey of industrial approaches. Environ Sci Technol 42:640–646
320. Meyer DE, Curran MA, Gonzolez MA (2009) An examination of existing data for the industrial manufacture and use of nanocomponents and their role in the life cycle impact of nanoproducts. Environ Sci Technol 43:1256–1262
321. Walser T, Demou E, Lang DJ, Hellweg S (2011) Prospective environmental life cycle assessment of nanosilver T-shirts. Environ Sci Technol. doi: 10.1021/es2001248
322. Botta C, Labille J, Auffan M, Borschneck D, Miche H, Cabie M, Masion A, Rose J, Bottero J-Y (2011) TiO_2-based nanoparticles released in water from commercialized sunscreens in a life-cycle perspective: structures and quantities. Environ Pollut 159:1543–1550
323. Cassee FR, Van Balen EC, Singh C, Green D, Muijser H, Weinstien J, Dreher K (2011) Exposure, health and ecological effects review of engineered nanoscale cerium and cerium oxide associated with its use as a fuel additive. Crit Rev Toxicol 41:213–229
324. Musee N (2011) Nanowastes and the environment: potential new waste management paradigm. Environ Int 37:112–128
325. Reijnders L (2006) Cleaner nanotechnology and hazard reduction of manufactured nanoparticles. J Clean Prod 14:124–133
326. DHHS (NIOSH) Publication No. 2010–104 Progress toward safe nanotechnology in the workplace, project updates for 2007 and 2008
327. DHHS (NIOSH) Publication No. 2007–123 Progress toward safe nanotechnology in the workplace
328. DHHS (NIOSH) Publication No. 2009–125 Approaches to safe nanotechnology
329. (2011) http://nanoparticlelibrary.net/index.asp, NIOSH topic: nanotechnology: nanoparticle information library. Administered by Oregon State University, created by NIOSH. Accessed 25 Apr 2011
330. Research & Markets (2011) The world market for nanotechnology and nanomaterials in consumer products, 2010–2015. http://www.researchandmarkets.com/research/5dd034/the_world_market_f 10 June 2010. Accessed 4 Feb 2011
331. Nano Tech Project website (2011) http://www.nanotechproject.org/ 9 June 2010. Accessed 4 Feb 2011
332. Nano Tech Project website (2011) http://www.nanotechproject.org/inventories/consumer/browse/products/iphone/. 9 June 2010. Accessed 4 Feb 2011
333. Nano Tech Project website (2011) http://www.nanotechproject.org/inventories/consumer/analysis_draft/ 10 June 2010. Accessed 4 Feb 2011
334. Serpone N, Dondi D, Albini A (2007) Inorganic and organic UV filters: their role and efficacy in sunscreens and suncare products. Inorg Chim Acta 360:794–802
335. Yuranova T, Laub D, Kiwi J (2007) Synthesis, activity, and characterization of textiles showing self-cleaning activity under daylight irradiation. Catal Today 122:109–117
336. Sundarajan G, Rao TN (2009) Commercial prospects for nanomaterials in India. J Indian Inst Sci 89:35–41
337. Guitterez JM, Gonzalez C, Maestro A, Sole I, Pey CM, Nolla J (2008) Nano-emulsions: new applications and optimization of their preparation. Curr Opin Colloid In 13:245–251
338. Nanotechnology Now website (2011) Nanotechnology Business Programs. http://www.nanotech-now.com/business.htm. Accessed 5 May 2011
339. IBM Research website (2011) http://researchweb.watson.ibm.com/nanoscience/ 8 June 2010. Accessed 4 Feb 2011
340. Rickerby DG (2007) Nanotechnological devices and nanopharmaceuticals: the European regulatory framework and research needs. J Nanosci Nanotechnol 7:4618–4625
341. U.S. EPA (2009) Nanomaterial research strategy, EPA 620/K-09/11. U.S. EPA, Washington, DC
342. Liu K, Sun Y, Lin X, Zhou R, Wang J, Fan S, Jiang K (2010) Scratch resistant, highly conductive, and high-strength carbon nanotube-based composite yarns. ACS Nano 4:5827–5834
343. Smith CJ, Shaw BJ, Handy RD (2007) Toxicity of single-walled carbon nanotubes to rainbow trout (Onchorhynchus mykiss): respiratory toxicity, organ pathologies, and other physiological effects. Aquat Toxicol 82:94–109
344. Brausch KA, Anderson TA, Smith PN, Maul JD (2010) Effects of functionalized fullerenes on bifenthrin and tribufos toxicity to Daphnia magna: survival, reproduction, and growth rate. Environ Toxicol Chem 29:2600–2606
345. ASTM Standard E2456, 2006 (2006) Standard terminology relating to nanotechnology. ASTM International, West Conshohocken. doi: 10.1520/E2526-08, www.astm.org
346. ASTM Standard E2578, 2007 (2006) Standard practice for calculation of mean sizes/diameters and standard deviations of particle size distributions. ASTM International, West Conshohocken, PA, doi: 10.1520/E2578-07, www.astm.org
347. ASTM Standard E2490, 2009 (2006) Standard guide for measurement of particle size distribution of nanomaterials in suspension by photon correlation spectroscopy (PCS). ASTM International, West Conshohocken, PA
348. ASTM Standard E2535, 2007 (2006) Standard guide for handling unbound engineered nanoscale particles in occupational settings. ASTM International, West Conshohocken, PA
349. ASTM Standard E2524, 2008 (2006) Standard test method for analysis of hemolytic properties of nanoparticles, ASTM International, West Conshohocken, PA

350. ASTM Standard E2525, 2008 (2006) Standard test method for evaluation of the effect of nanoparticulate materials on the formation of mouse granulocyte-macrophage colonies. ASTM International, West Conshohocken, PA
351. ASTM Standard E2526, 2008 (2006) Standard test method for evaluation of cytotoxicity of nanoparticulate materials in porcine kidney cells and human hepatocarcinoma cells. ASTM International, West Conshohocken, PA

Suggested Further Reading

Klabunde K, Richards R (2009) Nanoscale materials in chemistry, 2nd edn. Wiley, New York

Executive Office of the President (2010) Report to the President and Congress on the third assessment of the national nanotechnology initiative. President's Council of Advisors on Science and Technology, 12 Mar 2010

Organisation for Economic Co-operation and Development (2009) Inventory of National science, technology and innovation policies for nanotechnology 2008. Working Party on Nanotechnology, 17 July 2009

Cho G (2010) Smart clothing: technology and applications. Taylor & Francis, Boca Raton

ObservatoryNANO (2010) Medical textiles, sport/outdoor textiles. Report on Nanotechnology & Textiles, Apr 2010

Ventra M, Evoy S, Heflin JR (eds) (2004) Introduction to nanoscale science and technology (nanostructure science and technology), 1st edn. Springer, New York

Sahu SN, Choudary RK, Jena P (eds) (2006) Nano-scale materials: from science to technology. Nova Science, New York

Pagliaro M (ed) (2010) Nano-age: how nanotechnology changes our future. Wiley, Weinheim

Synthetic Organic Chemicals

10

Scott D. Barnicki

Introduction

Chapter Background

Synthetic organic chemicals are produced by the transformation of carbonaceous feedstocks into functionalized molecules through one or more chemical reactions. Such transformations are accomplished at vast industrial scales and the resulting products permeate every aspect of modern society. The molecules produced find use largely as monomers for polymer synthesis of ubiquitous plastics, or as task-specific ingredients for a myriad of applications as divergent as paint leveling agents to food preservatives. Advances in technology, significant increases in energy efficiency, as well as the utilization of fossil-fuel derived starting materials has resulted in unprecedented economy of scale and relatively stable product costs in spite of large relative increases in the price of oil and natural gas. The section entitled “Chemical Raw Materials and Feedstocks” covers the most important carbonaceous feedstocks currently utilized in the chemical processing industries; all derived from fossil-fuel based raw materials.

The volume of synthetic organic chemicals produced worldwide grew from less than 5 million metric tons in 1950 to more than 400 million metric tons in 2009, with an average annualized growth rate above 7% [1]. The value of all chemicals produced worldwide reached over $3.7 trillion dollars in 2008, before dropping somewhat to $3.4 trillion dollars during the 2009 recession as shown in Fig. 10.1. The average annualized worldwide growth rate in value of chemicals produced for the decade of 1999–2009 was 7.7% [2].

The United States has been the country with the largest chemical industry by volume and value for many years, due to inexpensive natural gas and very large consumer markets. However, this leadership position has deteriorated rapidly over the last decade with the dramatic growth in production of chemicals in China [2]. The value of all chemicals produced between 1999 and 2009 is shown in Fig. 10.2 for the traditional major chemical producing countries of the United States, Japan, and Germany in comparison to China. Average annualized growth rates in the value of chemicals produced over the decade for these countries are summarized in Table 10.1. The growth rate in value of chemicals produced in China far outpaces that in the United States. As a result, the US share of global chemical production value has dropped from 26% in 1999, to 20% in 2009, while that of China has rocketed from 5 to 19% in the same timeframe (Figs. 10.3 and 10.4).

The cost of a raw material can dramatically alter production patterns over a short period of time. Figure 10.5 shows chemical raw material usage of natural gas and annualized cost for natural gas for the years 1998, 2002, 2006 in the United States. After many decades near $2.00/million BTU, natural gas price in the United States began to rise and spiked dramatically in 2006 at above $9.50/MM BTU, leading to the shuttering of many older, small-scale marginal US facilities for the production of the ultra-commodities of ammonia-based fertilizers and methanol. As of 2007 only one US methanol facility based on coal, Eastman Chemical Company’s Kingsport plant, remained operational. Most of the US production moved overseas to the Middle East, Trinidad, and other regions with lower cost natural gas. Natural gas price since has dropped significantly in the US with the advent of substantial shale gas production.

Chapter Organization

The traditional approach to discussion of synthetic organic chemicals in most survey literature is to concentrate on families of derivatives from each specific chemical feedstock. Although many different chemicals can be made from each of the

---

S.D. Barnicki (✉)
Eastman Chemical Company, 1813 Buckingham Court,
Kingsport, TN 37660, USA
e-mail: sbarnick@eastman.com

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_10, © Springer Science+Business Media New York 2012

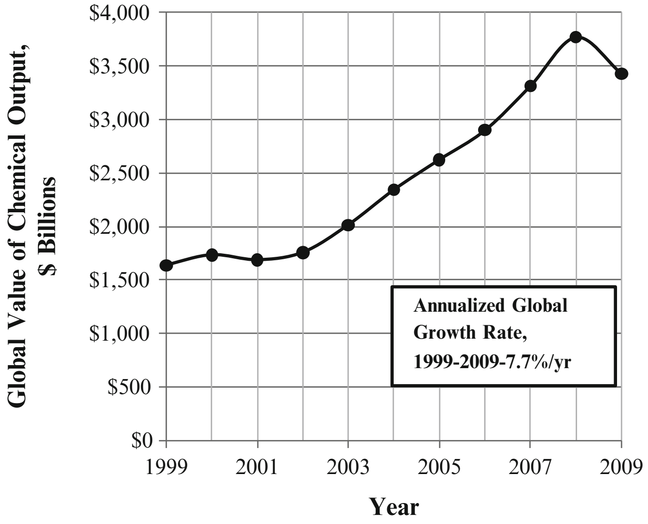


Fig. 10.1 Global chemical output, $ Billions, 1999–2009

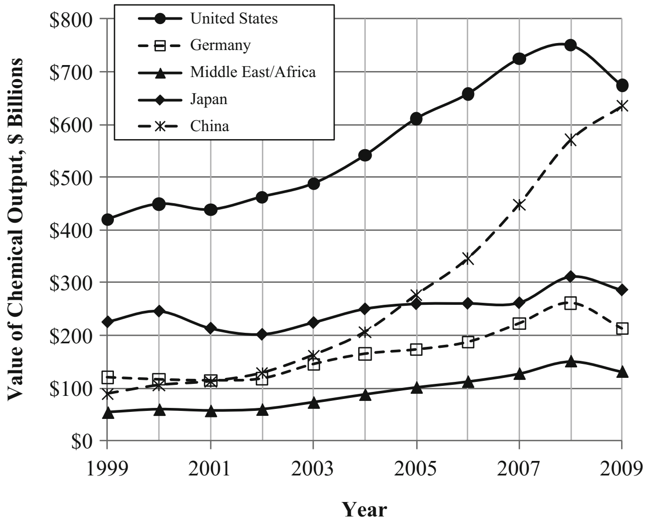


Fig. 10.2 Chemical output of selected countries, 1999–2009, $ Billions

Table 10.1 Average annualized growth rate of the value of chemical output, 1999–2009

	Average annualized growth rate (%)
United States	4.8
Other N. Amer & S. Amer	7.2
Germany	5.9
Other Europe	6.6
Japan	2.4
China	21.8
Other Asia Pacific	9.6
Middle East/Africa	9.3
Global total	7.7

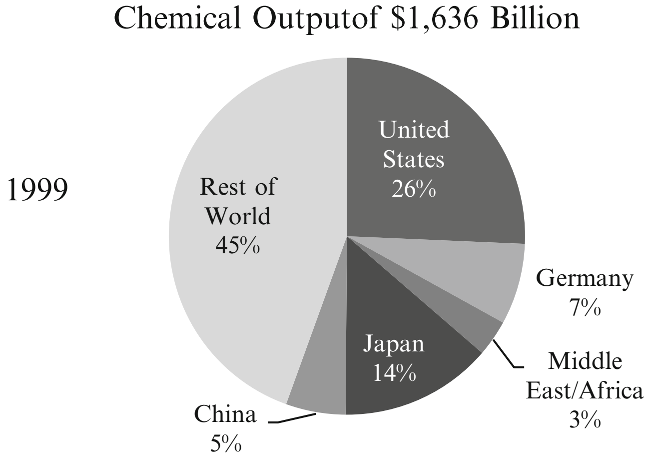


Fig. 10.3 Global chemical output, 1999, $ Billions

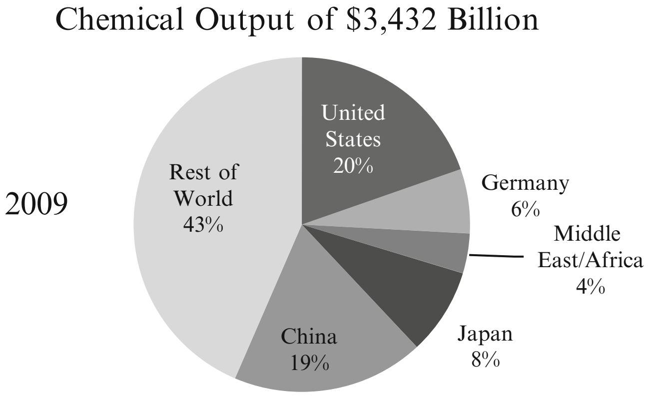


Fig. 10.4 Global chemical output, 2009, $ Billions

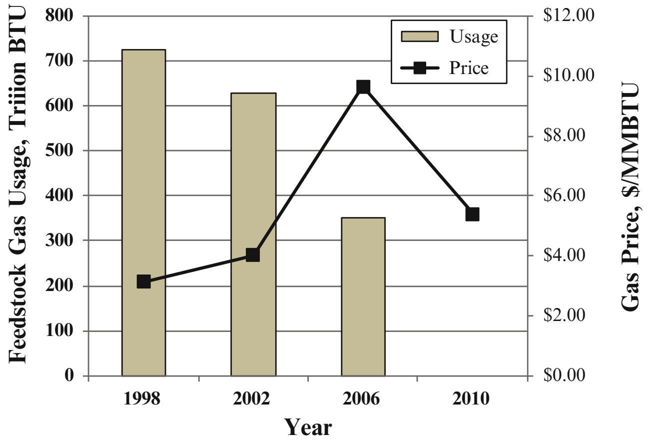


Fig. 10.5 U.S. usage of natural gas as chemical raw material vs. price, 1998–2006

feedstocks, only a limited number of different types of chemistry have found particular favor in large-scale commercial production of synthetic organic chemicals. These include: (1) oxidation, (2) carbonylation, (3) hydroformylation, (4) chlorination, (5) condensation, (6) hydration/hydrolysis, (7) esterification, (8) hydrogenation, (9) dehydrogenation, (10) sulfonation, and (11) ammonation/ammoxidation. The first nine of these are covered in this chapter. Discussion of the last two is found in other chapters of this book, in Chaps. 14, 22, and 36. Cross references within Chapter 10 are noted by subsection titles in quotation marks.

Organizing our discussion of the significant commercial processes by these chemistry themes allows one to see the commonality between many seemingly diverse production processes with very different raw materials and end products. In some cases a chemical is made commercially by very different chemistries, but the processes begin with the same feedstock. Examples include: (1) the production of alcohols by the direct or indirect hydration of olefins, (2) the various propylene-based routes to propylene oxide, and (3) the production of methyl methacrylate from ethylene by either carbonylation/esterification/condensation or hydroformylation/oxidation/esterification steps. In other cases, the feedstock may be entirely different, but end at the same product, such as the acetylene, propylene, and butane/butene-based processes for the production of 1,4-butanediol. More often, and to the point of the chapter organization, the chemistry is similar in spite of significant feedstock differences, such as hydroformylation of C_2-C_5 olefins, oxidation of aldehydes to carboxylic acids, or aldol condensations of a wide variety of carbonyls.

Often one basic chemistry route has become commercially dominant, for example ethylene oxide production via silver catalyzed oxidation of ethylene, terephthalic acid (TPA) by Co/Mn/Br catalyzed oxidation of *p*-xylene, or methanol by hydrogenation of carbon monoxide/carbon dioxide. In these cases, the key differences between commercial processes are more subtle and may amount to alternative catalysts, reactor conditions and designs, or separation sequences. A portion of these differences are highlighted in the discussion below where pertinent, but the roles of reaction engineering and separation process synthesis/design cannot be adequately covered in the space allotted. Many books and articles have been written on these process design subjects, with a few recommended texts given below [3–14].

The capacity, production, and consumption figures presented throughout the section, "Chemical Raw Materials and Feedstocks", for individual feedstocks and derivatives were derived from data gathered from a variety of literature sources, spanning the years 2006–2010 [15, 16]. Thus, some numbers reflect reduced production and consumption due to the 2008–2010 recession; some do not. The resulting variation in production/consumption numbers is estimated to be 5–10%.

The simplified flowsheets were drawn specifically for this chapter and are based on a compilation of open literature data (referenced in each subsection). The major processing steps of each flowsheet are named in italics and numbered. These designations appear as bolded numerals in parentheses in the corresponding process description in the text. Each process description includes a summary of temperature, pressure, yields, and selectivity, but open literature references can vary widely on the values.

Chemical Raw Materials and Feedstocks

Background

The formation of new a carbon–carbon bond is one of the more difficult chemical transformations to make. So the foundations of the modern chemical industry is very much a story of learning to exploit a small number of reactive skeletal carbon structures which can be derived from natural sources in relatively simple and efficient manner. In the early twentieth century the nascent chemical industry was based on utilization of coal tar liquids (a by-product of metallurgical coke production) and the fermentation of biologically derived feedstocks such as molasses, sugars, and whey. Coal tar liquids now provide a minor fraction of chemical feedstocks. Early fermentation routes, such as the ABE (acetone-butanol-ethanol) process, have all but disappeared due to poor economics. In spite of massive research efforts, newer bio-based processes have yet to compete favorably except in a few niche applications.

Rather, the substantial growth of the global chemical industry over the past 70 years would not have been possible without the concomitant rise in the fossil fuel energy infrastructure encompassing petroleum refining, natural gas production, and to a growing extent coal mining (primarily due to Chinese gasification projects). This close coupling of the chemical industry with the energy infrastructure is exemplified in the United States for the primary chemical raw materials of natural gas, petroleum products, and liquefied petroleum gases/natural gas liquids (LPG/NGL, C_2-C_4 hydrocarbons derived from crude natural gas and refinery gases). Figures 10.6, 10.7, and 10.8 show the relative magnitude of the total US energy consumption of these three key categories vs. their use as raw materials and energy sources for chemical production during the years 1998, 2002, and 2006. All quantities are expressed in trillions of BTUs of energy equivalent for easy comparison. Only LPG/NGL is used largely as a chemical raw material. Chemical raw material usages of natural gas and petroleum products represent just 1.6% and 4.5% of much larger power/heating and motor fuels markets respectively in the

Fig. 10.6 U.S. natural gas usage, 1998, 2002, and 2006

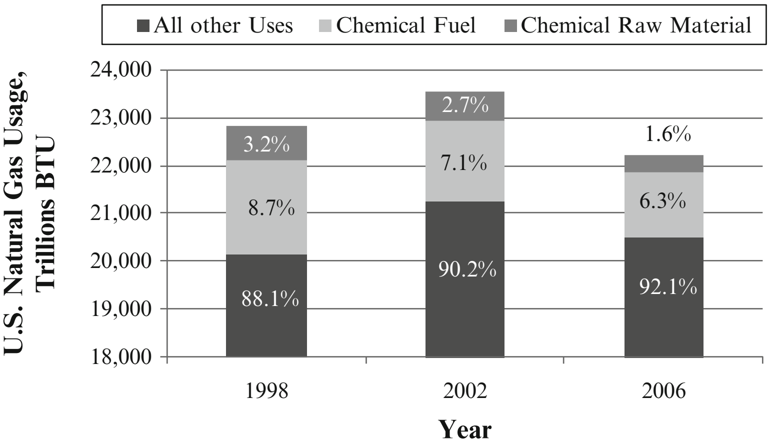


Fig. 10.7 U.S. LPG/NGL usage, 1998, 2002, and 2006

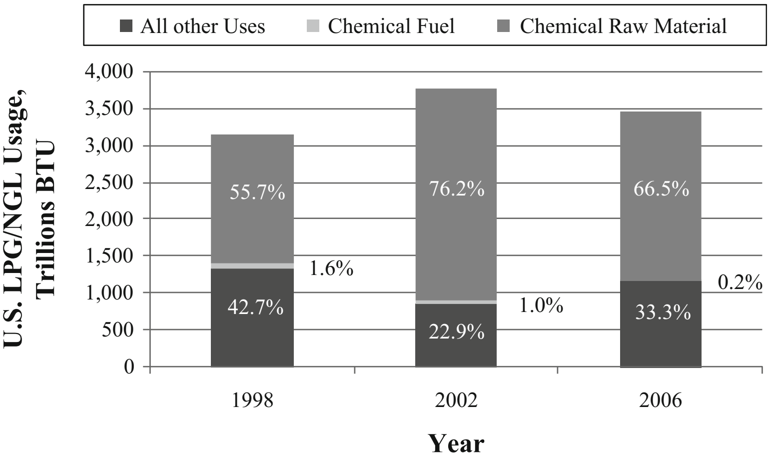
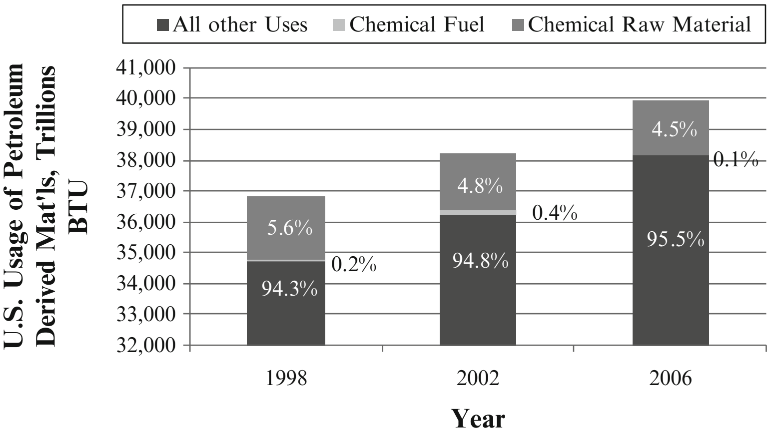


Fig. 10.8 U.S. usage of refined petroleum materials, 1998, 2002, and 2006



United States [17–19]. The situation is similar worldwide. Total global fossil fuel demand in 2007 was 282,000 trillion BTU's. About 9%, 25,500 trillion BTUs, was used as chemical raw materials, with 78.9% from petroleum, 16.9% from LPG/NGL, 3.9% from natural gas, and 0.3% from coal [20].

An overwhelming majority of the synthetic organic chemicals produced at commercial scale today begin with one of five major types of feedstocks:

1. Light olefins—ethylene and propylene
2. Aromatics—benzene, toluene, xylenes, or BTX

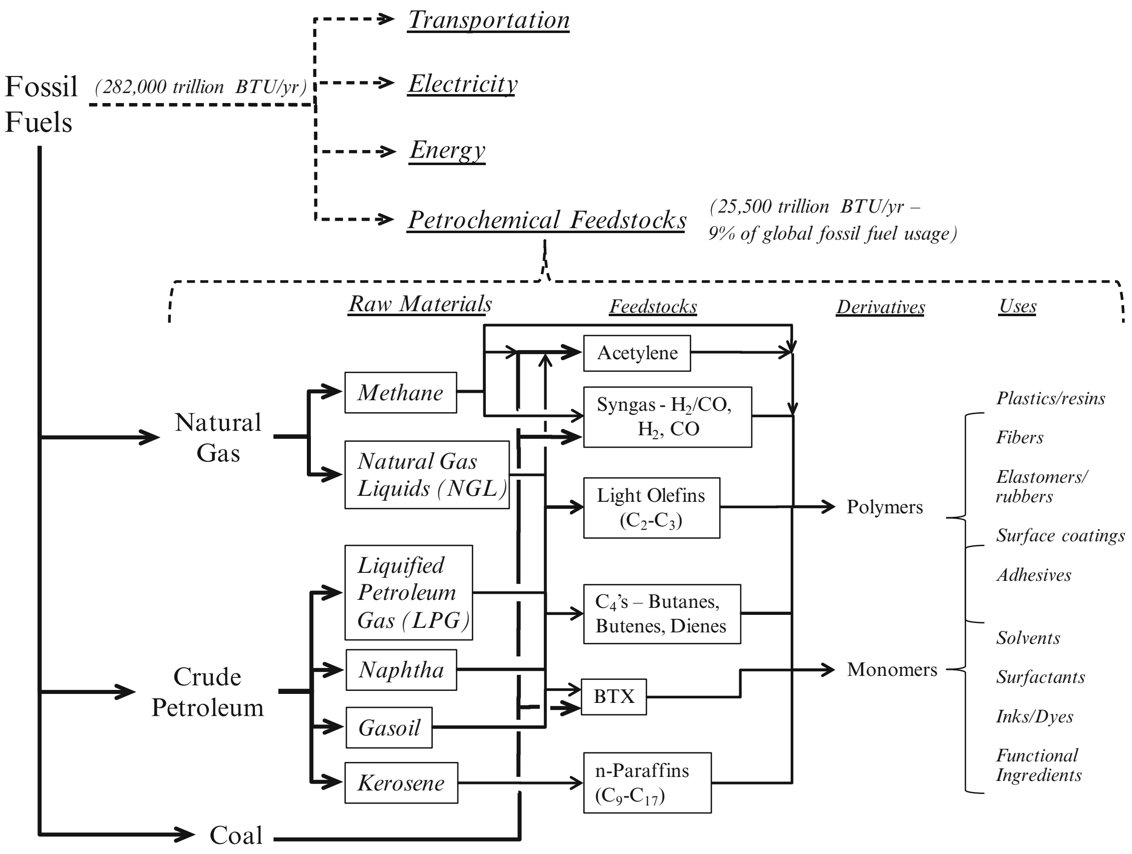


Fig. 10.9 Chemical raw material-feedstock-derivatives overview

3. C_4 hydrocarbons—butanes, butenes, butadiene
4. Kerosene derived C_9-C_{17} paraffins
5. Synthesis gas—a mixture of carbon monoxide and hydrogen

Figure 10.9 shows a simplified overview of the raw materials, feedstocks, and derivatives of the synthetic organic chemical industry. Major chemical feedstock capacity, production, and consumption for the world, United States, and China are summarized in Table 10.2.

As will become apparent below, the production of chemical feedstocks from fossil fuel raw materials typically involve highly endothermic processes. A large amount of energy must be expended for both driving endothermic reactions as well as for separation and purification. The resulting high energy, reactive feedstocks allow subsequent chemistry to proceeds favorably downhill to lower energy products, with mostly exothermic reactions and relatively high yields.

C_2-C_3 Light Olefins [21–25]

Ethylene and propylene are by far the most important building blocks of the petrochemical industry. The primary derivatives produced from ethylene and propylene feedstocks, the capacity/production/consumption of those primary derivatives, and their major applications are summarized in Tables 10.3, 10.4, 10.5, 10.6, 10.7, and 10.8. The family of derivatives produced from ethylene and propylene are presented in Figs. 10.10 and 10.11, respectively. Light olefins can be produced from a variety of raw materials and methods:

1. Steam cracking (thermal pyrolysis) of hydrocarbon raw materials ranging from LPG/NGL to naphthas and gas oils
2. Methanol to olefins
3. Recovery from refinery gases and FCC (fluid catalytic cracking) gases
4. Interconversion of butenes, ethylene, and propylene
5. Dehydrogenation of propane (propylene only)
6. Dehydration of bio-derived ethanol (ethylene only)

Steam cracking accounts for almost all of the ethylene and about 60% of the propylene produced worldwide. A simplified diagram of a naphtha-based cracker is shown in Fig. 10.12. In a thermal cracking unit, steam and the hydrocarbon raw material are heated in a short residence time tubular reactor/furnace (**1**) to 775–950°C, 0.17–0.24 MPa pressure, to initiate free radical cision, decomposition, isomerization, and aromatization reactions. The reactions are highly endothermic, requiring +1,600 to +2,800 kJ/kg of

Table 10.2 Capacity, production, and consumption of major chemical feedstocks

		Capacity, in kilo metric tons			Production, in kilo metric tons			Consumption, in KMT[a]	
Chemical feedstock		World	USA	China	World	USA	China	USA	China
Light olefins	Ethylene	128,000	27,300	11,000	113,000	24,000	10,600	22,500	11,000
	Propylene	84,000	26,500	10,500	74,000	17,000	10,000	16,800	11,000
C_4 Hydrocarbons	Total butanes	No data	No data	No data	170,000	49,000	15,000	50,000	17,000
	Butanes as Chem Fdstck	Not appl.	Not appl.	Not appl.	8,200 (est)	Not appl.	Not appl.	2,800 (est)	1,600 (est)
	Total butenes	No data	No data	No data	52,000	24,000	No data	22,000	No data
	Butenes as Chem Fdstck	No data	No data	No data	7,200	Not appl.	No data	1,700	No data
	Butadiene	12,000	2,600	1,700	9,000	1,600	1,400	2,000	1,500
BTX	Benzene	54,000	500	7,114	39,000	5,600	5,000	6,900	5,300
	Toluene	30,000	5,700	5,800	20,000	3,400	3,500	3,400	4,200
	Mixed xylenes[b]	55,000	8,400	10,500	43,000	5,800	6,300	4,900	7,100
	p-Xylene	34,000	4,500	7,000	27,000	3,300	4,300	2,400	7,700
	o-Xylene	5,300	380	1,000	3,600	360	690	310	1,400
	m-Xylene	700	180	50	560	160	40	100	75
Alkanes	C_9-C_{17} *n*-alkanes	3,400	440	420	2,900	360	370	230	530
Syngas (in MM GJ/yr equivalents of natural gas)	Hydrogen (ammonia syngas)	No data	No data	No data	3,914	No data	No data	No data	No data
	Other hydrogen	No data	No data	No data	670	No data	No data	No data	No data
	2-1 H_2/CO (methanol syngas)[c]	No data	No data	No data	1,300	No data	No data	No data	No data
	1-1 H_2/CO (oxo syngas)	No data	No data	No data	90	No data	No data	No data	No data
	Carbon monoxide	No data	No data	No data	110	No data	No data	No data	No data

[a] Gray shading indicates net importer

[b] Mixed xylenes includes *p*-,*m*-,*o*-xylenes along with materials not separated into pure isomers

[c] Does not include 2-1 H_2/CO syngas used for Fischer-Tropsch synthesis of fuels

Table 10.3 Ethylene feedstock: primary derivatives

Ethylene derivative	% of world consumption	Chemistry	Additional raw materials	Chapter cross reference
Polyethylene	59.4	Polymerization		Chapter 15
Ethylene oxide	14.6	Oxidation (epoxidation)	O_2	Chapter 10
Ethylene dichloride	11.9	Oxychlorination/addition chlorination	Cl_2, O_2	Chapter 10
Ethylbenzene	6.9	Alkylation	Benzene	Chapter 10
α-Olefns	2.3	Oligomerization		Chapter 36, Chapter 10
Vinyl acetate	1.3	Oxidative esterification	O_2, acetic acid	Chapter 10
Synthetic ethanol	0.7	Hydration	Water	Chapter 10
Acetaldehyde	0.7	Oxidation (Wacker)	Air, water	Chapter 10
Propionaldehyde	0.1	Hydroformylation	CO/H_2	Chapter 10
Others	2.1			

hydrocarbon fed, or roughly +55 to +219 kJ/gmole of ethylene/propylene produced. The yield of ethylene vs. propylene is highly dependent on the hydrocarbon raw material used as well as the severity of conditions. Ethylene production is favored by use of light hydrocarbons and higher temperature conditions. Typical yields for several common hydrocarbon feeds are given in Table 10.9. Major by-products are methane, hydrogen, butanes/butenes/butadiene, pyrolysis gasoline (benzene, toluene, C_8 aromatics), and heavy oils. Steam feed helps mitigate coking, with increasing amounts of water required for heavier feedstocks, typically 0.3/1 to 0.75/1 kg water/kg hydrocarbon.

The cracked gases are rapidly quenched to halt further reaction, first by indirect cooling against water, then by direct contact cooling with recycled heavy ends (**2**). The effluent is fractionated to separate the light gases (containing ethylene, propylene, C_4−C_5+ fractions) from pyrolysis gasoline, heavy oil fractions, and to remove water (**3**). The light gas is compressed to about 3.5 MPa, caustic washed to remove acid gases (CO_2, H_2S, COS,

Table 10.4 Ethylene feedstock: capacity/production/consumption of primary derivatives

	Capacity, in kilo metric tons			Production, in kilo metric tons			Consumption, in KMT[a]	
Chemical	World	USA	China	World	USA	China	USA	China
Polyethylene	91,00	23,000	10,000	67,000	14,000	7,200	12,000	12,000
Ethylene oxide	22,000	3,300	2,400	19,000	2,600	2,400	2,600	2,400
Ethylene dichloride	54,000	17,000	2,800	43,000	12,000	2,100	1,100	2,900
Ethylbenzene	35,000	6,000	4,200	28,000	4,800	2,800	4,800	2,800
α-Olefins	3,400	1,900	50	2,700	1,600	no data	1,500	no data
Vinyl acetate monomer (VAM)	6,200	1,700	1,300	5,200	1,400	1,100	1,000	1,400
Ethanol, synthetic and fermentation[b]	78,000	37,000	8,000	59,000	28,000	3,300	29,000	3,200
Acetaldehyde	2,600	230	1,000	1,300	110	480	110	480
Propionaldehyde	400	270	0	300	230	0	240	0

[a] Gray shading indicates net importer

[b] Only synthetic ethanol is based on ethylene. Roughly 3% of world capacity (~2,000 KMTA) and less than 2% of production (1,000 KMTA) is synthetic

Table 10.5 Major uses of primary ethylene derivatives

Chemical	Major uses
Polyethylene	31% film/sheet; 19% molding; 6% extrusion coating; piping/conduit; fibers
Ethylene oxide	77% ethylene glycol, DEG, TEG; 11% surfactants; 5% ethanolamines; 3% glycol ethers
Ethylene dichloride	>95% vinyl chloride monomer
Ethylbenzene	≈100% styrene
α-Olefins	48% polyethylene comonomers; 17% oxo alcohols; 16% poly α-olefins
VAM	42% VAM polymers/copolymers; 47% PVOH; 6% EVA
Ethanol	93% fuels; 2.3% solvent; 1.8% chemical intermediate; 2% human consumption
Ethanol as chem int. (USA)	28% vinegar; 26% acrylate esters; 17% ethanolamines; 11% ethyl acetate; 4% glycol ethers
Acetaldehyde	33% acetic acid; 21% esters; 17% pyridine derivatives; 12% pentaerythritol
Propionaldehyde	49% propanol; 32% propionic acid

Table 10.6 Propylene feedstock: primary derivatives

Propylene derivative	% of world consumption	Chemistry	Additional raw materials	Chapter cross reference
Polypropylene	62.0	Polymerization		Chapter 15
Acrylonitrile	8.2	Ammoxidation	NH_3, air	Chapter 22
Propylene oxide	7.1	Chlorohydrination/dehydrochlorination or Oxidation (hydroperoxide)	Cl_2/air or ethylbenzene/air, or isobutane/air	Chapter 10
n/*i*-Butyraldehyde	6.3	Hydroformylation	CO/H_2	Chapter 10
Cumene	6.1	Alkylation	Benzene, air	Chapter 10
Acrylic acid	3.8	Oxidation	Air	Chapter 10
Isopropanol	2.2	Dehydrogenation		Chapter 10
Propylene oligomers	1.6	Oligomerization		Chapter 18
Epichlorohydrin	0.4	Chlorohydrination/de hydrochlorination	Cl_2, water, HOCl	Chapter 10
Others	2.3			

mercaptans), further cooled, and partially liquefied (**4**). The demethanizer (**5**) removes methane and hydrogen overhead, followed by the distillation of ethane/ethylene/(acetylenes) in the deethanizer (**7**). If desired, methane and hydrogen are separated by Joule–Thomson expansion/cooling to give about 85–90% hydrogen (**6**). The ethylene-rich fraction from the deethanizer is first hydrogenated over Pd or Ni fixed bed catalysts (**8**) to convert alkynes to alkenes to prevent fouling. Propane/propylene are taken overhead in the depropanizer (**10**), and C_4 components split from

Table 10.7 Propylene feedstock: capacity/production/consumption of primary derivatives

	Capacity, in kilo metric tons			Production, in kilo metric tons			Consumption, in KMT[a]	
Chemical	World	USA	China	World	USA	China	USA	China
Polypropylene	50,000	9,200	7,500	44,000	7,800	7,100	6,600	9,600
Acrylonitrile	5,800	1,500	1,100	5,300	1,400	1,000	510	1,500
Propylene oxide	2,500	2,400	1,200	6,500	1,900	950	1,700	1,500
n-Butyraldehyde	7,600	1,600	1,000	6,600	1,300	1,000	1,300	1,000
i-Butyraldehyde	1,100	290	87	1,000	330	90	330	90
Cumene	14,000	4,300	740	12,000	3,700	710	3,500	750
Acrylic acid (and esters)	8,800	2,500	1,700	6,000	2,200	1,200	1,900	1,100
Isopropanol	2,400	730	130	1,800	720	120	480	220
Epichlorohydrin	900	460	530	580	240	320	220	370

[a] Gray shading indicates net importer

Table 10.8 Major uses of propylene primary derivatives

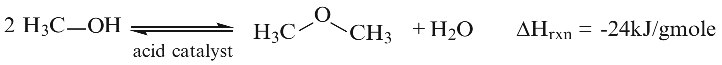
Chemical	Major uses
Polypropylene	31% injection molding; 31% fibers; 18% films/sheets
Acrylonitrile	43% acrylic fibers; 35% ABS/SAN resins; 9% acrylamide; 6% adiponitrile
Propylene oxide	65% polyether polyols for urethanes; 19% propylene glycols; 6% glycol ethers
n-Butyraldehyde	45% *n*-butanol; 47% 2-ethylhexanol
i-Butyraldehyde	40% isobutanol; 37% neopentyl glycol; 15% condensation intermediate
Cumene	≈100% phenol/acetone
Acrylic acid (and esters)	44% acrylate esters; 40% superadsorbent polymers; 7% water treatment
Isopropanol	62% solvent; 13% chemical intermediates; 7% acetone; 7% personal care
Epichlorohydrin	83% epoxy resins; 2% synthetic glycerine

remaining heavy materials in the debutanizer (**11**). A rather complex ethylene/propylene cycle is used to provide refrigeration needs in the plant.

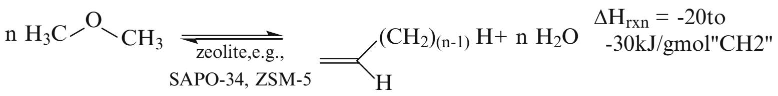
Ethane/ethylene are split in a column comprising over 100 stages (**9**), with ethane underflow recycled to cracking. The propane/propylene separation requires several 100 stages to achieve polymer-grade purity (**12**). Propane underflow is also recycled to cracking. The distillative separation of ethane/ethylene (relative volatility ≈ 1.4) and propane/propylene (relative volatility ≈ 1.2) to produce polymer grade ethylene and propylene (typically 99.5% or greater purity) requires significant staging and reflux. The production of ethylene and propylene are the largest single users of energy in the chemical industries, equating to (not an insignificant) 1% of total energy demand globally.

About a third of propylene is recovered from refinery operations, such as FCC of heavy oils. FCC produces primarily motor gasoline components, but also 5–9% propylene, which can be recovered by fractionation [26]. See Chap. 18. Less than 10% of propylene currently comes from on-purpose propane dehydrogenation, ethylene–butene disproportionation/metathesis, and other sources [27]. During the next decade in the United States in particular, the significant increase in availability of NGL from shale gas is expected to move to higher usage of light feedstocks for cracking. This will exacerbate the relative shortage of propylene and may lead to more serious investment in metathesis and dehydrogenation processes.

It has been known for many years that methanol could be reacted over acidic zeolites at high temperatures to produce primarily an aromatic-heavy gasoline as well as some C_2-C_4 olefins. This type of process was commercialized briefly in New Zealand in the 1980s for gasoline production and later shut down due to poor economics [28]. The catalyst was the ubiquitous ZSM-5 zeolite. In all process variations, methanol is first converted into dimethyl ether with release of water. DME then further generates a "pool" of CH_2 equivalents on the catalyst surface.



$$2\ H_3C{-}OH \underset{\text{acid catalyst}}{\rightleftharpoons} H_3C{-}O{-}CH_3 + H_2O \qquad \Delta H_{rxn} = \text{-24kJ/gmole}$$



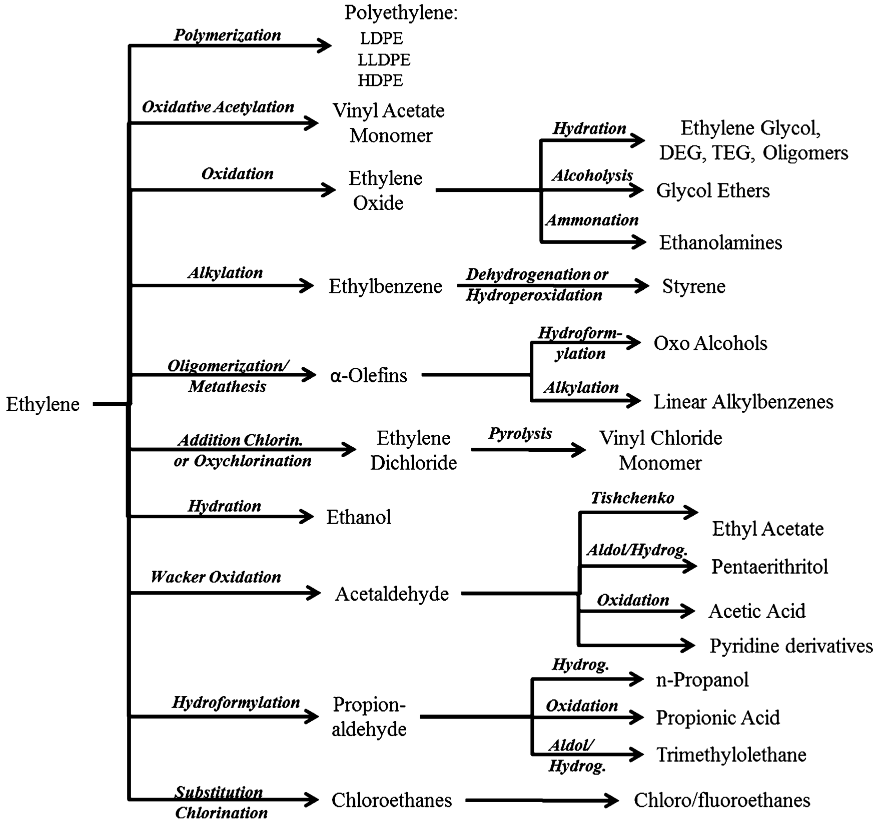


Fig. 10.10 Ethylene derivative tree

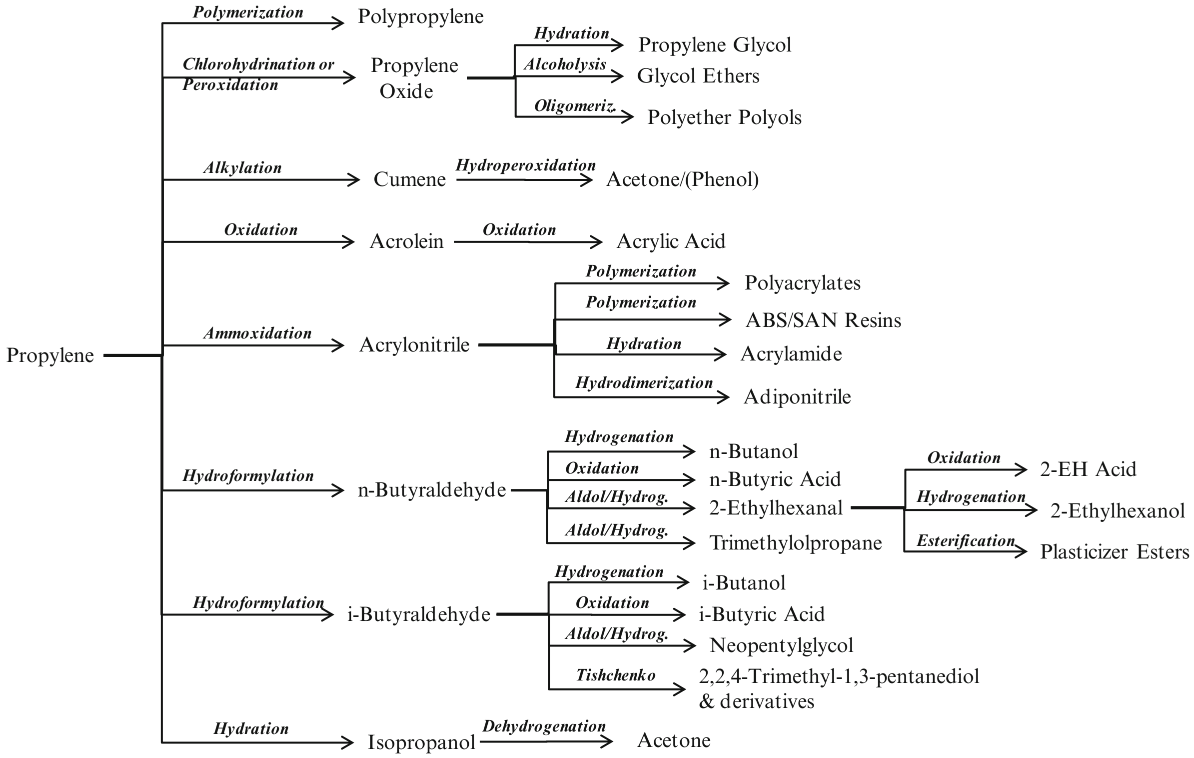


Fig. 10.11 Propylene derivative tree

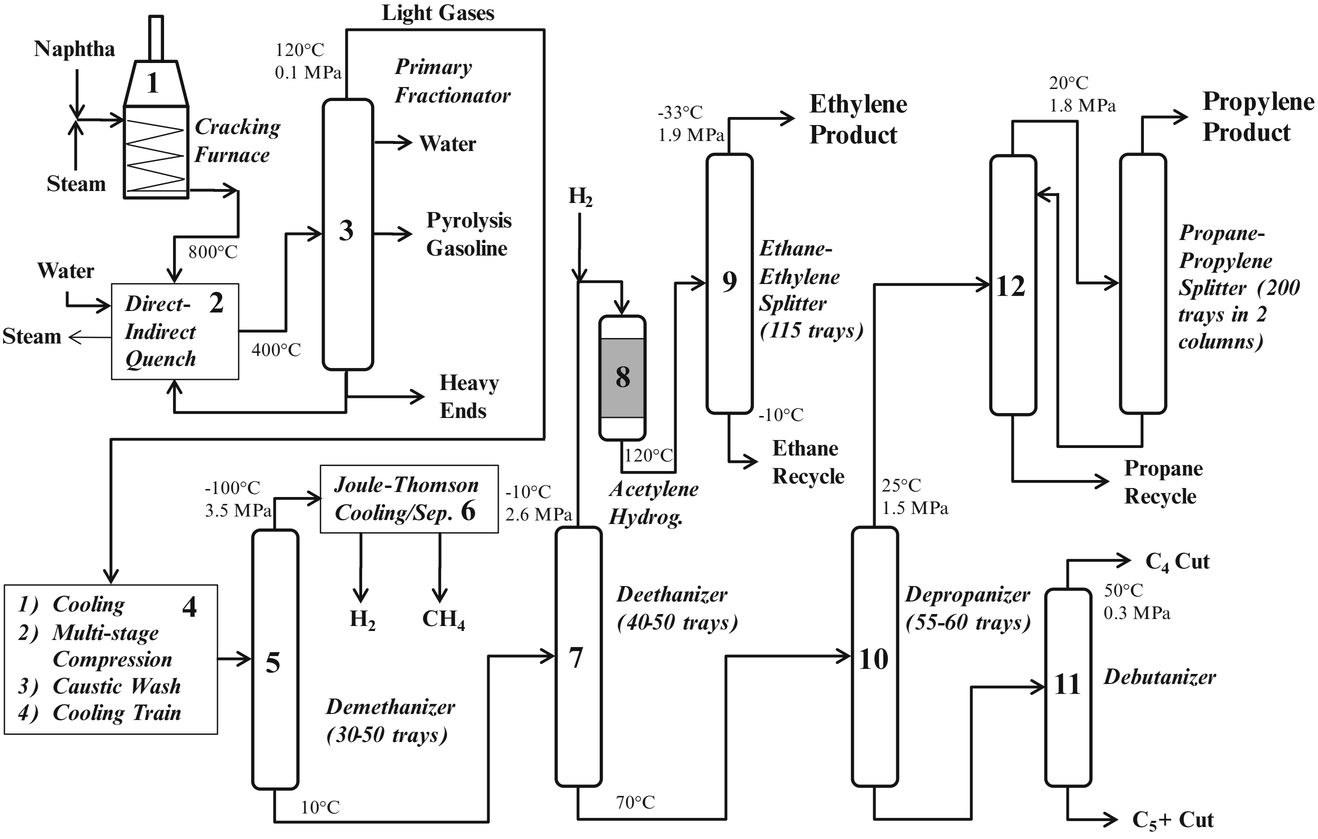


Fig. 10.12 Flowsheet for naphtha cracker

Table 10.9 Typical stream cracking yields

	Ethane (%)	Propane (%)	*n*-Butane (%)	Naphthas (%)	Gas oils (%)
Hydrogen rich gas	5.9	1.7	1.3	0.9	1.3
Methane rich gas	7.1	25.4	22.4	14.6	10.5
Ethylene	80.5	41.2	40.6	31.5	25.9
Propylene	2.0	21.8	16.7	15.3	14.9
Butadiene	2.0	2.3	3.5	4.6	4.5
Butanes/butenes	0.9	1.6	6.7	6.8	4.8
Benzene	0.9	2.0	2.6	4.7	6.2
Toluene	0.2	0.5	1.2	3.2	3.2
Xylenes/C_8 aromatics	0.0	0.0	0.4	2.4	1.4
Acetylenes and others	0.7	2.9	3.0	13.2	7.4
Fuel oil	0.0	0.6	1.7	2.8	19.9
Total	100	100	100	100	100

These CH_2 equivalents then participate in the formation of larger molecules of many possible types, such as *n*-olefins, aromatics, naphthenes, relatively small amounts of *n*-paraffins, CO_2, hydrogen, methane, and trace amounts of lower carboxylic acids [29–31]. The product distribution is heavily dependent on process conditions as well as zeolite geometry. High temperatures favor aromatic formation, while the pore size of the zeolite, if properly chosen, can significantly reduce or virtually eliminate aromatic formation. For example, SAPO-34, a template-based silica aluminophosphate molecular sieve zeolite with a pore opening of 3.8 Å, controls the size of the olefin produced to mostly C_2-C_3, as well as minimizing aromatics formation. ZSM-5 and MFI zeolites, with pore openings of 5.1–5.6 Å, lead to much higher aromatics production.

Modern methanol-to-olefins processes capable of producing relatively high yields of ethylene and propylene, are just beginning to be commercialized, in particular in China, where the methanol is largely derived from coal [32–34]. In the UOP/Hydro Methanol-to-Olefins (MTO) process methanol is reacted at 350–550°C, 0.2–0.4 MPa, over SAPO-34 zeolite to produce predominantly ethylene and propylene. Approximately 80% of the input carbon is converted into ethylene and propylene, and as high as about 90% with recycle of the butanes fraction. The ethylene/propylene product mass ratio can be varied between 0.75 and 1.05, dependent on process conditions. For 100 kg of methanol input, typical mass yields are 17 kg ethylene, 17 kg propylene, 5 kg mixed C_4's, 1.9 kg C_5+, 1.7 kg fuel gases, 57 kg water, 1 kg and minor materials (CO_2, coke, carboxylic acids). This process uses a fluid bed reactor with a separate regenerator vessel (using air), much like an FCC unit. The reactor effluent is separated from the water by-product and unreacted DME is recovered. The hydrocarbon fraction is washed with caustic to remove acids, dried, and compressed. Recovery of pure ethylene and propylene occurs in a distillation train very similar to that used in steam cracking for olefin production.

In the Lurgi MTP® process, methanol is converted to DME over γ-alumina in a separate reactor, followed by zeolite-based conversion to hydrocarbons in a six-stage adiabatic reactor system. Propylene yield is about 70% on input carbon, with relatively minor amounts of ethylene. A gasoline fraction of relatively modest octane is the major by-product.

As of 2011, at least one plant has been built in Brazil for dehydration of sugarcane-derived ethanol to ethylene (200 KMTA of ethylene, <0.2% of world capacity), and additional facilities are planned in Brazil and India [35–37].

C_6-C_8 Aromatics: BTX [38–42]

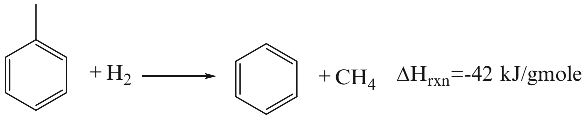
In the early twentieth century, the primary source of the C_6-C_8 aromatics (benzene, toluene, *o*-, *m*-, *p*-xylenes, BTX) was coal tar liquids (the aromatics-rich fraction is also called benzole) from the carbonization of coal to produce metallurgical coke. The rise of the petroleum refining industry has significantly diminished the contribution of this source of aromatics, less than 2% of aromatics are still derived from coal tars. Globally BTX is produced primarily by catalytic reformer operations (55–60%), and by recovery from pyrolysis gasoline (40–45%). Because of the large gasoline demand (as well as the tradtional use of LPG/NGL for steam cracking), about 70–75% of aromatics produced in the United States comes from reformates. The situation is reversed in Europe, Japan, and China. The predominance of naphtha steam cracking in these locations leads to larger amounts of aromatics from pyrolysis gasoline.

In catalytic reformers, dehydroisomerization and cyclization/dehydrogenation reactions convert cycloparaffins and alkylcyclopentanes (naphthenes), along with paraffins into predominately aromatic liquid products at high temperature over platinum-based catalysts (with H_2 and lighter hydrocarbon as by-products). Reforming significantly improves the octane, or gasoline blending quality of the stream. A typical reformate contains 50–60% aromatics, mostly toluene and xylenes. In the United States about 90% of this reformate is used for gasoline production, the other 10% becomes chemical feedstocks.

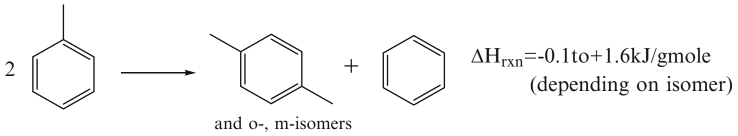
The amount and composition of the BTX fraction recovered from pyrolysis gasoline is highly dependent on the steam cracking feedstock and severity of conditions (see Table 10.9). Due to its high content of polymerizable components (not good for internal combustion engines), virtually all of the pyrolysis gasoline is used for chemical feedstock production rather than gasoline production.

The most important aromatics for chemical production are benzene and *p*-xylene. Reformer and pyrolysis gasoline (pygas) aromatic fractions as produced typically do not meet the demand patterns for these aromatics. Four families of processes are widely used to adjust the natural distribution between benzene, toluene, xylene isomers, and lesser amounts of polyalkylbenzenes and ethylbenzene [43–47]. See also the discussion of alkylation in Section 10.3 below.

1. *Hydrodealkylation* involves the thermal or catalytic reaction of alkylated aromatics (normally methyl aromatics) with hydrogen to produce light alkanes (normally methane) and benzene.



   Toluene is the most common feed substrate, but higher alkylated benzenes (e.g., xylenes and trimethyl benzene) are also used. In the catalytic process, temperatures stay below 500°C to prevent metal sintering, with pressures from 2.5 to 7.0 MPa. Common catalysts are Group VIII metals and metal oxides, e.g., chromium oxides, Pt, Pt oxide, on alumina. Thermal dealkylation requires a temperature of 600–660°C, pressures of 3.5–7.0 MPa. The heat of reaction is controlled by recycling cold hydrogen at high (6:1 to 8:1) molar ratio to substrate. Because of the formation of the highly stable by-product of methane, the reaction goes very far to completion. Selectivity to benzene approaches 99% in the thermal process and 98–99% in the catalytic processes. Many versions of hydrodealkylation have been commercialized and are available for license.
2. *Transalkylation/disporportionation* involves the migration of methyl groups among aromatic rings in the presence of hydrogen. For example two toluene molecules produce a mixture of xylene isomers.

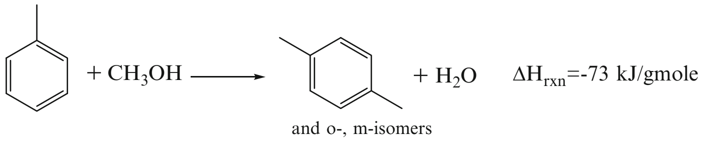


The reaction proceeds over aluminosilicate or silicoaluminophosphate zeolites (often containing noble metal (Pt) or rare earths) at 400–470°C, pressures of 1.4–3 MPa, with a hydrogen:aromatic molar ratio of 5:1 to 12:1. Yield is well above 90%. Under normal reaction conditions the xylene isomer composition approaches the equilibrium value. At 400°C the distribution is about 21.6% *ortho*-, 24.4% *para*-, and 54% *meta*-xylene. Thus, if *p*-xylene is the desired product (usually), conversion to *para*-xylene is low per pass. Separation and recycle are critical (and expensive).

A fairly recent innovation has been the use of shape selective zeolite catalysts that allow toluene into the zeolite cage, but only allow *para*-xylene, with its narrow cross sectional area, to leave. *Ortho*- and *meta*-isomers are trapped and isomerized until *para*-xylene is formed. Significantly greater than equilibrium levels of *para*-xylene can be produced with much less separation and recycle requirements.

The other significant C_8 isomer, ethylbenzene (EB), can be present in appreciable amounts, depending on the origin of the stream. Difficult to separate cleanly (requiring superfractionation with over 200 stages), it can be converted via several reaction pathways: (1) de-ethylation to benzene (similar to HDA); (2) exhaustive hydrogenation to ethylcyclohexane, rearrangement to dimethylcyclohexanes, and dehydrogenation to xylene isomers; (3) disproportionation with xylenes to give benzene and ethyl xylenes. All three are operational on most transalkylation and disproportionation catalysts, typically giving at least 30% up to almost 100% EB conversion per pass.

3. *Isomerization* is a related process using much the same catalysts and reaction conditions as transalkylation. Some versions do not require hydrogen co-feed and can operate at 200–260°C. Shape selective catalysts that enhance *para*-xylene formation are also employed for isomerization.
4. *Toluene methanation* is a fourth process to convert toluene to xylenes that has been commercialized in the last decade to a limited extent. Methanol is reacted with excess toluene at high temperature over acidic or zeolite catalysts to produce xylene isomers and water.



   Toluene is distilled and recycled, while the product mixture of xylene isomers must be separated and isomerized to maximize *p*-xylene. Shape selective, *para*-enhancing zeolites are also used here. Note that this reaction cannot be accomplished with methane as the alkylating agent.

Recovery and purification of individual BTX components are challenging separations problems. Separation of paraffins and olefins from aromatics is typically done by some combination of simple distillation, extraction, and extractive distillation. Benzene and toluene (normal boiling points of 80.1°C and 110.6°C respectively) can be separated from C_8 and higher aromatics by simple distillation. However, the C_8 isomers all boil within about 8°C of each other (ethylbenzene at 136.2°C, *p*-xylene at 138.3°C, *m*-xylene at 139.1°C, and *o*-xylene at 144.4°C), rendering simple distillation impractical. Crystallization and shape-selective molecular sieve adsorption have become standard techniques for xylene isomer purification. A rich literature on these separations is readily available [48–53]. An aromatics complex often has many of these reaction and separation processes working in concert. Further information on aromatics production is given in Chap. 18.

The primary derivatives produced from BTX feedstocks, the capacity/production/consumption of those primary derivatives, and their major applications are summarized in Tables 10.10, 10.11, and 10.12. The family of derivatives produced from BTX is presented in Fig. 10.13.

C_4 Hydrocarbons: Butanes, Butenes, Butadiene [54–58]

The two saturated C_4 hydrocarbons, *n*-butane and isobutane (2-methylpropane), are obtained directly as components of NGL or from a variety of petroleum refining operations. Crude natural gas typically contains 1–2 mole% butanes, with a normal to iso ratio between 1:2 and 2:1. Much of these butanes become raw material for stream cracking to light olefins. Refinery processes giving appreciable amounts of butanes include crude distillation, catalytic cracking, catalytic reforming, delayed coking, fluid coking, visbreaking, and hydrocracking. Further information on these processes is given in Chap. 18. In the United States, refinery operations account for close to 80% of the butanes produced. A significant portion of the refinery butanes is used for fuel, gasoline blending, and other gasoline-producing operations. Only about 10% is diverted as a raw material for chemical synthesis.

Table 10.10 BTX feedstocks: primary derivatives

	BTX derivative	% of world consumption	Chemistry	Additional raw materials	Chapter cross reference
Benzene	Ethylbenzene	52.4	Alkylation	Ethylene	Chapter 10
	Cumene	19.8	Alkylation	Propylene	Chapter 10
	Cyclohexane	12.3	Hydrogenation	Hydrogen	Chapter 10
	Nitrobenzene	8.2	Nitration	Nitric acid	Chapter 22
	Alkylbenzenes	2.7	Alkylation	C_{10}.C_{16} olefins	Chapter 36, Chapter 10
	Chlorobenzenes	0.0	Substitution chlorination	Chlorine	Chapter 10
	Other	4.6			
Toluene	Benzene/xylenes	60.3	Disprop/transalkyl/ hydrodealkyl	Hydrogen	Chapter 10
	Solvent use	12.7			
	Toluene diisocyanate	5.6	Nitration/hydrog/phosgenation	Nitric acid/H_2/phosgene	Chapter 22
	Benzoic acid	1.1	Oxidation	Air	Chapter 10
	Other	20.3			
Xylenes	Terephthalic acid/ DMT	79.1	Oxidation/(esterification)	Air/(methanol)	Chapter 10
	Phthalic acid/ anhydride	10.2	Oxidation	Air	Chapter 10
	Isophthalic acid	2.6	Oxidation	Air	Chapter 10
	Other	8.1			

Table 10.11 BTX feedstocks: capacity/production/consumption of primary derivatives

		Capacity, in kilo metric tons			Production, in kilo metric tons			Consumption, in KMT[a]	
	BTX derivative	World	USA	China	World	USA	China	USA	China
Benzene	Ethylbenzene	35,000	6,100	4,200	28,000	4,800	2,800	4,800	2,800
	Cumene	14,000	4,600	740	12,000	3,700	710	3,500	750
	Cyclohexane	6,800	1,500	650	5,100	1,100	650	1,100	650
	Nitrobenzene	5,700	1,400	1,400	4,500	1,300	730	1,300	720
	Alkylbenzenes	3,400	310	550	2,900	160	510	190	520
	Monochlorobenzene	500 (est)	20	No data	370	20	380	50	270
	Ortho-dichlorobenzene	100 (est)	10	No data	22	7	No data	3	No data
	p-Dichlorobenzene	200 (est)	20	No data	95	18	No data	34	9
	Higher chorobenzenes	25 (est)	No data	No data	6	No data	No data	2	No data
Toluene	Toluene diisocynate	2,200	480	440	1,800	410	270	220	460
	Benzoic acid	360	75	130	250	50	73	45	63
Xylenes	Terephthalic acid/DMT (TPA equivalents)	No data	4,500	No data	40,000	3,300	11,000	3,600	17,000
	Phthalic acid/phthalic anhydride	4,800	470	1,400	3,500	290	900	280	1,100
	Isophthalic acid	1,000	270	35	600	150	30	120	100

[a]Gray shading indicates net importer

Butenes comprise C_4 mono-olefin isomers, 1-butene, cis-2-butene, trans-2-butene, and isobutylene. These isomers are co-produced with butanes in refinery and steam cracking operations. Typical yields of butanes, butenes, and butadiene for several processes are given in Table 10.13. These C_4 fractions cannot be separated economically into pure components by simple distillation due to close boiling points. Rather, separation and reaction methods that exploit property differences due to molecular structure or between functional groups are employed, such as shape-selective adsorption, extraction, extractive distillation, isomerization, and dehydrogenation. Separation sequences typically begin with removal of 1,3-butadiene via extraction or extractive distillation (see below). The remaining butanes/butenes can be treated by a variety of methods including adsorption (to separate branched and normal species); reactions that exploit the reactivity of the olefins, especially isobutylene, over paraffins; extraction or extractive distillation (to separate butenes from butanes), or isomerization (to convert *n*-butenes to isobutylene or visa versa, *n*-butane to isobutane or visa versa, or 2-butene to 1-butene). A large body of information on these technologies is available in the open

Table 10.12 Major uses of BTX primary derivatives

	BTX derivative	Major uses
Benzene	Ethylbenzene	≈100% styrene
	Cumene	≈100% phenol/acetone
	Cyclohexane	>95% KO oil (cyclohexanol-cyclohexanone)—mostly for nylon intermediates
	Nitrobenzene	97% aniline, 3% pharma, dyes, intermediates
	Alkylbenzenes	≈100% linear alkylbenzene sulfonate surfactants
	Monochlorobenzene	79% nitrochlorobenzene production; 7% sulfone polumers; 3% phenyl compounds
	Ortho-dichlorobenzene	87% dichloroaniline; 12% solvent
	p-Dichlorobenzene	57% polyphenylene sulfide polymer; 16% moth control; 14% air deodorants
	Higher chorobenzenes	Solvents and intermediates
Toluene	Toluene diisocynate	90% flexible polyurethane foams; rigid foams, adhesives, paints, concrete sealers, cross-linker [321]
	Benzoic acid	43% benzoate salts; 22% benzoate plasticizers; 11% alkyd resins; 8% feed additives, 4% benzoyl chloride
Xylenes	Terephthalic acid/DMT (TPA equivalents)	60% PET Fibers; 31% Bottles/Packaging; 5% Films; 2% PBT
	Phthalic acid/phthalic anhydride	53% plasticizers; 21% alkyd resins; 16% unsaturated polyester resins
	Isophthalic acid	33% unsat polyesters; 39% co-polyesters; 19% surface coating resins

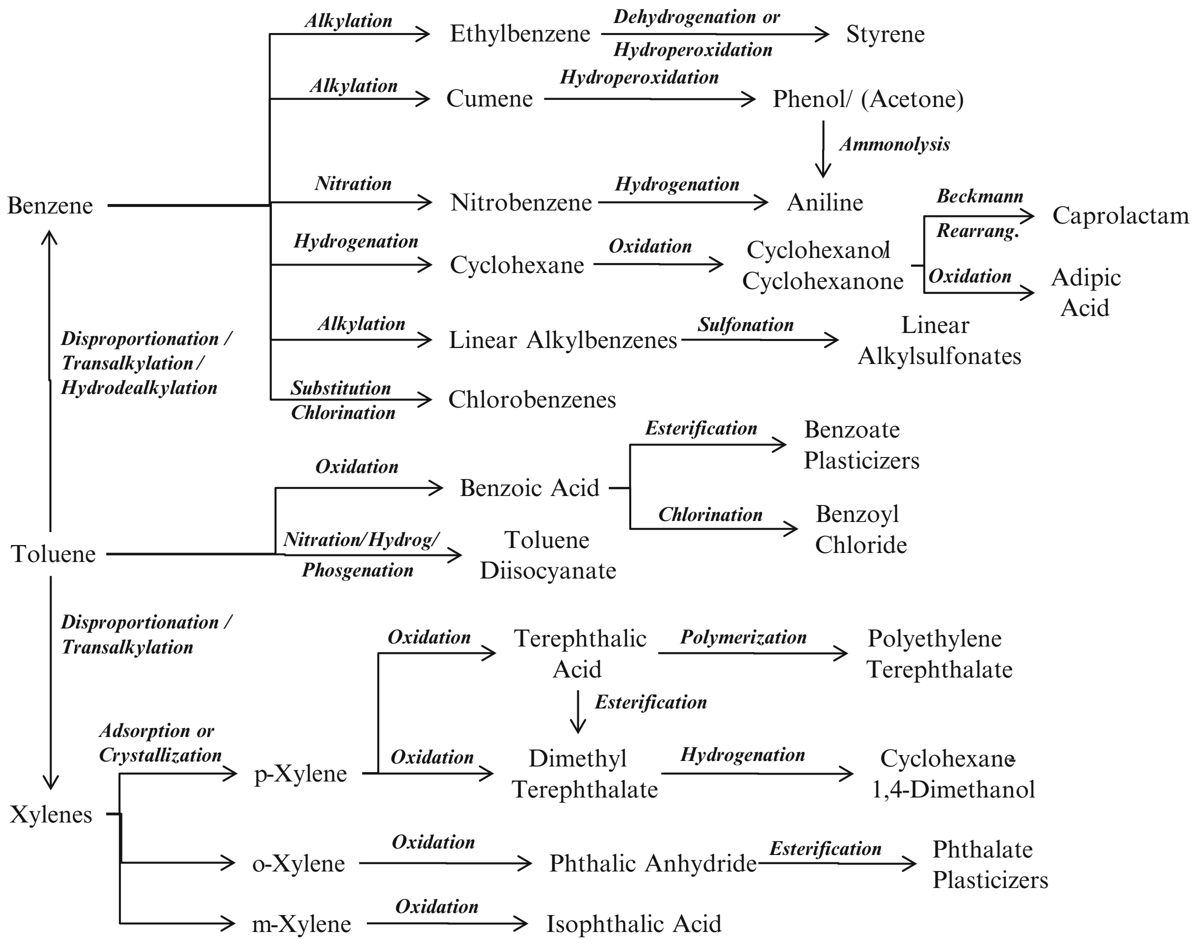


Fig. 10.13 BTX derivative tree

literature [55, 58–60]. Dehydrogenation of C_4 species is covered briefly in the section, “Butenes and 1,3-Butadiene.”

During the 1990s in the United States, a significant fraction of the butenes pool was used for production of ether oxygenates, such as MTBE (methyl tert-butyl ether) for gasoline blending to meet specifications for reformulated gasoline. MTBE use in gasoline was later discontinued and replaced with corn-based ethanol. A large portion of the butenes is now used for alkylation (isobutane reacted with propylene and/or butenes) and for the production of higher octane butene dimer blending components. Further information on these processes is given in Chap. 18.

Over 95% of the 1,3-butadiene (1,3-BD) produced globally is recovered as a by-product of light olefins production from steam cracking. A number of extractive distillation processes, such as the *n*-methylpyrrolidone (NMP)-based process illustrated in Fig. 10.14, have been commercialized for recovery of 1,3-BD from such mixed C_4 streams [61–64]. The mixed C_4 stream, containing *n*/*i*-butanes, *n*/-*i*-butenes, 1,3-BD, 1,2-butadiene, and alkynes, is fed to a predistillation column (**1**), where propyne, propadiene, and other lights are removed overhead. The underflow of this column is fed to the bottom of the stripping section of the extractive distillation column using NMP as the extractive distillation solvent (**2**). The distillate is a butadiene-free butanes/butenes mixture (commonly referred to as raffinate-1). The bottom butadiene-rich stream is sent to the rectifying section of the extractive distillation, a divided wall column (**3**), where butenes are stripped out of the column feed mixture in one compartment (**3a**), with the stripped materials returned to the extractive stripping column. In the second compartment (**3b**), the more soluble C_4 alkynes are removed from the crude 1,3-BD by extraction into NMP. The bottoms of the rectifying section are distilled in the stripping column (**4**) to remove light hydrocarbons from the recycle solvent. Acetylenes banding in this column are removed as a sidedraw, scrubbed of NMP (**6**), and added to the raffinate-1 stream if desired. The stipped hydrocarbons are recompressed (**5**) and recycled to the rectifying section of extractive column. The overhead from the rectifying section (**3b**) is distilled in the final butadiene column (**7**), with water removed overhead, 1,3-BD taken as a liquid sidedraw product, and C_4-C_5 hydrocarbons underflowed. Recovery of 1,3-BD is above 98%.

Table 10.13 Typical C_4 fractions from some cracking operations

	Catalytic cracking of gas oil (%)	Delayed coking (%)	Naphtha steam cracking (%)
C_4 product as % of feed	3–10	1.5–2.0	2–12
n-Butane	7–13	47	2–5
i-Butane	28–52	12	0.5–1.5
1-Butene	7–8	13	14–16
Cis-2-butene	20–30	5	5–6
Trans-2-butene	20–30	7	6–7
Isobutylene	26–28	16	22–27
1,3-Butadiene	0.1–0.5	0.5	37–45
1,2-Butadiene, C_3-C_4 acetylenes	1–2	<3	3–8

A number of older processes for on-purpose synthesis of 1,3-BD, such aldol condensation of acetaldehyde, followed by hydrogenation/dehydration, ethanol to acetaldehyde to butadiene(via aldol), or acetylene plus formaldehyde [65], are largely obsolete. A small amount of butadiene is produced by dehydrogenation or oxidative dehydrogenation of butane/butenes, as discussed in section, "Butenes and 1,3-Butadiene."

The primary derivatives produced from C_4 hydrocarbon feedstocks, the capacity/production/consumption of those primary derivatives, and their major applications are summarized in Tables 10.14, 10.15, and 10.16. The family of derivatives produced from C_4 hydrocarbon feedstocks is presented in Fig. 10.15.

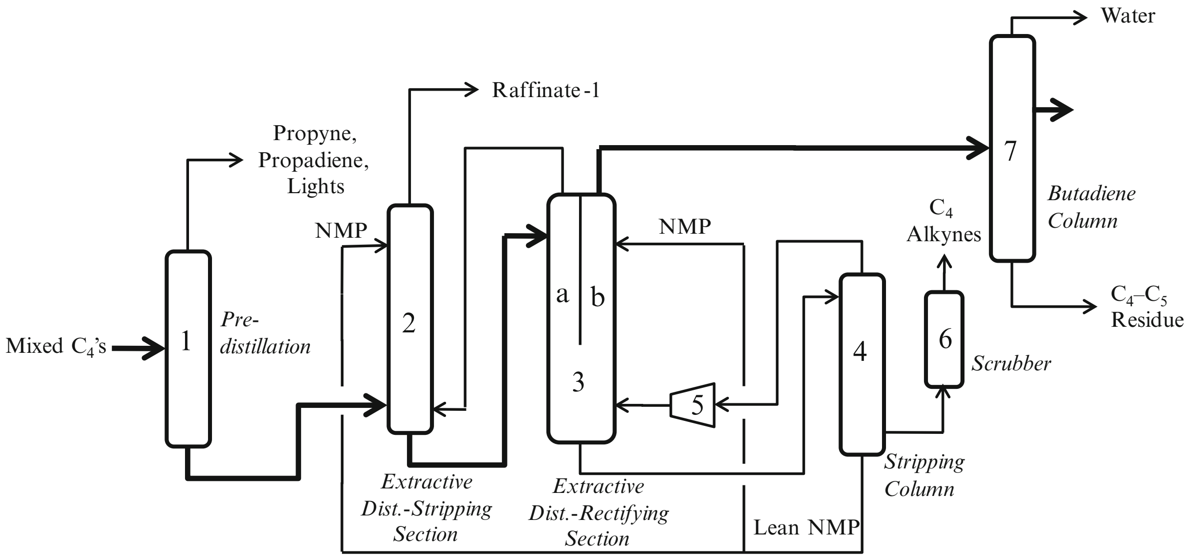


Fig. 10.14 Recovery of 1,3-butadiene from mixed C_4's by extractive distillation

Table 10.14 C_4 feedstocks: primary derivatives

	Butenes derivative	% of world consumption	Chemistry	Additional raw materials	Chapter cross reference
Butanes	Gasoline/fuel	83.3	Alkylation and direct use	Other fuel components	Chapter 18
	Steam cracker feed	11.8	Thermal cracking	NGL/LPG, naphtha	Chapter 10
	MTBE	2.3	Etherification	Methanol	Chapter 18
	Propylene oxide/ tert-butanol	1.9	Hydroperoxidation	Propylene	Chapter 10
	Maleic anhydride	0.4	Oxidation	Air	Chapter 10
	Acetic acid	0.0	Oxidation (no longer used)	Air	Chapter 10
	Other	0.3			
1,3-Butadiene	Elastomeric rubbers	65.2	Polymerization	Styrene	Chapter 16
	ABS resins	12.4	Polymerization	Acrylonitrile/styrene	Chapter 15
	Styrene-BD latexes	12.4	Polymerization	Styrene	Chapter 16
	Adiponitrile/HMDA	4.3	Hydrocyanation/ hydrogenation	HCN/H2	Chapter 22
	Other	5.7			
Mixed butenes	Mixed butenes fuel uses	82.0	Alkylation, etherification, others	Methanol, other fuel components	Chapter 18
n-Butenes	Co-monomer, polyethylene	2.2	Polymerization	Ethylene	Chapter 15
	sec-Butanol/MEK	1.9	Hydration/ dehydrogenation	Water	Chapter 10
	1,3-Butadiene	0.5	Dehydrogenation		Chapter 10
	n/*i*-Valeraldehydes	0.3	Hydroformylation	1-1 CO/H_2	Chapter 10
	Other	3.1			
Isobutenes	Butyl rubber	1.7	Polymerization	Isoprene	Chapter 16
	Polybutenes	1.6	Polymerization	306834	Chapter 15
	Neopentanoic acid	(Incl in other)	Koch carbonylation	CO, water	Chapter 10
	t-Butanol	(Incl in other)	Hydration	Water	Chapter 10
	Methyl methyacrylate	0.8	Oxidation, esterification	Air, methanol	Chapter 22, Chapter 10
	Other	1.8			

Table 10.15 C_4 Feedstocks: capacity/production/consumption of primary derivatives

	Capacity, in kilo metric tons			Production, in kilo metric tons			Consumption, in KMT[a]	
Chemical	World	USA	China	World	USA	China	USA	China
Maleic anhydride	2,100	280	720	1,400	230	380	220	310
MEK	1,400	140	430	1,000	100	240	150	270
Valeradehydes	450	190	0	260	160	0	160	0
Methyl methacrylate[b]	3,600	940	400	2,800	770	330	600	360

[a] Gray shading indicates net importer

[b] Values for all routes to methyl methacrylate

Table 10.16 Major uses of C_4 hydrocarbons primary derivatives

Chemical	Major uses
Maleic anhydride	55% unsaturated polyester resins; 13% 1,4-BDO; 4% fumaric acid; 3% lubes
MEK	49% coatings solvent; 17% adhesives; 10% inks
Valeradehydes	C_5 alcohols, acids; aldol chemistry
Methyl methacrylate	59% PMMA resins/sheets; 21% surface coatings

C_9-C_{17} Paraffins [66–68]

Higher C_9-C_{17} alkanes derived from petroleum, find limited but important uses as raw materials for the production of higher C_9-C_{17} *n*-olefins. These higher *n*-olefins are intermediates in the manufacture of detergents and anionic surfactants. Because of their wide availability and generally low cost, heavy naphthas (boiling range of 150–200°C) and

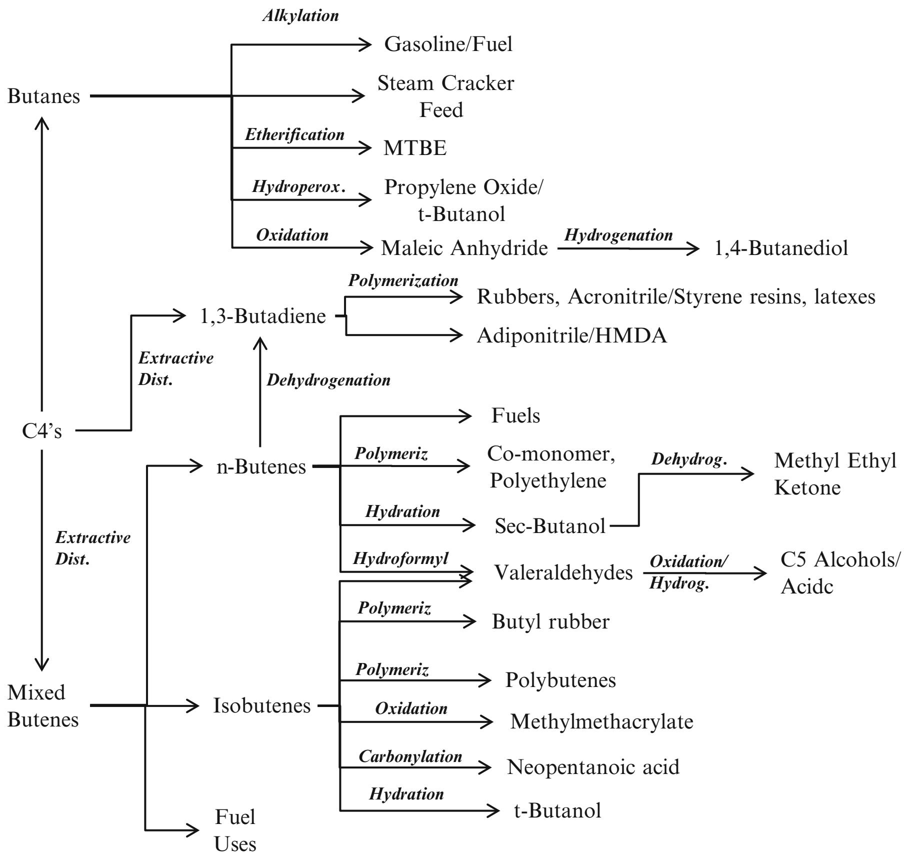


Fig. 10.15 C_4's derivative tree

Table 10.17 C_9-C_{17} hydrocarbon feedstocks: primary derivatives

C_9-C_{17} *n*-paraffinic derivatives	% of world consumption	Chemistry	Additional raw materials	Chapter cross reference
Linear alkylbenzenes	79.0	Dehydrogenation/alkylation	Benzene	Chapter 10
Detergent alcohols	5.1	Dehydrogen/hydroformyl/ hydrog	Syngas/H_2	Chapter 36, Chapter 10
Chloroparaffins	5.4	Dehydrog/addition chlorination	Cl_2	Chapter 10
Others—solvents, etc.	10.5			

kerosene (boiling range of 190–250°C) are the raw materials for production these olefins. The naphtha and kerosene fractions typically contain 20% or more *n*-paraffin content. After hydrotreating to remove sulfur which poisons downstream catalysts, the naphtha and kerosene cuts are separated by shape-selective adsorption that takes advantage of the larger cross-sectional area of the branched isomers. The *n*-alkanes are preferentially adsorbed, while the branched materials largely pass through. The *n*-alkanes are desorbed by a light hydrocarbon, and distilled to recover the desorbent. Both vapor and liquid phase adsorption processes are practiced. The high *n*-alkane product is then dehydrogenated to give most internal olefins. Further details on the dehydrogenation process are given in the section, "Internal Olefins from Higher *n*-Alkanes." The primary derivatives produced from C_9-C_{17} hydrocarbon feedstocks, the capacity/production/consumption of those primary derivatives, and their major applications are summarized in Tables 10.17, 10.18, and 10.19. The family of derivatives produced from C_9-C_{17} hydrocarbon feedstocks is presented in Fig. 10.16.

Synthesis Gas [69–77]

Synthesis gas, or "syngas," is a mixture of hydrogen and carbon monoxide produced by the partial oxidation of carbonaceous feedstocks. The ratio of hydrogen to carbon monoxide in the crude syngas (as well as carbon dioxide content and impurities) is highly dependent on the carbonaceous raw material as well as the process used to generate the syngas. Natural gas is the most important raw material for syngas generation worldwide, with limited coal-based production in

Table 10.18 C_9-C_{17} hydrocarbons feedstocks: capacity/production/consumption of primary derivatives[a]

	Capacity, in kilo metric tons			Production, in kilo metric tons			Consumption, in KMT[b]	
Chemical	World	USA	China	World	USA	China	USA	China
Linear alkylbenzenes	3,000 (est)	420	420	2,300	240	370	240	440
Detergent alcohols	250 (est)	0	0	150	0	0	0	20
Chloroparaffins	200 (est)	No data	0	160	45	0	30	70

[a] Figures are for production using C_9-C_{17}-derived olefins. Ethylene oligomers also used for production of these derivatives

[b] Gray shading indicates net importer

Table 10.19 Major uses of C_9-C_{17} hydrocarbons primary derivatives

Chemical	Major uses
Linear alkylbenzenes	≈100% surfactants
Detergent alcohols	42% alcohol ethoxylates; 27% alcohol ether sulfates; 13% alcohol sulfates
Chloroparaffins	Intermediate for linear alkylbenzenes, solvents

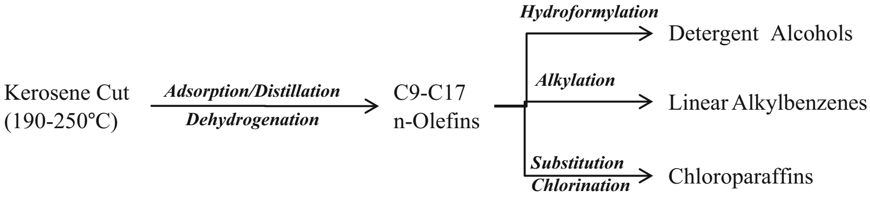


Fig. 10.16 C_9-C_{17} normal paraffins derivative tree

the United States, South Africa, India, and significantly growing production in China. Liquid hydrocarbons (naphtha, heavy fuel oils, and vacuum resid) are also used to some extent.

Syngas production involves the breaking of C–C and C–H bonds of the raw material molecules at high temperature via reaction with water (steam reforming), oxygen (partial oxidation), or carbon dioxide (carbon dioxide reforming), or combinations therein (autothermal reforming or gasification). The primary reactions in steam methane reforming (SMR), stream hydrocarbon reforming (SHR), and carbon dioxide reforming (CMR—not yet commercial) are endothermic, with only the water-gas shift reaction providing heat. Heat must be added externally, normally by superheating steam and raw materials in a fired furnace.

$$CH_4 + H_2O \leftrightarrow CO + 3H_2$$
$$\Delta H_r = +206 \text{ kJ/gmole(methane steam reforming)}$$

$$C_nH_m + nH_2O \rightarrow nCO + (n + 1/2m)H_2$$
$$\Delta H_r \approx +170 \text{ kJ/gmole - } CH_2 - \text{(hydrocarbon reform)}$$

$$CH_4 + CO_2 \leftrightarrow 2CO + 2H_2 \quad \Delta H_r = +247 \text{ kJ/gmole}$$
$$\text{(methane } CO_2 \text{ reforming)}$$

$$CO + H_2O \leftrightarrow CO_2 + H_2$$
$$\Delta H_r = -41 \text{ kJ/gmole(water - gas shift)}$$

Syngas can also be produced by partial oxidation, typically with high purity oxygen (to avoid dilution with large quantities of hard to separate nitrogen). In partial oxidation, a fraction of the raw material is simultaneously combusted to completion to provide in situ heat input. Roughly 70–85% of the energy in the raw material is preserved in the syngas produced. For methane, pertinent reactions are given below, with reforming and water-gas shift reactions also occurring to some extent.

$$CH_4 + \frac{1}{2}O_2 \rightarrow CO + 2H_2$$
$$\Delta H_r = -36 \text{ kJ/gmole(partial oxidation)}$$

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$$
$$\Delta H_r = -803 \text{ kJ/gmole(combustion)}$$

$$C_nH_m + (n/2)O_2 \rightarrow nCO + (m/2)H_2$$
$$\Delta H_r \approx -80 \text{ kJ/gmole - } CH_2 - \text{(hydrocarbon part. ox.)}$$

$$C_nH_m + (n + m/4)O_2 \rightarrow nCO_2 + (m/2)H_2O$$
$$\Delta H_r = -650 \text{ kJ/gmole(combustion)}$$

$$CO + \frac{1}{2}O_2 \rightarrow CO_2 \quad \Delta H_r = -283 \text{ kJ/gmole(combustion)}$$

$$H_2 + \frac{1}{2}O_2 \rightarrow H_2O \quad \Delta H_r = -242 \text{ kJ/gmole(combustion)}$$

Table 10.20 H_2/CO ratios of syngas processes

Syngas Production Method	Feedstock	Feedstock C/H wt ratio	Product H_2/CO molar ratio	Stoicheometric number
Syngas production method	Natural gas	3.2	2.8–4.7 (avg 4.3)	2.1 avg
	Naphtha	8.4	2.1–4.4	2.0 avg
Autothermal reforming	Natural gas	3.2	1.8–3.8	2.0 avg
	Naphtha	8.4	1.1–3.4	2.0 avg
POX	Natural gas	3.2	1.75	1.6
	Naphtha	8.4	1.1	1.0
	Heavy fuel oil	5.2	0.95	0.8
Gasification—carbonaceous solids				
Entrained flow—O_2 blown, water slurry (general electric)	Coal—Pittsburgh #8	7.3	0.89	0.34
	Coal—Illinois #6	7.7	0.86	0.30
	Petcoke	11.6	0.62	0.24
Entrained flow—O_2 blown, dry feed (shell)	Coal—appalachian	8.1	0.47	0.45
	Coal—Illinois #6	6.4	0.47	0.42
	Lignite	6.2	0.49	0.38
Biomass gasifier (silva gas—POX)	Wood	8.6	0.55	0.17
Flash pyrolysis	Wood	8.6	0.48	−0.3
Underground gasification	Coal—Illinois #6	7.4	2.02	0.26
	Coal—powder river basin	6.9	2.53	−0.15

Gasification of solid carbonaceous raw materials adds additional endothermic reactions of solid carbon with steam and exothermic reactions with oxygen. Methanation may also occur under appropriate conditions.

$$C + CO_2 \leftrightarrow 2CO$$
$$\Delta H_r = +172\ \text{kJ/gmole(Boudouard reaction)}$$

$$C + H_2O \leftrightarrow CO + H_2$$
$$\Delta H_r = +131\ \text{kJ/gmole(water gas reaction)}$$

$$C + \frac{1}{2}O_2 \rightarrow CO$$
$$\Delta H_r = -111\ \text{kJ/gmole(partial oxidation)}$$

$$C + 2H_2 \leftrightarrow CH_4$$
$$\Delta H_r = -75\ \text{kJ/gmole(methanation)}$$

Typical H_2/CO ratios for the most important syngas-generating processes are summarized in Table 10.20. Further information on syngas production from natural gas and via coal gasification is given in Chaps. 19, 20, and 22.

It is not much of an exaggeration to say that coal contains every element on the periodic table, many of which end up in the crude syngas, such as mercury, arsenic, and sulfur as H_2S and COS. To a lesser extent natural gas often contains H_2S, COS, as well as mercaptans. Many of the downstream catalysts used in syngas derivative processes are particularly susceptible to poisoning by these trace elements. A wide variety of techniques, mostly involving absorption and adsorption, have been developed and implemented for removal of such trace elements.

Most syngas derivatives are produced optimally at a specific H_2/CO ratio, with a limited acceptable CO_2 content [78]. See Fig. 10.17 for the syngas family of derivatives produced from different H_2/CO ratios. If not produced directly from the syngas generator, a number of reaction and separation techniques are used to adjust the composition of the gas.

1. The water-gas shift reaction is the primary mechanism for reactive adjustments of the H_2/CO ratio. Both sweet (very low amounts of H_2S and other sulfur moieties present) and sour (high sulfur levels) shifts are practiced. The feed gas enters the shift reactor at 200°C or higher.

$$CO + H_2O \leftrightarrow CO_2 + H_2$$
$$\Delta H_r = -42\ \text{kJ/gmole(water gas reaction)}$$

   The reaction is normally run adiabatically with a significant temperature, often of 200°C or more. As temperature is increased, the equilibrium shifts more toward the left. Multiple stages, with heat interstage steam generation for cooling, is required for high levels of conversion, i.e., for hydrogen production.
2. Carbon dioxide is removed by absorption with either physical or chemical absorbents, the choice depending on gas pressure, CO_2 concentrations, and desired level of removal.
3. Small adjustments of H_2/CO may be accomplished by membrane permeation (hydrogen permeates primarily), decreasing the H_2/CO ratio of the retentate. Membrane permeation is particularly common for production of 1/1 H_2/CO syngas for hydroformylation.

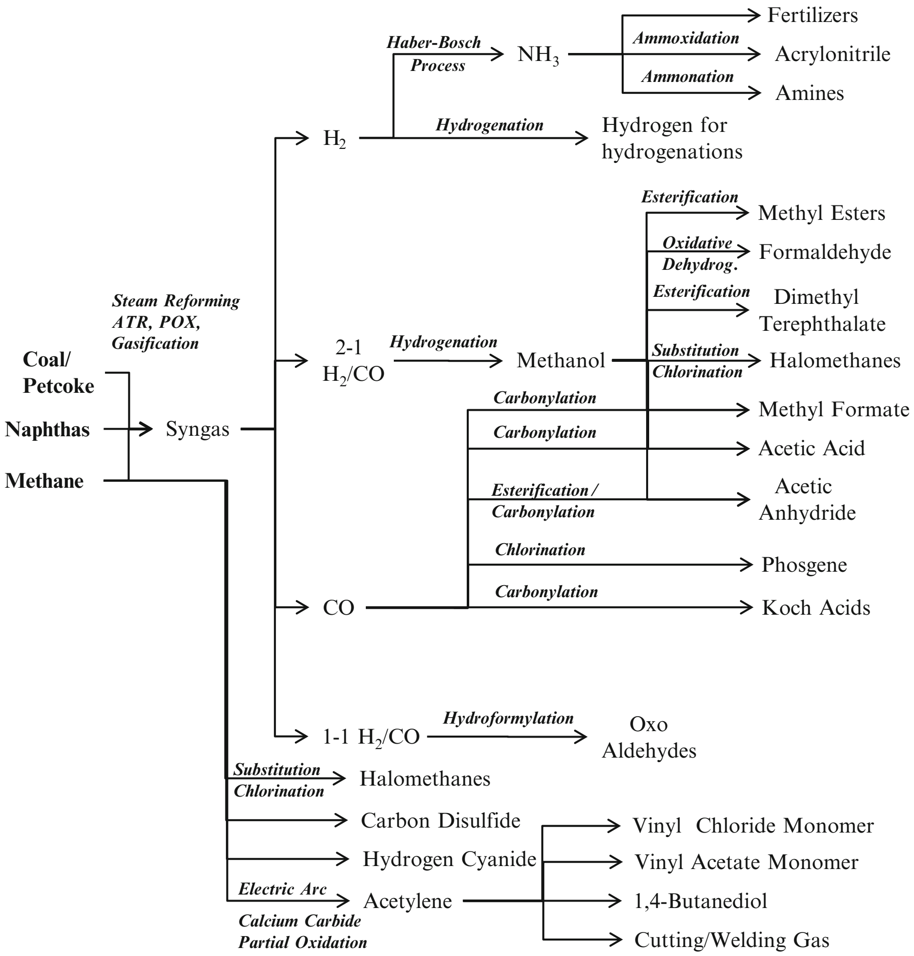


Fig. 10.17 Syngas and methane derivative tree

Table 10.21 Breakdown of syngas types (see also Chaps. 19, 20, and 22)

Type of syngas	Estimated % of worldwide production
Ammonia syngas	65
Hydrogen	11
2-1 H_2/CO (methanol syngas)[a]	21
1-1 H_2/CO (oxo syngas)	2
Carbon monoxide	2

[a] Does not include 2-1 H_2/CO syngas used for Fischer-Tropsch synthesis of fuels

4. High purity hydrogen is produced by water-gas shift, followed by CO_2 absorption, and pressure swing adsorption.
5. CO may be produced at up to about 95% purity by membrane permeation (more extreme case of 3). High purity CO (>99.9%) is made by cryogenic distillation of syngas, or less commonly by chemical absorption processes (CO absorbed).

Estimated global production of syngas, at common H_2/CO ratios, is given in Table 10.21. The capacity/production/consumption of primary derivatives produced from syngas feedstocks and their major applications are summarized in Tables 10.22 and 10.23.

Minor Feedstocks: Methane and Acetylene

Methane itself is relatively unreactive. Much effort has been expended to synthesize functionalized chemicals from methane directly. Examples of large-scale industrial syntheses are carbon disulfide, hydrogen cyanide, and to some extent chloromethanes. However, most other direct methane processes have been ultimately uneconomical and have fallen by the wayside or never been commercialized.

Until the rise of the massive global petrochemical infrastructure in the late 1940s and 1950s, acetylene was one of the most important feedstocks for the production synthetic organic chemicals. Although highly reactive, high conversion and generally very good selectivities can be achieved in acetylene-based chemistries, Such routes thus were favored for many years. Acetylene was once a major feedstock for the production of vinyl chloride monomer (VCM), acrylic acid, acrylonitrile, and many chlorinated solvents. Olefin-based processes have largely supplanted those starting with acetylene. Between 1965 (near its peak demand) and 2007, acetylene usage for

Table 10.22 Syngas feedstocks: capacity/production/consumption of primary derivatives

		Capacity, in kilo metric tons			Production, in kilo metric tons			Consumption, in KMT[a]	
Chemical	Year	World	USA	China	World	USA	China	USA	China
Methanol	2008	64,000	1,000	24,000	42,000	900	11,500	6,100	13,000
Acetic acid	2009	13,000	2,500	5,700	8,000	2,100	2,200	2,100	3,200
Acetic anhydride	2009	2,600	1,000	680	1,600	700	290	670	300
Ammonia	2009	200,000	11,000	67,000	160,000	9,900	55,000	17,000	55,000
Formic acid	2009	720	0	300	570	0	270	10	170
Hydrogen[b]	2009	No data	No data	No data	1,900	No data	No data	No data	No data

[a] Gray shading indicates net importer

[b] Figures include only on purpose hydrogen for non-refinery consumption. Does not reflect hydrogen produced in MeOH, ammonia, CO, oxo syngas synthesis

Table 10.23 Major uses of syngas primary derivatives

Chemical	Major uses
Methanol	35% formaldehyde; 13% MTBE; 10% acetic acid/anhydride; 7% fuel; 3% methyl methacrylate; 3% solvent
Acetic acid	33% VAM; 18% acetic anhydride; 17% acetate esters; 17% solvent for *p*-xylene oxidation to TPA
Acetic anhydride	64% cellulose esters; 12% pharmaceutical synthesis; 6% tetraethylenediamine
Ammonia	54% urea; 7% nitric acid; 7% ammonium nitrate; 6% ammonium phosphate, 3% caprolactam; 1.5% acrylonitrile
Formic acid	30% silage/preservative; 23% leather processing; 14% pharmaceuticals; 9% textile dyeing/finishing
Hydrogen	≈100% reductions

Table 10.24 Acetylene capacity/consumption statistics

In kilo metric tons/yr	World	USA	China
Capacity	3,600	150	3,100
Consumption	3,500	130	3,100

Table 10.25 Large-scale acetylene uses

		% of derivative production based on acetylene		
	% of acetylene consumption	World	USA	China
Vinyl chloride	59	14	0	57
Vinyl acetate	9	20	0	79
1,4-Butanediol	4	40	<50	0
Cutting/welding	22	n/a	n/a	n/a
Other	6			

chemical synthesis has dropped at least 70% in all areas of the world except China. Acetylene demand, especially for VCM, has enjoyed somewhat of a renaissance in China, with capacity growth above 15% per year over the last decade. The fact that acetylene can be made from a wide variety of feedstocks, including coal, is a particular advantage in oil-poor China. Currently about 75% of global acetylene feedstock usage occurs in China. Capacity/consumption statistics for acetylene are given in Table 10.24 for the year 2007. Remaining large scale applications of acetylene are summarized in Table 10.25.

Acetylene can be produced by a wide variety of processes [79, 80]. These include:

1. *Calcium carbide processes*, in which calcium oxide and coke are electrothermally reacted at 2,250°C to produce calcium carbide. The reaction is highly endothermic at +465 kJ/gmole. Calcium carbide is reacted with water in an exothermic reaction acetylene and calcium hydroxide. Yield of acetylene is about 76% on calcium carbide.
2. *Electric arc processes*, in which a carbonaceous raw material is cracked at very high temperatures up to 20,000°C in an electric arc-generated plasma with very short residence times of 0.001–0.01 s. Rapid quenching to 200–300°C freezes the composition and prevents decomposition of acetylene to thermodynamically favorable soot and hydrogen. Typical yields are around 45 wt% of feed carbon to acetylene and 25% to ethylene.
3. *Autothermal partial oxidation processes*, in which light hydrocarbons (predominantly methane in practice) are reacted with oxygen in a controlled fashion. In one version methane and oxygen at a 2:1 M ratio are preheated separately to 500–600°C, mixed, and reacted in proprietary burner configurations at 1,500°C, millisecond residence times to produce syngas and acetylene. The endothermic acetylene formation reaction is coupled with the highly exothermic partial combustion of methane to provide in situ heat integration:

$$CaC_2 + 2H_2O \longrightarrow HC{\equiv}CH + Ca(OH)_2 \qquad \Delta H_{rxn} = -130 kJ/gmole$$

Table 10.26 Raw materials and chemistries of downstream derivatives

Chemical	Raw materials	Chemistry
Vinyl chloride monomer	Ethylene dichloride	Pyrolysis
Ethyl acetate	Ethanol, acetic acid	Esterification; Tishchenko
Ethylene glycol	Ethylene oxide; ethylene carbonate	Hydration; hydration
Ethylene oxide glcyol ethers	Ethylene oxide, alcohols	Alcoholysis
Chloroethanes	EDC, Cl_2	Substitution chlorination
Phenol	Cumene, air	Hydroperoxidation
Acetone	Cumene, air; isopropanol	Hydroperoxidation; dehydrogenation
Hydroquinone	Diisopropylbenzene, air	Hydroperoxidation
MIBK	Acetone, H_2	Aldol, hydrogenation
2-Ethylhexanol	*n*-Butyraldehyde, H_2	Aldol, hydrogenation
Propylene glycol	Propylene oxide	Hydration
Propylene oxide glycol ethers	Propylene oxide, alcohols	Alcoholysis
n-Butanol	*n*-Butyraldehyde, H_2	Hydrogenation
Isobutanol	*i*-Butyraldehyde, H_2	Hydrogenation
Bisphenol A	Phenol, acetone	Condensation
Styrene	Ethylbenzene; propylene, ethylbenzene	Dehydrogenation; hydroperoxid.
Cyclohexanone-cyclohexanol	Cyclohexane	Oxidation
Formaldehyde	Methanol, air	Oxidation
Neopentyl polyols—NPG, PE,TMP, TME	Formaldehyde, C_2-C_4 aldehydes	Aldol, hydrogenation
Ketene	Acetic acid	Thermal dehydration
Ketene derivatives, diketene	Ketene, many others	Many
Methyl chloride	Methanol/HCl; methane/Cl_2	Hydrochlorination; chlorination
Methylene chloride	Methyl chloride	Substitution chlorination
Chloroform	Methylene chloride	Substitution chlorination
Carbon tetrachloride	Chloroform; higher polychlorinated species	Substitution chlorination; perchlorination
Acetate esters (methyl, ethyl, *i*/*n*-propyl, *i*/*n*-butyl)	Acetic acid; C_1-C_4 alkanols	Esterification
1,4-Butanediol	Formaldehyde/acetylene/H_2; maleic anhydride/H_2	Condensation/hydrog; hydrog

$$2\,CH_4 \xrightarrow[1500°C]{\Delta} HC\equiv CH + 3H_2 \qquad \Delta H_{rxn}=+377\,kJ/gmole$$

$$CH_4 + 1.5\,O_2 \longrightarrow CO + 2\,H_2O \qquad \Delta H_{rxn}=-520\,kJ/gmole$$

The reaction gases are rapidly quenched with water or oil to prevent decomposition to soot and hydrogen. The cooled gas, comprising 8% acetylene, 57% H_2, and 26% CO, is contacted with *N*-methylpyrrolidone (NMP) in a countercurrent absorber to recover acetylene (see the section, “C_4 Hydrocarbons: Butanes, Butenes, Butadiene,” for a similar use of NMP). About 25–30% of the carbon in the raw material is converted into acetylene, with about 3–5% going to coke. Most of the rest ends up as CO.

Downstream Derivatives

Many of the primary derivatives of the major petrochemical feedstocks are intermediates for the synthesis of additional downstream derivatives The raw materials and chemistries of these downstream derivatives are summarized in Table 10.26. Capacity/production/consumption figures and major uses of the downstream derivatives are presented in Tables 10.27 and 10.28, respectively.

Manipulation of Hydrocarbons: Oligomerization, Metathesis, Alkylation

Oligomerization

Linear α-Olefins [81–84]

Linear α-olefins can be produced from higher *n*-paraffins by dehydrogenation, as described in the section, “Internal Olefins from Higher *n*-Alkanes.” Linear terminal olefins with an even number of carbon atoms can be produced by a family of related processes based on oligomerization/metathesis of ethylene using trialkyl aluminum or titanium (Ziegler) catalysts or liganded Ni catalysts. The reaction takes place in two steps:

Table 10.27 Capacity/production/consumption of downstream derivatives

	Capacity, in kilo metric tons			Production, in kilo metric tons			Consumption, in KMT[a]	
Chemical	World	USA	China	World	USA	China	USA	China
Vinyl chloride monomer	45,000	9,200	14,000	33,000	6,800	8,300	6,100	9,100
Ethyl acetate	3,000	95	1,200	1,800	70	820	95	590
Ethylene glycol	24,000	2,900	3,000	17,000	1,500	1,900	1,800	7,300
Ethylene oxide glcyol ethers	No data	No data	No data	1,100	440	100	260	220
Chloroethanes	1,200	450	95	870	350	45	370	120
Phenol	10,000	2,900	990	8,600	2,600	560	2,000	1,000
Acetone	6,400	1,800	350	5,700	1,600	340	1,400	820
Hydroquinone	60	25	10	60	20	9	15	10
MIBK	460	90	60	290	60	10	50	60
2-Ethylhexanol	3,100	400	510	2,800	310	510	270	780
Propylene glycol	2,100	760	190	1,600	610	120	460	170
Propylene oxide glycol ethers	No data	No data	No data	520	140	90	110	95
n-Butanol	3,800	1,100	480	2,800	850	470	660	710
Isobutanol	520	110	60	500	100	40	40	70
Bisphenol A	5,400	1,000	510	4,000	820	150	760	660
Styrene	30,000	5,000	3,400	24,000	4,000	2,400	3,500	5,200
Cyclohexanone-cyclohexanol	6,700	1,600	1,500	5,800	1,400	1,100	1,400	1,200
Formaldehyde	42,000	5,300	13,000	29,000	3,700	9,000	3,700	9.000
Neopentyl polyols—NPG, PE,TMP, TME	1,400	220	350	1,000	160	230	160	220
Ketene	62,016	210	130	510	190	130	190	130
Ketene derivatives, diketene	100	20	40	80	15	17	15	17
Methyl chloride	2,600	660	540	1,900	650	340	620	340
Methylene chloride	1,300	300	390	830	120	290	70	280
Chloroform	1,600	270	540	1,200	370	410	220	570
Carbon tetrachloride	220	45	70	130	40	25	30	25
Acetate esters (methyl, ethyl, *i*/*n*-propyl, *i*/*n*-butyl)	5,500	860	2,000	3,500	570	1,400	520	1,100
1,4-Butanediol	1,800	360	350	1,300	280	190	290	260

[a] Gray shading indicates net importer

1. In the oligomerization, or chain growth step, ethylene is inserted progressively into the carbon-catalyst metal bond, leading to a longer linear alkyl chain.
2. To terminate the chain growth, ethylene displaces the longer chain alkyl unit from the C-metal bond, releasing a longer chain α-olefin, and starting the oligomerization cycle over again.

The product composition follows the Shulz–Flory distribution of chain lengths, with substantial amounts of undesirable high and low molecular weight tails. Both one-reactor and two-reactor processes are practiced commercially. The one-reactor process leads to a somewhat broader distribution of molecular weights, with typical conditions of 170–290°C, 14–28 MPa. For a two-reactor process with a Ziegler catalyst, typical conditions are 100–120°C, 7–20 MPa for oligomerization and termination at 260–300°C, 2–5 MPa. With a nickel catalyst, more mild oligomerization conditions are possible, 80–120°C, 7–14 MPa. A typical product distribution is 5–14 wt% C_4, 25–50 wt% C_6-C_{10}, 15–20 wt% C_{12}-C_{14}, 11–15 wt% C_{16}-C_{18}, and 3–15 wt% C_{20+} One process can produce a very narrow distribution of closer to 75% C_6-C_{10} olefins. Depending on the process conditions, catalyst, and molecular weight of the olefin in question, the product comprises 88–98% linear α-olefins. Distillation is used to separate the crude olefin product into molecular weight ranges required for various applications (e.g., oxo for detergent alcohols, LAB production).

Two methods are commonly used for narrowing the molecular weight range of the ultimate product. Low molecular weight olefins may be recycled to participate in further oligomerization (see above). Secondly, light and heavy fractions may be metathesized to produce two smaller olefinic molecules (see section "Metathesis", directly below).

Metathesis [85]

An intriguing reaction in which the carbon–carbon double bonds of two olefins are broken and remade, resulting in the switching of their alkylidene groups, is referred to as olefin metathesis.

Table 10.28 Major uses of downstream derivatives

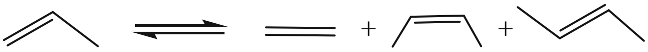
Chemical	Major uses
Vinyl chloride monomer	≈100% polyvinyl chloride
Ethyl acetate	50% surface coatings; inks, solvent applications
Ethylene glycol	85% polyethylene terphthalate; 10% antifreeze
Ethylene oxide glycol ethers	Solvents, paints, inks, cleaning fluids
Chloroethanes	Flurocarbon intermediates
Phenol	44% BPA; 27% phenol/formaldehyde resins; 7% caprolactam; 3% alkylphenols
Acetone	29% methyl methacrylate; 29% solvent applications; 22% bisphenol A; 10% aldol chemicals
Hydroquinone	41% antioxidant; 21% polymerization inhibitor; 14% photographic processing
MIBK	67% solvent applications; 22% rubber antiozonants
2-Ethylhexanol	70% plasticizer esters; acrylate esters; solvent applications
Propylene glycol	30% unsaturated polyester resins; functional fluids; personal care; P type glycol ethers
Propylene oxide glycol ethers	Solvents, paints, inks, cleaning fluids
n-Butanol	33% acrylate esters; 15% acetate esters; 15% glycol ethers; solvent applications
Isobutanol	Acetate esters; glycol ethers; other esters; solvent applications
Bisphenol A	69% polycarbonate; 27% epoxy resins
Styrene	59% polystyrene; 16% ABS/SAN resins
Cyclohexanone-cyclohexanol	62% caprolactam; 34% adipic acid; solvent applications
Formaldehyde	Resins: 46% urea/formald, 11% phenol/formald, 8% polyacetal, 6% melamine'/formald; 5% pentaerthyritol; 4% 1,4-BDO
Neopentyl polyols—NPG, PE,TMP, TME	70% coatings resins; 10% lube esters; 7% unsaturated polyester resins; 3% acrylates/methacrylates
Ketene	70% acetic anhydride; 14% diketene; 5% sorbic acid
Ketene derivatives, diketene	Acetoacetate esters; arylides
Methyl chloride	Intermediate for production of silicones; methyl cellulose solvent, butyl rubber solvent
Methylene chloride	Adhesives, paint removal; solvent for chemical/pharma processes; aerosols; metal cleaning
Chloroform	Intermediate for production of HCFC-22 (R-22); raw material for substitution fluorination for fluorpolymers, e.g., Teflon
Carbon tetrachloride	Intermediate for production of HFC-245fa and HFC-365mfc fluorocarbons
Acetate esters (methyl, ethyl, *i*/*n*-propyl, *i*/*n*-butyl)	Solvents
1,4-Butanediol	50% THF/spandex; 24% polybutylene terephthalate; 14% gamma-butyrolactone; 9% polyurethanes

$$\mathrm{RHC{=}CHR + R'HC{=}CHR' \rightleftharpoons 2\ RHC{=}CHR'}$$

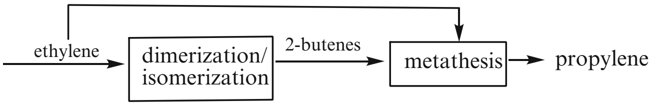
The analogous reaction for carbon–carbon triple bonds, alkyne metathesis, was also discovered at about the same time, in the early 1960s.

$$\mathrm{RC{\equiv}CR + R'C{\equiv}CR' \rightleftharpoons 2\ RC{\equiv}CR'}$$

One of the earliest applications of olefin metathesis was developed at Phillips Petroleum and is known as the Triolefin Process [86]. It illustrates the simplest case of olefin metathesis in which propylene is converted to ethylene and both cis and trans isomers of 2-butene.



The process was practiced as described (propylene-to-ethylene/butene) by Shawinigan Chemical from 1964 to 1972 but the demand for propylene vs. ethylene shifted. Now processes have been developed by ARCO (now Lummus OCT: olefin conversion technology) and Axen (Meta-4) to execute variations of the opposite in which C_4 streams, derived from hydrocarbon cracking, are reacted with ethylene to produce propylene. This technology thus provides the opportunity to enrich the ethylene/propylene product ratio of a cracker to favor propylene. In OCT, ethylene is dimerized to butenes and then cracked with additional ethylene ("ethenolysis") to yield propylene, allowing total conversion of such a stream to propylene. Several different metal oxides, such as $CoO\text{-}MoO_3/Al_2O_3$ oxide (120–210°C), WO_3/SiO_2 (450–500°C) and Re_2O_7/Al_2O_3 (450–500°C), are employed as catalysts.



Another example of ethenolysis to produce valuable products is the cracking of unsaturated fatty acid esters for production of fuels from renewable materials [87].

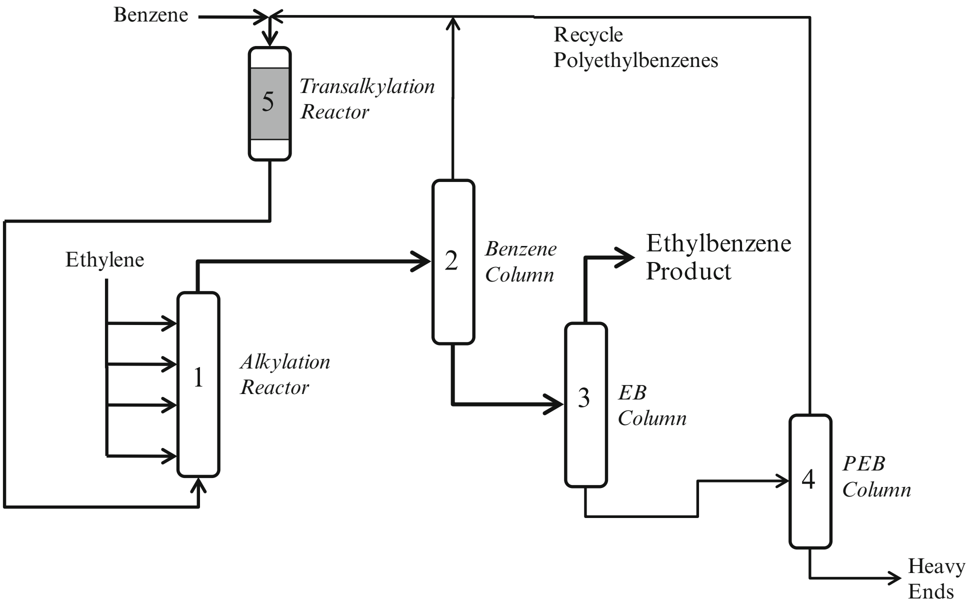
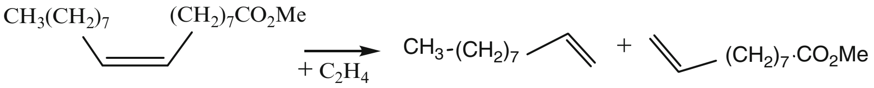
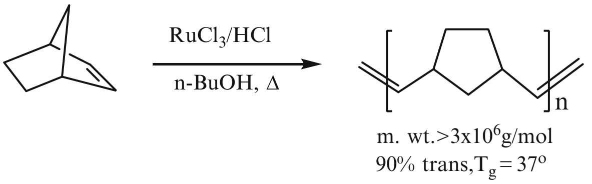


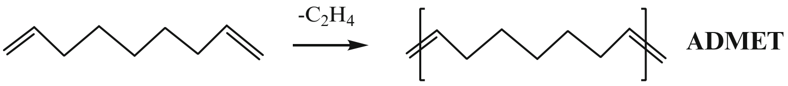
Fig. 10.18 Flowsheet alkylation of benzene to ethylbenzene

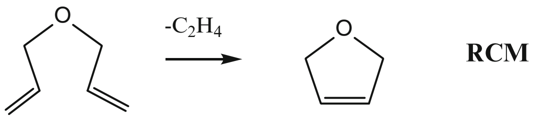


Olefin metathesis is an equilibrium controlled reaction, and the equilibria of a related set of olefins are typically close to a statistical mixture, thus limiting the conversion of reactants. However, if one olefin of the equilibrating mixture can be selectively removed then a metathesis reaction can be driven to high conversion. A simple strategy which exploits this concept is the reactions of a-olefins which produce low boiling ethylene (the oppositie of ethenolysis). As the volatile alkene escapes from solution the reaction proceeds to completion. This approach is exploited in polymer synthesis by reacting various α-, ω-dienes to generate ethylene and a growing polymer chain (referred to as "Acyclic Diene Metathesis" ADMET, polymerization). Likewise, dienes which can produce a favored ring size yield "Ring Closing Metathesis." (RCM).

A second strategy for driving a metathesis reaction toward complete conversion is demonstrated by polymerization of cyclic monomers, such as norbornene, which release enough energy upon ring-opening to drive the polymerization. This method is known as ring-opening metathesis polymerization (ROMP) [88].







The reaction above gives an early example of the homogeneous catalysis of metathesis reactions by metals such as ruthenium or iridium [89, 90]. The development of single-component metal-carbene metathesis catalysts, in particular those of ruthenium, and their applications in organic synthesis have been the subject of an enormous body of research in the last 20 years and the 2005 Nobel prize in chemistry was awarded to three of the pioneers in this area: Yves Chauvin, Robert Grubbs and Richard Schrock. ROMP is one of the most promising techniques to emerge in this work with ready applications in polymer and material science. The structures of some of the most widely utilized catalysts are given in Fig. 10.18.

In 1977, Shell commercialized a process known as the SHOP Process (Shell Higher Olefin Process) for the production of detergent alcohols at their plant in Geismar, LA. The 2010 capacity of this plant for olefin production is over 900,000 M-tons/year. Detergent alcohols are produced in four steps from ethylene. First, the ethylene is oligomerzed using a homogeneous nickel-based catalyst with a chelating P-O ligand ([Ni] ~50 ppm). Oligomerization is carried out at 80–140°C and 7–14 MPa in a polar organic solvent (1,4-butanediol) which phase separates from C_4-C_{30} α-olefin products [91]. The nickel catalyst solution phase is separated in a high-pressure separator and the C_4-C_{20} α-olefins are distilled from the mixture and are purified and sold as Shell's Neodene linear α-olefins.

In the next step the mixture of α-olefins are isomerized in the liquid phase using a heterogeneous catalyst such as MgO granules. This produces all possible isomers of the even-numbered linear olefins. Then, in the third step, olefin metathesis of this mixture over a molybdenum/alumina catalyst (100–125°C, 10 MPa) further complicates this mixture by doubling the number of olefins within the range to include not only the even (C_4, C_6, C_8, etc.) but also all of the odd-numbered (C_4, C_5, C_6, C_7, etc.) olefins in a thermodynamically controlled equilibrium. The yield of desired C_{11}-C_{14} olefins is 10–15% per pass. These olefins are treated in the final step with syngas ($CO:H_2 = 1:2$) using a phosphine-modified cobalt catalyst (170°C, 10 MPa). In this process internal double bonds are isomerized rapidly, resulting in a highly linear alcohol product (75–90% linear).

Alkylation [92, 93]

Chemistry

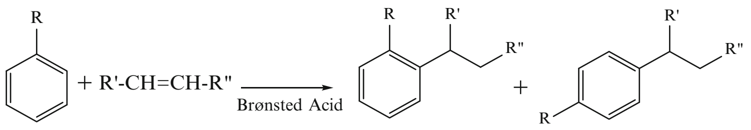
Catalytic alkylation of aromatics is a substitution reaction in which a hydrogen atom on the ring or side chain of the aromatic is replaced by an alkyl group derived from an alkylating agent. Alkylation may occur on both unsubstituted as well as substituted aromatics. The reaction is shown below for ring substitution with an olefinic alkylating agent:

Acid-catalyzed electrophilic substitution (by way of a carbonium ion) gives replacement on the ring. If the alkylation is to occur on the side chain, then nucleophilic (base-catalyzed), or free radical reactions are involved. Most large-scale commercial alkylations are electrophilic in nature, e.g., cumene, ethylbenzene, and linear alkylbenzenes (LAB's) syntheses, but a few are nucleophilic, e.g., isobutylbenzene, dimethylnaphthalene syntheses.

Alkylating agents can be olefins, alkynes, dienes, alcohols, and ethers. For electrophilic substitution, any substrate capable of forming a carbonium ion may be used, although the use of olefins predominates for all applications except introduction of a methyl group. Methanol is a common methylating agent as it can give a carbonium ion equivalent, whereas methane itself is unreactive.

Acid catalysts useful for aromatic alkylations are Brønsted acids containing acidic protons. Examples include protonic acids such as sulfuric, HF, phosphoric, solid phosphoric acid (SPA), polysulfonate cation-exchange resins; acidic aluminas, silicas, alumina-silicas, zeolites (aluminosilicates and silicoaluminophosphates); and acidic halides, such as $AlCl_3$ and BF_3. The latter are relatively inactive in their pure states, but become catalytically active when coupled with trace concentrations of water, alcohols, or hydrogen halides (giving the needed Brønsted acidity). These Lewis acids were the original highly effective Friedel-Crafts catalysts, with high activity at low temperature (<100°C), but have largely been replaced commercially with more selective, less corrosive, and environmentally friendly zeolites.

With strong Brønsted acid catalysis, the olefin is protonated to form an activated electrophile (carbonium ion), which attacks the electron-rich π-complex of the aromatic ring. A new C–C-bond is formed, followed by regeneration of an acidic proton. Acid-catalyzed skeletonal rearrangement of the alkylating agent leads to formation of the most substituted and stable ion (tert > iso > normal). Relative rates of alkylation of the ring and isomerization of the alkylating agent dictate the product distribution. Typically for small alkylating agents like butenes, (especially isobutene) rearrangement is fast, giving almost exclusively tert-butyl aromatics. Thus synthesis of isopropylbenzene (cumene) and diisopropylbenzene from propylene and benzene is highly selective while isobutylbenzene formation must be base catalyzed. Alklyation rate also roughly follows carbonium ion stability, with isopropylation roughly 1,500 times faster than ethylation and 20,000 times faster than methylation.



ΔH_{rxn} = -80 to -115kJ/gmole
(depending on alkylating agent)

R, R', R" = H, methyl, ethyl, propyl, etc.

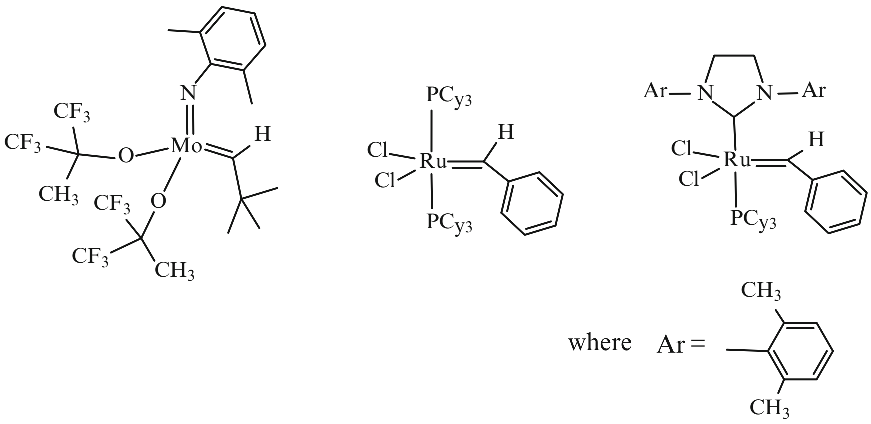
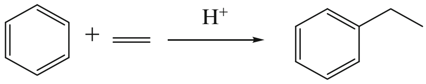


Fig. 10.19 Single-component olefin metathesis catalysts

The rate of alkylation of benzene is slower than alkylation of alkylaromatics, so polyalkylation is kinetically favored. However, many of the most important alkylation products (e.g., cumene and ethylbenzene) are monosubstituted. The traditional method of limiting polyalkylation is to use high ratios of aromatic to alkylating agent, leading to low per pass aromatics conversion, and high separation/recycle costs. One of the most intriguing breakthroughs in alkylation technology in the last two decades has been the commercialization of shape-selective zeolite catalysts, which can largely eliminate polyalkylation, as well as allow much higher than equilibrium production of *para*-isomers (see the section, "C_6-C_8 Aromatics: BTX," above).

Ethylbenzene [94–98]

Well over 90% of the ethylbenzene consumed worldwide is formed by the alkylation of benzene with ethylene.



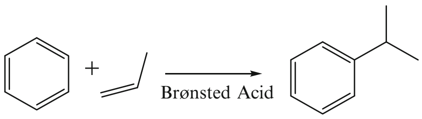
ΔH_{rxn} = -113 kJ/gmole

Early processes employed $AlCl_3$ in liquid phase processes operating below 130°C. Significant polyalklyation occurred, but was controlled by high benzene to ethylene feed ratios and recycle of polyethylbenzene (PEB) materials for equilibration in the alkylation reactor. High temperature (>300°C) vapor phase zeolite (ZSM-5) processes made significant inroads starting in the 1970s, but reactors coked quickly and required frequent regenerations.

Greenfield plants predominantly rely on next generation large pore zeolite catalysts to allow liquid-phase operation at 170–240°C (more than 100°C lower than the older vapor phase processes), at 3.5 MPa, with a much smaller excess of benzene to ethylene required. Low temperature operation decreases formation of by-product xylenes, polyaromatics, and PEB. In a typical commercial process (Fig. 10.19), benzene at less than 7:1 B:E molar ratio is fed to zeolite-filled alkylation beds in series (**1**), with cooled ethylene added between stages. Older vapor phase processes required a B:E of 18:1. Interstage cooling provides additional temperature control. Purification is accomplished in a three column distillation sequence. Unreacted benzene (for recycle) is taken overhead in the benzene column (**2**), EB overhead in the product column (**3**), and PEB for recycle in the third column (**4**), with polyaromatics and other heavies as underflow. PEB are converted to EB by reaction with benzene in a separate transalkylation unit (**5**). Ethylene and benzene consumption are respectively 0.265 and 0.738 kg/kg EB. Ethylene containing up to about 30% ethane can be used with the addition of a column for recovery of ethylene from off-gas ethane by absorption into benzene feedstock.

Cumene and Diisopropylbenzene [99–103]

Cumene production via alkylation of benzene with propylene, began in World War II as a high octane aviation fuel component. This early vapor phase process operated at 200–250°C, 1.5–3.5 MPa, over a silicophosphoric acid (SPA), with a 7:1 benzene:propylene ratio, with benzene-based yields above 95%.

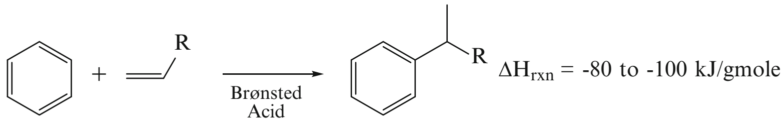


ΔH_{rxn} = -106 kJ/gmole

The SPA process was widely used for decades with little change, competing well against even zeolite based processes. Although successful for ethylbenzene synthesis, similar early generation zeolites were found to be unsuitable for cumene production. These narrow pore zeolites promote excessive *n*-propylbenzene formation due to (unwanted) shape selectivity, as well as propylene oligomerization (propylene is more reactive to oligomerization than ethylene). In the mid-1990s, a proprietary zeolite was invented with unique crystal structure that largely eliminated these former disadvantages, even at benzene:propylene ratios less than 2:1. Commercialized liquid phase processes using this zeolite are capable of greater than 99.7% yield of cumene, at 99.97% purity. The process flowsheet for cumene production looks very similar to that of ethylbenzene, with a mult-stage series alkylation system, a separate transalkylation step, and a three-stage distillation sequence for purification. Benzene, cumene, polyisopropylbenzene are taken overhead in successive columns, and heavies as underflow in the final column. Para-DIPB can be made in essentially an identical process, but with a higher propylene to benzene feed ratio.

Higher Alkylbenzenes [104–107]

Linear alkylbenzenes (LAB's) with a C_{10}-C_{14} side chain are important intermediates for the production of alkylbenzenesulfonate anionic detergents (see Chap. 36). The name is somewhat of a misnomer, as the alkylation reaction produces predominately branched 2-phenylalkane derivatives, rather than true linear 1-phenylalkanes.



Alkylation with the propylene tetramer has largely become obsolete, as the resulting branched alkylbenzene sulfonates are not readily biodegradable. The higher alkylbenzenes are synthesized by the liquid-phase reaction of C_{10}-C_{14} α-olefins and internal olefins with benzene, using typical Friedel-Crafts catalysts such as HF, $AlCl_3$, or zeolites. The higher chain olefins are sourced from either: (1) catalytic dehydration of the *n*-paraffin fraction recovered from the kerosene crude distillation cut to give largely internal olefins (see the section, "Internal Olefins from Higher *n*-Alkanes"), or (2) ethylene oligomerization processes (see above, the section, "Oligomerization"), which give largely α-olefins. Newer zeolite catalysts favor the more desirable 2-phenylalkane isomers with both internal and α-olefin feedstocks. Most greenfield LAB plants are now zeolite-based, with their clear advantage in safety, catalyst and capital cost, with elimination of catalyst neutralization stepsand corrosion concerns.

Oxidation

Chemistry [108–115]

Roughly two-thirds of high value synthetic organic chemicals are oxygenated. These include aldehydes, ketones, alcohols, organic acids, anhydrides, esters, and epoxides. Many of these functional groups ultimately are derived from some form of selective oxidation. In its broadest sense, oxidation is the process whereby the oxidation state of a substrate (typically carbon in this chapter) is increased. Although conceptually simple the controlled addition of oxygen to a hydrocarbon feedstock, i.e., partial oxidation, is a tricky feat to master. The principle technical challenges in partial oxidation of organic substrates are (1) finding a catalyst or conditions that give high selectivity toward the desired oxidation state and low selectivity for over or complete oxidation; and (2) carefully controlling reaction conditions, especially temperature and composition, to prevent loss of selectivity, runaway exothermicity, or explosive conditions.

Selective oxidations can be classified into two broad categories:

1. *Electrophilic oxidations* involve activation of dioxygen into electron-seeking forms such as O_2^- or O^- and formation of free radical intermediates.
2. *Nucleophilic oxidations* involve activation of the organic substrate followed by stepwise nucleophilic addition of O_2^- and hydrogen abstraction.

In order of increasing severity, electrophilic oxidations include; (1) addition of atomic oxygen to a high electron density double bond to form an epoxide, (2) formation of an alkyl hydroperoxide via hemolytic cleavage of a C–H bond by an oxygen radical to form an alkyl radical followed by reaction with dioxygen, (3) fission of C–C bonds to form two aldehydic groups, (4) rapid stepwise fission of C–H and C–C bonds to form CO_2 and water—i.e., total combustion. Nucleophilic oxidations in increasing severity include; (1) abstraction of hydrogen from an aldehyde followed by nucleophilic addition of oxygen to form an acid, (2) abstraction of hydrogen from an alcohol substrate to form an aldehyde or ketone, (3) abstraction of hydrogen from a substrate to form olefins, diolefins, aromatics from hydrocarbons, (4) abstraction of hydrogen from an olefin followed by successive nucleophilic additions of heteroatoms such as O, Cl, S, and N.

Electrophilic oxidations alter the carbon skeleton and π electron systems, while nucleophilic oxidations do not. Nucleophilic oxidations often involve a redox mechanism, in which oxygen is supplied indirectly to the substrate from the oxide lattice of the catalyst, and the catalyst lattice is then reoxidized by dissolved or gaseous dioxygen. Important industrial oxidation reactions described in this chapter are classified and summarized in Table 10.29.

The design and operation of reactors for oxidation service are demanding tasks. The desired products of selective partial oxidation are invariably not the thermodynamically

Table 10.29 Summary of oxidation reactions

Product	Substrate	Oxidant	Oxidation type[a]	Catalyst	Catalyst form	Phase	Conversion per pass	Selectivity
Ethylene oxide	Ethylene	>99% O_2	E-1	Ag (K, Cl) on α-Al_2O_3	Heterogeneous	Vapor	7–15%	Up to ~90%
Propylene oxide/styrene	Propylene/ ethylbenzene	Air	E-2	Mo complex or Si-supported TiO_2	Homog or hetero	Liquid	13% EB, ~100% propylene	~90%
Propylene oxide/t-butanol	Propylene/ isobutene	Air	E-2	Mo complex	Homogeneous	Liquid	~30% i-C_4 ~9% propylene	>90%
Phenol/acetone	Cumene	Air	E-2	Co/Mn salts	Homogeneous	Liquid	>99% of cumene	>95%
Hydroquinone/ acetone	Disopropylbenzene	Air	E-2	Co/Mn salts	Homogeneous	Liquid	>90%	>90%
Maleic anhydride	Benzene	Air	E-3, N-1	V-Mo oxides	Heterogeneous	Vapor	96%	73%
Adipic acid	Cyclohexanone/ cyclohexanol	HNO_3	E-3, N-2	Co salts, V-Cu salts	Heterogeneous	Liquid	>95%	92–96%
Acrylic acid	Acrolein	Air	N-1	V/Mo/W oxides	Heterogeneous	Vapor	~100%	90–95%
Methacrylic acid	Methacrolein	Air	N-1	V/Mo/W oxides	Heterogeneous	Vapor	90%	80–90%
C_2-C_8 carboxylic acids	C_2-C_8 saturated aldehydes	Air	N-1	None or Co,Mn, Cu salts	Homogeneous	Liquid	80–95%	>90–95%
Formaldehyde	Methanol	Air	N-2	Ag	Heterogeneous	Vapor	70–80% or >99%	89–92%
Formaldehyde	Methanol	Air	N-2	Fe-Mo oxides	Heterogeneous	Vapor	70–80% or >99%	93–96%
Maleic anhydride	Butane or butenes	Air	N-3,4	V-P oxide	Heterogeneous	Vapor	80%	70%
Acetaldehyde	Ethylene	Air	N-4	Pd/CuCl	Homogeneous	Liquid	>99%	92–94%
Acrolein	Propylene	Air	N-4	Bi/Mo/Fe/Co/K oxides	Heterogeneous	Vapor	>95%	80% acrol. 10% acry acid
Methacrolein	Isobutylene	Air	N-4	Bi/Mo/Fe/Co/K oxides	Heterogeneous	Vapor	>95%	~75–85%
Terephthalic acid	*p*-Xylene	Air	N-4	Co/Mn/Br	Homogeneous	Liquid	>99%	>98%
Phthalic anhydride	*o*-Xylene	Air	N-4	Ti-V oxides	Heterogeneous	Liquid	92–96%	75–85%
Benzoic acid	Toluene	Air	N-4	Co salts	Homogeneous	Liquid	>50%	90%
Cyclohexanone/ cyclohexanol	Cyclohexane	Air	N-4	Co/Cr salts	Homogeneous	Liquid	5–7%	75–80%
VAM	Ethylene/acetic acid	Air	N-4	Pd/Ag/K on α-Al_2O_3 or silica	Heterogeneous	Vapor	8–15% ethylene, 15–35% Acetic acid	94–96% on ethylene, >98% on acetic
VCM	1,2-Dichloroethane		N-4	Cu(II)Cl2 on γ-Al_2O_3	Heterogeneous	Vapor	93–97%	91–96%
1,3-Butadiene	Butanes/butenes	Air	N-3	Ferrite/Zn, ferrite/Mn	Heterogeneous	Vapor	>60%	90–93%

[a] *E* electrophilic; *N* nucleophilic; see text for number designation

stable products. Thus kinetic controls must be exploited to avoid overoxidation or complete combustion. Moreover, complete combustion releases typically four to ten times the energy of selective partial oxidations. Heat management is of paramount importance as non-selective reactions often have higher activation energies than the selective pathways, and are thus accelerated disproportionally by poor heat control. A large body of literature has been built up on heat and mass transfer, as well as reactor design. Some representative references are given below [116–118].

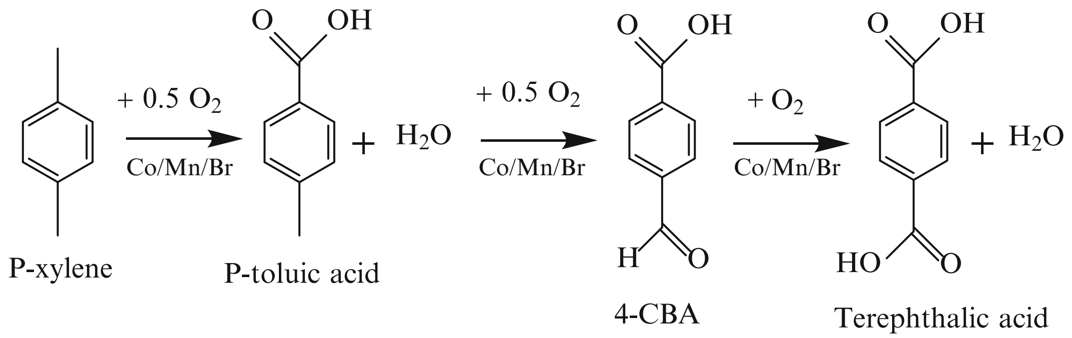
Responsible industrial implementation of oxidation chemistry requires understanding the implications of generating potentially flammable or explosive mixtures. The literature on safety issues, data and discussions of explosive limits is quite extensive. A few key references are cited here [119–123].

Homogeneous Aromatic Oxidations

Terephthalic, Isophthalic, Trimellitic Acids [124–126]

A number of aromatic carboxylic acids such as TPA, isophthalic acid (IPA), and trimellitic anhydride (TMA), are produced by the liquid phase air oxidation of the parent aromatic (*p*-xylene, *m*-xylene, and 1,2,4-trimethyl benzene respectively) with a homogenous catalyst comprising cobalt, manganese, and bromine salts. The key breakthrough in the development of this family of processes was the addition of the bromine promoter which resulted in a remarkably higher reaction rate and selectivity than with Co/Mn alone [127]. TPA is by far the largest volume product of this group and is produced via the stepwise oxidation of *p*-xylene:

The oxidation sequence begins by with hydrogen abstraction from one of the methyl groups of *p*-xylene by a bromine radical to form a benzyl radical. Dioxygen reacts with this species to form a hydroperoxide intermediate that passes through tolualdehyde on the way to toluic acid. Toluic acid is somewhat deactivated toward oxidation compared to *p*-xylene, but conversion of the second methyl group proceeds via a similar pathway, albeit at a slower rate.



ΔH_{rxn} = -1363 kJ/gmole

A simplified flowsheet for one commercial process is given in Fig. 10.20. Air is compressed (**1**) and fed along with *p*-xylene, acetic acid solvent, and catalyst salts into a bubble column reactor or CSTR (**2**). The reaction occurs in the liquid phase at 175–230°C, 1.5–3.0 MPa. The exothermic heat of reaction is removed by evaporation of the reaction mixture. These vapors are carried into the water column (**3**), where the water of reaction is removed overhead, and off-gases are vented to a scrubber to prevent emission of bromine species. The underflow of the water column, comprising acetic acid is returned to the process. Conversion of *p*-xylene is typically above 99%, with greater than 98% selectivity to TPA. Small amounts of acetic acid (about 36 kg/metric ton of TPA) and *p*-xylene are lost by complete oxidation to carbon dioxide and water. The inlet air rate is controlled so that the off-gases contain less than the minimum oxygen composition needed to form explosive mixtures.

TPA is highly insoluble at reaction conditions and largely precipitates as it forms. Efficient absorption and dispersion of oxygen is critical for the reduction of partially oxidized intermediates which can cause purity problems in the final TPA. 4-Carboxybenzaldehyde (4-CBA), generated at up 5,000 ppm in the oxidizer, tends to occlude in the TPA crystals and must be removed to give high quality TPA monomer. A class of compounds generated in ppm levels, known as florenones, are highly colored and result in yellowed polymer if not removed or destroyed in the TPA purification process.

Upon exiting the oxidizer, the crude TPA slurry is centrifuged and washed with acetic acid (**4**). Recycle of the Co/Mn/Br catalyst and purging of by-products in the filtrate and wash streams typically involve proprietary processes (**5**) that are extremely important for good economics. In order to reduce 4-CBA and toluic acid to acceptable levels, the crude TPA is subjected to a series of long residence time post oxidations (**6**) with air at high temperature and pressure. Although still a slurry in the post oxidizer, the crystals continually remelt and recrystallize in a steady state fashion, allowing occluded impurities to be further oxidized in solution. The TPA slurry effluent from the post oxidizers is fed to the flash crystallizer (**7**) for concentration and to build crystal size. Further solvent is removed in the vacuum flash drum (**8**). Crystals are collected on the rotary

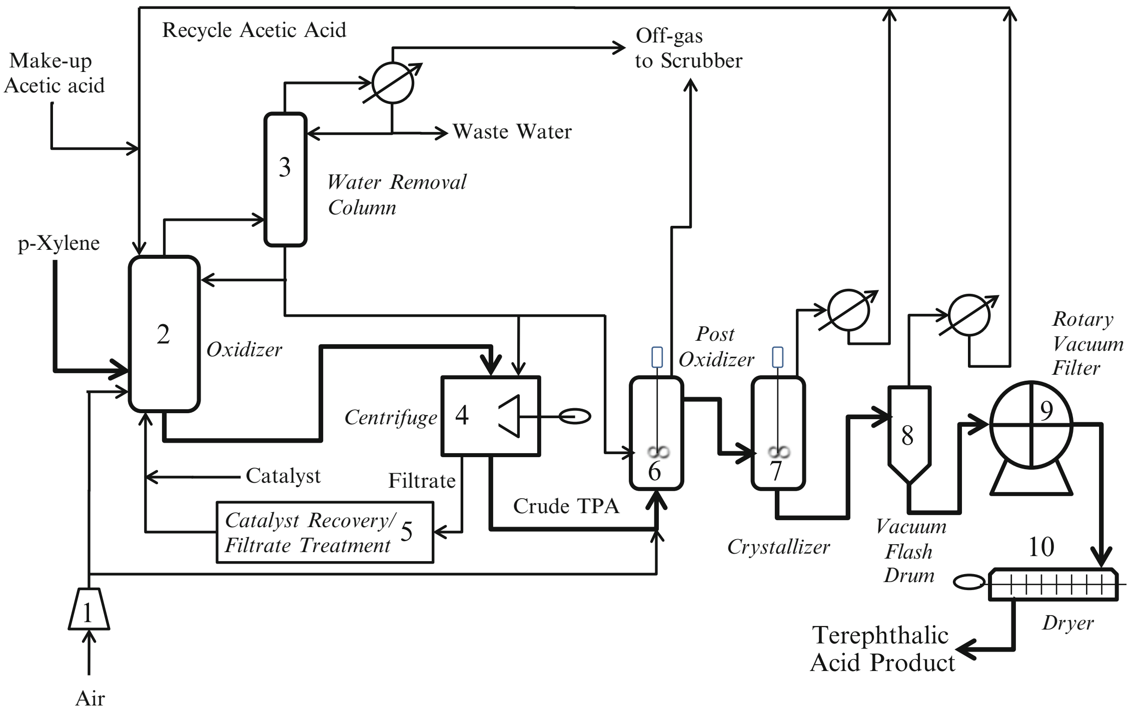


Fig. 10.20 Flowsheet for oxidation of *p*-xylene to terephthalic acid

vacuum filter (**9**) and the dryer (**10**) produces the final polymer grade TPA product with less than 200 ppm CBA and less than 10 ppm toluic acid. The high corrosivity of process streams necessitates the use of expensive titanium reactors and hastelloy in most other locations.

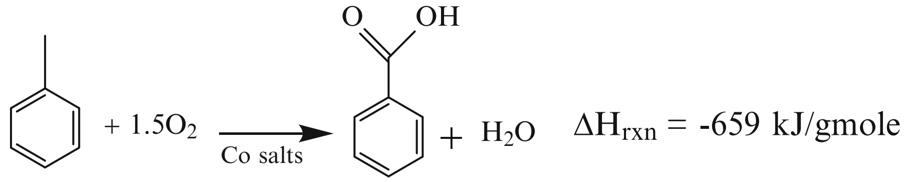
In another widely practiced process, the crude TPA from the centrifuge is dissolved in water at around 260°C. The aqueous TPA solution is hydrogenated over a Pd/carbon catalyst at high temperature and pressure to convert the 4-CBA to *p*-toluic acid and for color body destruction. The hydrogenated effluent is cooled, crystallized, centrifuged, washed, and dried to produce polymer grade TPA.

For many years the best method for purifying TPA to remove color bodies and monofunctional by-products was via esterification with methanol to produce dimethyl terephthalate (DMT), with subsequent distillative purification of DMT. However, the hydrogenation and post oxidation steps have demonstrated superior economics, although giving not quite as pure a product as the DMT route. DMT still finds use in a number of niche applications.

The production of IPA is very similar to TPA. However, IPA is 8–12 times more soluble (depending on temperature) in the reaction media than TPA, does not precipitate as readily, and has less issues with 3-CBA occlusion. On the negative side, IPA's higher solubility requires more stringent crystallization and purging protocols for high recovery.

Benzoic Acid [128, 129]

In a common industrial process, benzoic acid is produced via the liquid-phase air oxidation of toluene with a homogenous cobalt catalyst at 130–170°C, 0.3–1.0 MPa.

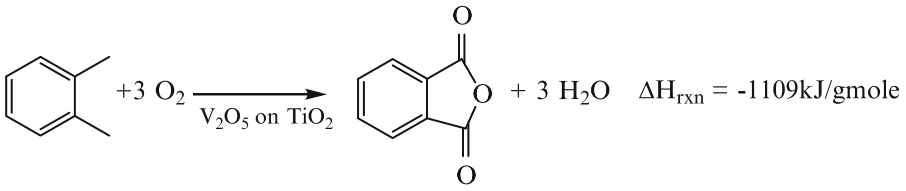


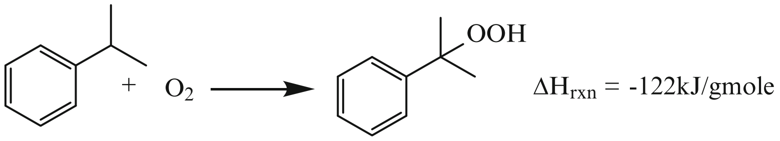
The reaction is carried out normally in a bubble column or CSTR. The exothermic heat of reaction is removed by evaporation of the reaction mixture, or via a circulating cooler. Partial conversion of toluene is typical, with selectivity on toluene of about 90% to benzoic acid, 7–9% to benzaldehyde, and the remainder a mix of benzyl alcohol, benzyl formate, methyl biphenylsm and other minor by-products. Unreacted toluene is flashed off and benzoic acid is purified by distillation. The heavy ends are extracted to recover cobalt. High purity benzoic acid may also be produced by recrystallization from water. As with TPA, the reaction proceeds through a free-radical chain process, with a hydroperoxide intermediate. Cobalt shortens the induction period, and retards by-product formation. Oxygen in the off-gas is kept below the critical explosive concentration to ensure safe operation.

Benzoic acid may be produced with the same catalyst system (Co/Mn/Br) at similar conditions to TPA and IPA, with essentially complete conversion and high yield. However, the use of Br requires expensive titanium and hastelloy metallurgy, whereas the cobalt system can be practiced in cheaper stainless steels.

Phthalic Anhydride [130–133]

In the early 1960s naphthalene from coke oven gases was the predominant raw material for phthalic anhydride production. Since the mid-1990s, over 90% of phthalic anhydride has been made from *o*-xylene over TiO_2-supported V_2O_5 catalsysts with promoters such as K, Cs, P, Sb, or Nb. The formation of phthalic anhydride from *o*-xylene is highly exothermic, with a heat of reaction of $-1{,}109$ kJ/gmole. Nonselective complete combustion is about four times more energetic per mole of xylene than the desired reaction. The reaction takes place at 360–390°C, less than 0.1 MPa.





Conversion of *o*-xylene is 92–96% per pass, with the molar yield on *o*-xylene is around 75–85%. Major by-products are CO_2, benzoic acid, phthalide, *o*-toualdehyde, and maleic anhydride. Formation of potentially explosive feed and product mixtures dictates a very low *o*-xylene to air mass ratio of 1/60 to 1/120 (or about 1.2 mole% in air) [134].

A simplified flow sheet for the high temperature partial oxidation of *ortho*-xylene in air is given in Fig. 10.21. The *o*-xylene feed is vaporized into air and fed into the fixed bed tubular reactor (**1**) where the heat of reaction is controlled by shell-side generation of high pressure steam or exchange against molten salt. The reactor effluent, containing around 1% phthalic anhydride and other by-products, is cooled by steam generation (**2**), and desublimated in a device called a switch condenser (**3**) (a cyclically operated finned heat exchanger in which the anhydride is solidified and periodically melted). The crude anhydride is thermally treated (**4**) at 230–300°C to decompose or polymerize some of the by-products, followed by distillation. The first column removes light ends such as benzoic acid and maleic anhydride (**5**), and the product column (**6**) takes phthalic anhydride overhead from heavies. Off-gases are scrubbed with water (**7**), to recover maleic and fumaric acids.

Peroxidations of Secondary or Tert-Alkyl Benzene Derivatives

Phenol/Acetone or MEK and Hydroquinone/Acetone [135–140]

Greater than 90% of phenol and acetone are produced worldwide by the air oxidation of cumene via a free radical chain mechanism in which cumene hydroperoxide (CHP) is the chain initiator. [1] The reaction is autocatalytic and reasonably exothermic. As shown in Fig. 10.22, cumene and air are reacted in two to four bubble columns in series (**1**) at 85–110°C, 0.5–0.8 MPa, to produce CHP at 25–30 wt% concentration in excess cumene.

---

[1] Most of the remainder of world acetone production is derived from dehydrogenation of isopropanol. See the section, "Dehydrogenation."

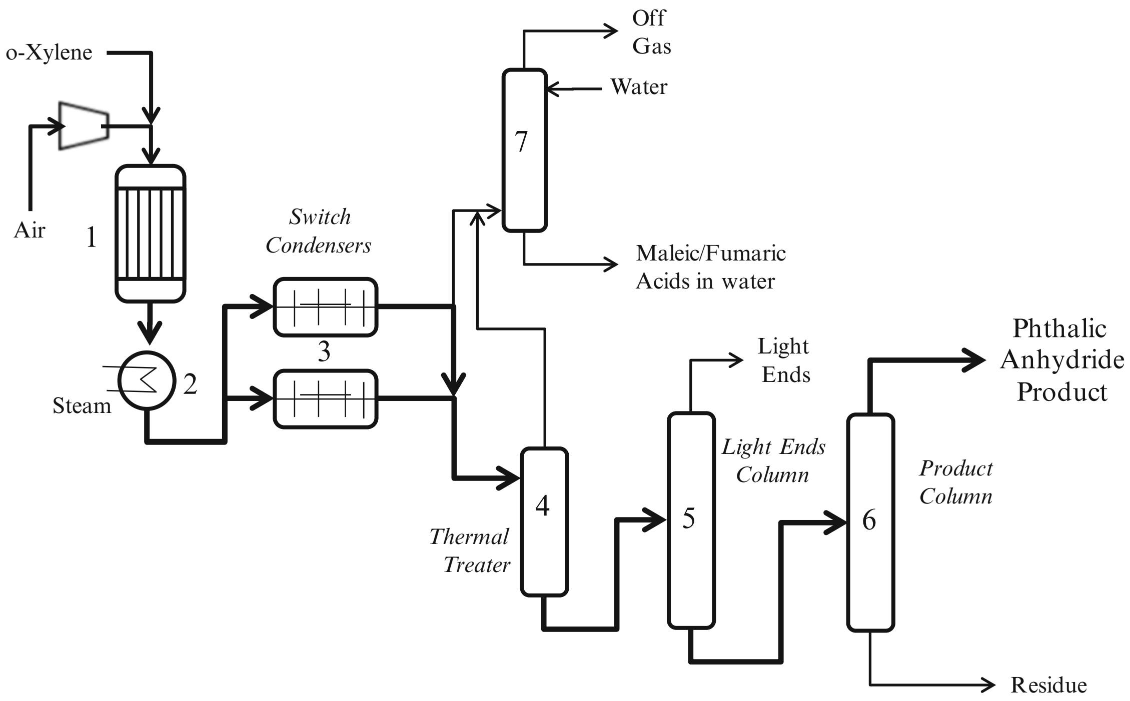


Fig. 10.21 Flowsheet for oxidation of *o*-xylene to phthalic anhydride

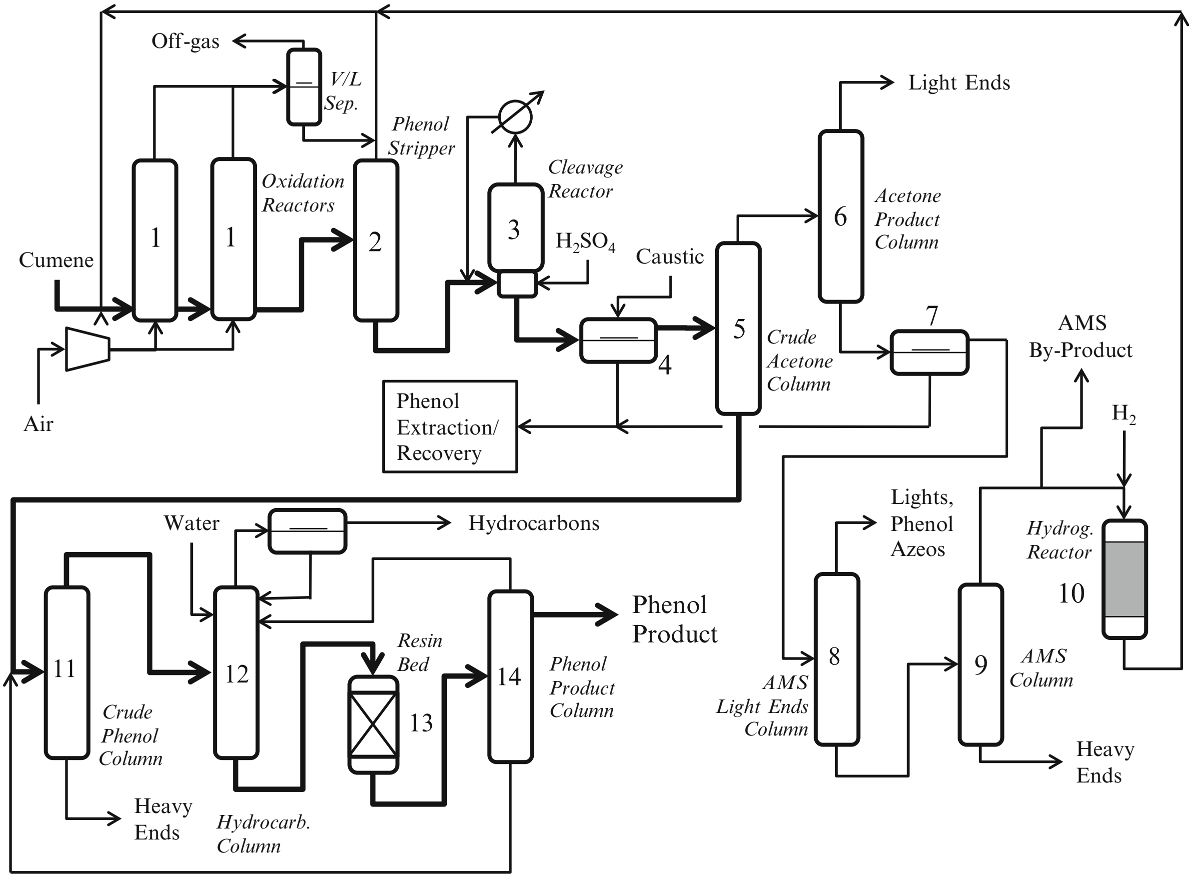
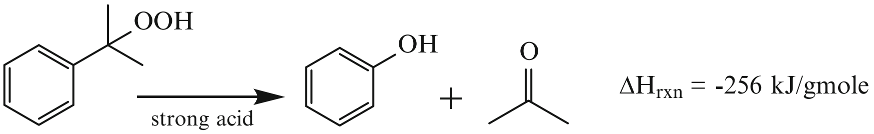


Fig. 10.22 Flowsheet for cumene hydroperoxidation to phenol/acetone

Temperature is controlled by evaporation of cumene and water as well as by cooling in external heat exchangers. Conversion of cumene is essentially complete with a selectivity of greater than 95% to CHP. Phenol is an inhibitor to oxidation, so cleavage to phenol is undesirable at this point. In some processes a small amount of sodium carbonate or sodium hydroxide is added to maintain a neutral solution and prevent the acid-catalyzed decomposition of CHP. Thermal decomposition of CHP (rapid above 130°C) leads to the by-products α,α-dimethylphenyl methanol (DMPM), acetophenone, methanol, formic acid, and acetic acid. Alkaline earth bismuthates, stannates, antimonates, and metal phthalocyanines are known to catalyze the reaction, but are not used industrially as they tend to also catalyze the decomposition of CHP to by-products. As both cumene and CHP are potentially explosive, the oxygen in the reactor off-gas is maintained at 1–6 volume %, below its critical concentration of 8.5–9.5% under reaction conditions.

The crude CHP from the bubble columns is vacuum-distilled (**2**) to remove cumene and concentrated to 65–85 wt% CHP. The concentrated solution is cleaved at 70–90°C in the presence of a strong acid catalyst such as sulfuric acid, to produce phenol and acetone in an exothermic reaction.



Temperature is controlled by boiling of acetone. Under these conditions DMPM dehydrates to α-methylstyrene (AMS), while diacetone alcohol and mesityl oxide form via aldol/dehydration of acetone. The cleavage mixture is neutralized with caustic or sodium phenolate, phase separated (**4**) and the organic layer is refined in a series of distillations. Many distillation sequences are possible and are practiced. One variation is shown in Fig. 10.22. The neutralized organic layer is fed to the crude acetone column (**5**) to distill acetone, mesityl oxide, water, lights, cumene, and AMS overhead from crude phenol and heavies. This distillate is further fractionated under vacuum in the acetone product column (**6**) to remove lights (aldehydes) overhead, acetone as a sidedraw near the top, and heavy water/hydrocarbon azeotropes out the bottom. Traces of caustic are sometimes added to this column to catalyze aldol condensations of aldehyde impurities into higher boiling, easily separable materials. The bottoms stream separates into two liquid phases (**7**). The organic phase is distilled to remove low-boiling AMS-cumene-phenol azeotropes overhead (**8**). After removal of heavies in the AMS column underflow (**9**), the AMS is subjected to hydrogenation (**10**) over Ni, Cu, or Pd catalysts and converted back into cumene for recycle.

In the crude phenol column (**11**), the bulk of the phenol is removed overhead with light organics. The crude phenol bottoms further distilled (**12**) for additional phenol recovery, with the residue used as fuel. Hydrocarbons are azeotroped out of the distilled phenol (**13**) with water and the bottoms phenol product is passed through a resin bed (**14**) to convert any carbonyls present to heavy ends. This treated material is subjected to a final distillation (**15**), with trace hydrocarbon/water removal overhead, and phenol product recovered as a sidedraw.

Decanter water streams are acidified, if needed, to convert the sodium phenolate contained therein back into phenol. Phenol is then recovered by extraction into cumene or other suitable solvents, distilled, and recycled. Overall selectivity on cumene is about 97.5%. Many other process variations are practiced.

Hydroperoxide Co-product Processes for Propylene Oxide

Although differing from ethylene oxide by only one extra carbon in the backbone, propylene oxide (PO) has not been made successfully at commercial scale by the direct oxidation of propylene. The presence of allylic hydrogens dramatically and adversely affects the efficacy of silver-based catalysts for the selective oxidation of propylene to PO with oxygen [141]. In spite of significant research over the last several decades, no catalysts have been found that show commercially viable yields or rates. Instead, all world scale PO plants rely on indirect oxidation of propylene by either the chlorohydrin route (50% of total, see the section, “Propylene Oxide via Propylene Chlorohydrin”) or co-product hydroperoxide routes (50% of total, described herein).

Styrene/Propylene Oxide Co-production [142]

About 15% of commercial styrene is produced by the rather complex air oxidation of ethylbenzene with coproduction of propylene oxide. The remaining 85% is produced via dehydrogenation of ethylbenzene (see the section, “Styrene from Ethylbenzene”). The process consists of four steps:

1. Ethylbenzene is oxidized directly by air to produce ethylbenzene hydroperoxide at around 130°C, 0.2 MPa, and 13% conversion per pass. Yields are about 90% to the

hydroperoxide, 5–7% to acetophenone and α-methylbenzyl alcohol.

2. The hydroperoxide is reacted with propylene (molybdenum or titanium catalysts) at about 110°C, 4.0 MPa, with essentially complete peroxide conversion to give α-methylbenzyl alcohol and propylene oxide. Propylene oxide and ethylbenzene for recycle are separated by distillation from the acetophenone and α-methylbenzyl alcohol.
3. α-Methylbenzyl alcohol is dehydrated in the vapor phase at about 250°C and low pressure over an acidic oxide such as alumina, to produce styrene. Styrene is purified by distillation.
4. Recovered acetophenone is hydrogenated in the liquid phase (Zn-Cu oxide catalyst) at 90–150°C, 8.0 MPa, to give α-methylbenzyl alcohol at 92% selectivity.

This route offers a pathway to propylene oxide without chlorohydrin production, but produces a weight ratio of styrene to propylene oxide of 2.5 to 1, which does not match the relative market demands of the two products and is quite capital intensive.

Tert-Butanol/Propylene Oxide Co-production [143]

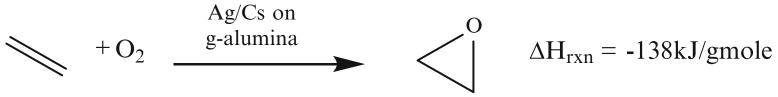
This route is similar to the styrene co-product process:

1. Isobutene is oxidized in the liquid phase with pure oxygen to produce tert-butyl hydroperoxide at around 130°C, 2.5–3.5 MPa, and 15–40% conversion per pass. Yields are about 50% to the hydroperoxide, 46% to tert-butanol, and a minor amount of aldehydes and ketones.
2. The hydroperoxide is reacted with propylene at a 1:10 M ratio (molybdenum catalyst in toluene) in a two-stage process. The first stage is held at about 110°C, 4.0 MPa, with the second stage elevated to 120°C to ensure essentially complete peroxide conversion. Propylene conversion is about 9% per pass at greater than 90% selectivity. Propylene and other lights are removed overhead by distillation for recycle, followed recovery of crude propylene oxide overhead and tert-butanol, catalyst, and other higher boiling by-products as underflow.
3. Production of high purity PO requires further distillation to remove lights such as ethylene oxide, and extractive distillation with a heavy hydrocarbon to remove close-boiling or azeotrope-forming hydrocarbons or carbonyls.

This route also offers a pathway to propylene oxide without chlorohydrin production, but produces a weight ratio of tert-butanol to propylene oxide of about 3 to 1. Economic utilization of the tert-butanol relies on demand for methyl-tert-butyl ether, or other secondary isobutylene markets.

Ethylene Oxide [144–149]

Ethylene oxide (EO) is one of the most versatile and reactive chemicals produced commercially at large scale. Formerly, EO was produced via the chlorohydrin of ethylene, similar to propylene oxide (see below). This route has been abandoned commercially in favor of the direct oxidation of ethylene. The use of high purity oxygen (>99%) instead of air as the source of the oxidant has become standard. A simplified flowsheet of a typical oxygen-based EO plant is shown in Fig. 10.23. Ethylene is reacted with compressed (**1**) oxygen over a silver-based catalyst in a multi-tubular fixed bed reactor (**2**), with a hot spot temperature of 250–275°C, 1–2.2 MPa pressure.



The complete combustion of ethylene to CO_2 and water releases −1,688 kJ/gmole. Commercial selectivities range from 80 to 90% on ethylene over the lifetime of the catalyst, resulting in a net heat load of 230–450 kJ/gmole. Ethylene conversion is 7–15% per pass. The off-gas from the reactor typically comprises 1–2 vol% EO, 5–6 vol% CO_2, 5–9 vol% O_2, 15–40 vol% ethylene, 5–15% inerts (Ar, N_2, ethane), and 1–60 vol% methane. The methane ballast gas acts as a heat sink and keeps the recycle gas above the upper explosive limit at the oxygen-recycle gas mix point.

The reactor effluent is cooled and EO is absorbed into water (**3**), or in some specialized plants into ethylene carbonate. The EO is recovered by distillation (**4**, **5**, **6**) from the water, if desired in pure form. Alternatively, the EO is reacted with water to produce a mixture of mono-, di-, and tri-ethylene glycols directly, or with CO_2, then water, to produce higher purity monoethylene glycol (see Section 10.8 below). The safe handling of concentrated EO requires an appreciation of its volatility (normal boiling point of 10.8°C), and ability to form explosive mixtures without an additional oxidant (the upper explosive limit goes up to 100% EO) [150, 151].

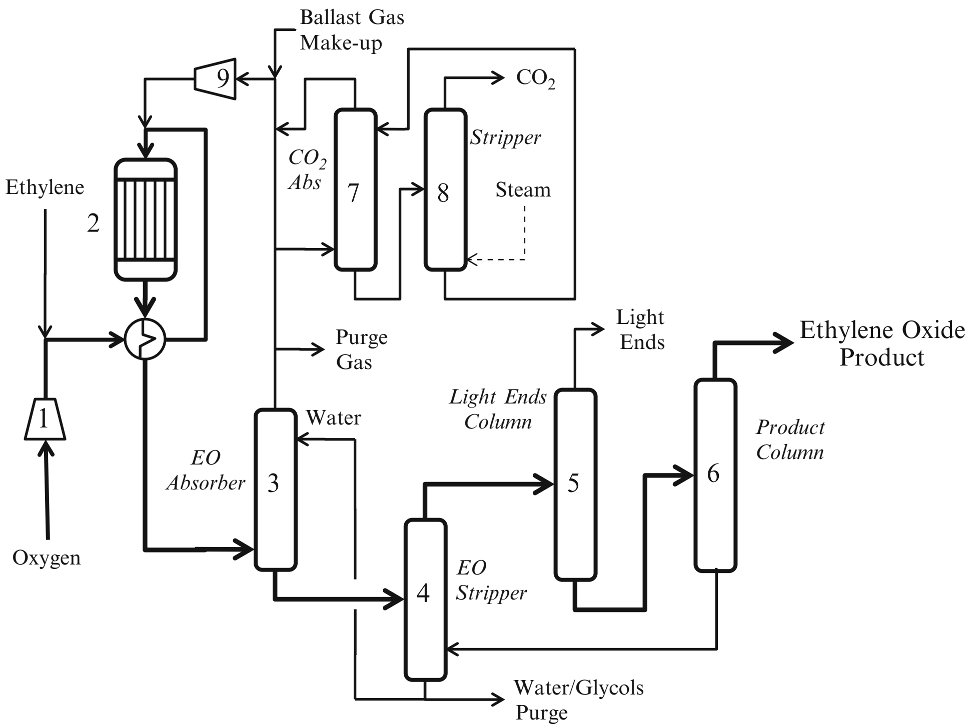


Fig. 10.23 Flowsheet for oxidation of ethylene to ethylene oxide

Carbon dioxide is a modest reversible retardant to the epoxidation reaction; a slipstream of EO-free recycle gas is treated with hot potassium carbonate to absorb out the CO_2 (**7**). The CO_2-rich absorbent is regenerated by vacuum steam stripping (**8**) [152]. A small purge is also removed to prevent the build-up of inerts from the ethylene and oxygen feeds, before recompression and recycle (**9**).

The only effective catalyst discovered thus far for this epoxidation reaction comprises silver on ultrapure γ-alumina, typically with a cesium promoter. Catalyst selectivity has dramatically improved over the past 40 years, from about 65% in 1966 to a maximum of about 90% today. A small continuous feed of a chlorinated hydrocarbon (1–10 ppmv), such as 1,2-dichloroethane, vinyl chloride, or ethyl chloride helps to suppress the combustion of ethylene, significantly improving yield and controllability. The beneficial effect of the chlorinated hydrocarbon was found serendipitously in the early days of commercial silver-catalyzed ethylene oxidation. The air intake for an EO plant was located near the off-gas outlet of a vinyl chloride plant. Unexpected improved performance of the EO reactor was eventually traced back to small quantities of chlorinated hydrocarbons present in the inlet air.

By adjustment of the Cs/Ag ratio on the catalyst, the analogous epoxide of butadiene, 1,2-epoxybutene, can be made by a similar vapor phase oxidation process [153].

Wacker Oxidation: Acetaldehyde [154–158]

Acetaldehyde, a very reactive and versatile molecule, historically was an important intermediate for the production of a wide variety of chemicals. Large volume applications included:

(a) *n-Butanol*, via aldol condensation to crotonaldehyde, followed by hydrogenation—now supplanted by hydroformylation of propylene (see section "Hydroformylation, Lower C_2-C_5 Olefins").

(b) *Terephthalic acid*, added as a peroxidant rate accelerator for the oxidation of *p*-xylene to TPA—now supplanted by more effective Co/Mn/Br catalyst systems.

(c) *Vinyl acetate*, via reaction with acetic anhydride or ketene to form ethylidene diacetate (EDA), with subsequent cracking to VAM—now supplanted by oxidative esterification of ethylene with acetic acid (see section "Esters via Oxidative Acetylation of Olefins, VAM").

(d) *Acetic acid*, via oxidation (see section "Oxidation of Aldehydes to Acids")—largely supplanted by carbonylation of methanol (see "Acetic Acid").

(e) *Ethyl acetate*, via Tishchenko reaction (see section "Tishchenko Reactions: Ethyl Acetate and 2,2,4-Trimethyl-1,3-Pentanediol Derivatives")—still practiced, competes against acetic acid-ethanol esterification (see section "Low-Boiling Esters").

(f) *Pyridine derivatives*, via reaction of paraldehyde (acetaldehyde cyclic trimer) with aqueous ammonia and formaldehyde [159]—still a significant outlet for acetaldehyde. Current applications are dominated by the latter three, while the first three, once the biggest uses, are essentially gone.

Acetaldehyde has been made by dehydrogenation of ethanol over Ag (similar to formaldehyde, See section, "Formaldehyde"). However, much of the remaining world capacity is based on the homogeneous $PdCl_2/CuCl_2$-catalyzed

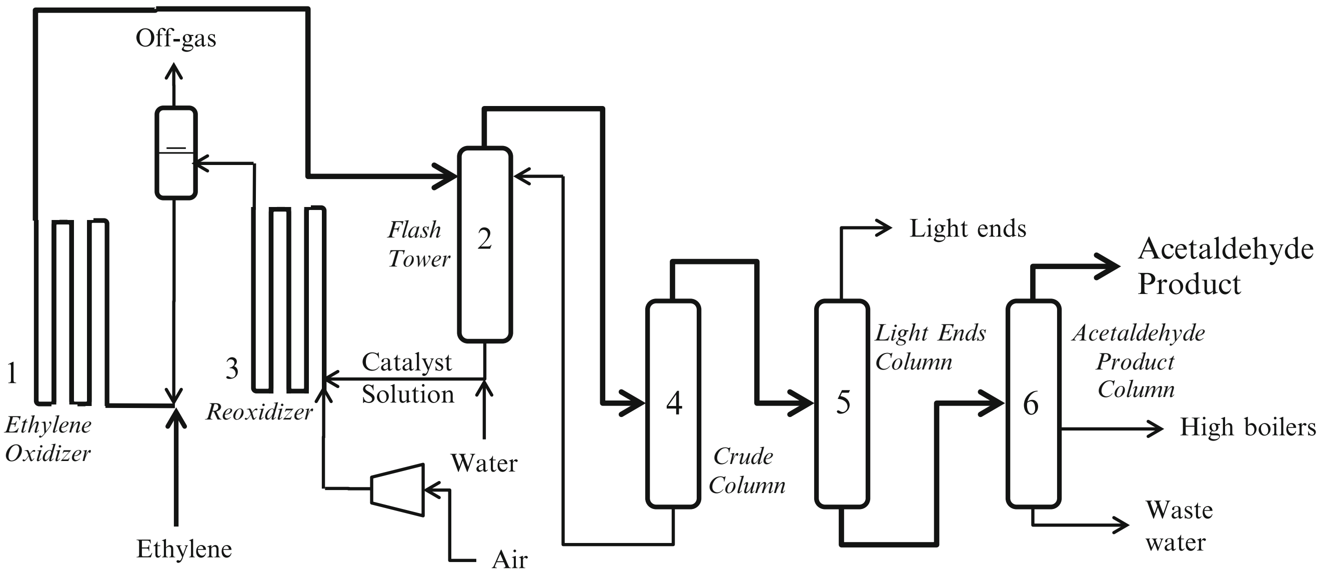
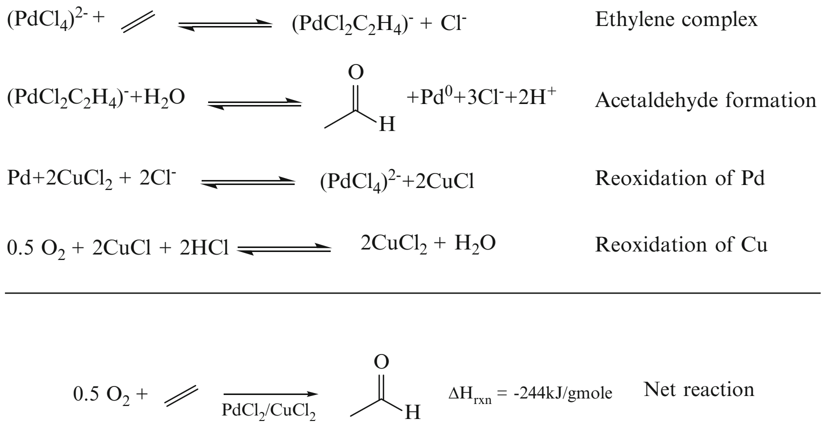


Fig. 10.24 Flowsheet for Wacker oxidation of ethylene to acetaldehyde

direct oxidation of ethylene with air developed in the late 1950s. The catalytic cycle begins with the complexation of Pd with ethylene. Water addition leads to hydride transfer and reductive elimination to produce acetaldehyde, reduced Pd, and hydrochloric acid. Cupric chloride reoxidizes the reduced Pd, regenerating the active Pd species. Finally, cuprous chloride is re-oxidized by oxygen back to Cu(II). The net reaction is the addition of 0.5 moles of oxygen to ethylene.



The palladium catalyzed oxidation of an olefin is highly efficient and only a small concentration of expensive Pd is required for commercially viable rates if sufficient (and comparatively inexpensive) reoxidant is present.

Both one-stage (ethylene reaction and entire redox cycle occur in one vessel) and two-stage (ethylene reaction and reoxidation of cuprous chloride with air in separate vessels) processes have been commercialized, although the two-step process was and is more prevalent. In a typical commercial two-step process, as illustrated in Fig. 10.24, ethylene gas is mixed with freshly reoxidized aqueous catalyst solution and passed through a tubular serpentine reactor (**1**) at 105–110°C, 0.9–1.0 MPa, where acetaldehyde is formed with close to 100% conversion of ethylene. The catalyst solution contains a substantially higher molar concentration of Cu(II) than Pd (typically 150 to 200 to 1), allowing multiple Pd turnovers per reactor pass. The reactor effluent is then flashed (**2**) to remove most of the acetaldehyde as vapor, along with water, and by-products, prior to the lean reduced catalyst solution entering the second tubular reactor (**3**). Air is compressed to about 1.0 MPa and allowed to react with Cu(I) ions at about 100–110°C, to regenerate Cu (II). Recycled and make-up water is added to maintain the catalyst concentration before reentering the ethylene reactor. Conversion of ethylene is close to 99%, with a selectivity of greater than 94%. By-products include chloroacetaldehydes (1–2 mole%), ethyl chloride, chloromethane (0.5–1.0%), carbon dioxide (~0.8%), and acetic acid (2–4%) resulting from over-oxidation of acetaldehyde.

The crude acetaldehyde is purified by a series of distillations. Acetaldehyde is concentrated in the crude column (**4**), with underflow returned as reflux to the flash tower. In the light ends column (**5**), low-boiling substances (chloromethanes, chloroethane, carbon dioxide) are removed overhead. Aqueous acetaldehyde and high boilers are distilled further in the product column (**6**), with high purity acetaldehyde collected as distillate, water/acetic acid as underflow. A fraction of the underflow is purged to prevent buildup of acetic acid, the rest recycled. Mid-boiling chloroaldehydes are removed as azeotropes with water in a sidedraw. This stream may be further concentrated if desired in a sidedraw stripper. The catalyst solution contains appreciable free hydrochloric acid and is quite corrosive. Significant parts of the plant must be constructed of corrosion-resistant materials such as titanium. Small amounts of chloride are lost in the chlorinated by-products. A constant makeup of chloride is fed as hydrochloric acid.

This process may also be adapted for higher olefins, with production of the corresponding internal carbonyl (i.e., ketone) predominating. Thus, propylene gives acetone and 1-/2-butenes give methyl ethyl ketone (MEK), typically 88–92% yield on the olefin. Acetone was produced commercially in Japan for a number of years via the Wacker oxidation of propylene, but the plant has subsequently shut down.

Considerable literature has been published in the past two decades on the use of vanadium-molybdenum-phosphorus homogeneous heteropolyacids (HPA) as replacements for the copper-based reoxidant. The required $PdCl_2$ concentration is considerably reduced, HCl concentration is essentially zero, and formation of chlororinated by-product species is claimed to be virtually eliminated [160, 161]. No commercialized applications of HPA systems are known.

Acrolein/Acrylic Acid and Methacrolein/ Methacrylic Acid [162–165]

Acrylic and methacrylic acids are produced commercially by similar two-step oxidations of propylene and isobutene/isobutene. Propylene oxidation is now the dominant route for acrylic acid, while several methacrylic acid processes are competitive including acetone cyanohydrin (see Chaps. 14 and 22), propionaldehyde/formaldehyde aldol/oxidation (see section “Mannich Base Condensation: Methacrolein”), ethylene carbonylation, and isobutene oxidation. For acrylic acid, propylene/air/steam at a ratio of 1:8:4 (plus recycled gas) is fed over a bismuth molybdate oxide catalyst in a multi-tubular fixed bed reactor at 300–400°C, 0.15–0.25 MPa.

Conversion of propylene is greater than 95% per pass, with a yield of 80% acrolein, 5–10% acrylic acid. By-products are CO, CO_2, acetic acid, acetaldehyde, and polyacrolein. The reactor is cooled by molten salt circulation and steam generation. A world scale reactor with 40,000 tubes in a single shell can produce 100,000 metric tons/year, with a catalyst lifetime up to 10 years.

Pure acrolein can be obtained by water washing of the reactor gases to knock out acids, followed by quench absorption in water. The crude, wet acrolein is distilled overhead, with additional steps for light/heavy ends removal. Acrolein is prone to oligomerization; hydroquinone or a derivative is added as a stabilizer.

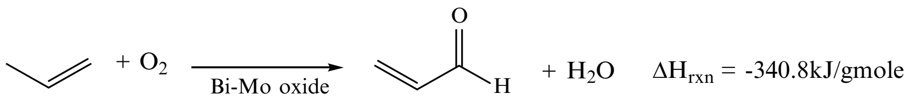
More commonly acrylic acid is the desired end product. The crude acrolein gaseous effluent is sent directly to a second salt/steam cooled multi-tubular fixed bed reactor operating at around 260°C to complete the oxidation to acrylic acid.

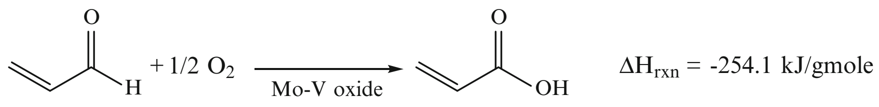
Conversion of acrolein is essentially 100% at greater than 90% yield. Modern catalysts comprise Molybdenum– vanadium oxides with copper, tungsten, or cerium promoters. The reactor effluent is cooled further by steam generation.

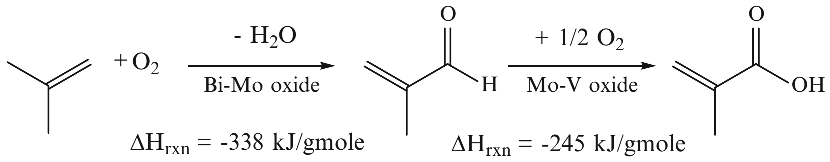
Many different approaches are practiced for the recovery and purification of the acrylic acid. In one variation, the crude acrylic acid is quenched and absorbed into water to produce a 20–70 wt% solution in water. The acrylic acid is concentrated by extraction into an organic solvent (examples are ethyl acetate, MEK, butyl acetate, ethyl acrylate), followed by azeotropic distillation of the extract phase. Water is removed overhead as the solvent-water azeotrope, and the dry acrylic acid bottoms product is distilled overhead in a second column, with oligomers, dimers, and polymers taken as underflow. Additional acrylic acid is recovered from the heavy ends by decomposition of acrylate dimers in an evaporator.

In a second recovery scheme, the cooled acrylic acid reactor effluent is absorbed into a high boiling ester rather than water. This eliminates the extractor, but azeotropic distillation is required for drying, and the acrylic must be distilled from the ester. In both processes, an inhibitor (phenothiazone, hydroquinone, or hydroquinone monomethyl ether) is added to prevent polymerization of the acrylic acid.

In an analogous process for methacrylic acid, isobutene/air/steam at a ratio of 1:11.5:7.5 (plus recycled gas) is fed over a bismuth molybdate oxide catalyst at about 350°C, followed by second stage oxidation over Mo–V oxides at about 280°C







The multi-tubular fixed bed reactors are cooled by molten salt circulation and steam generation. Conversion of methacrolein is around 90% per pass to keep selectivity to 80–90% methacrylic acid. Product recovery by distillation is similar.

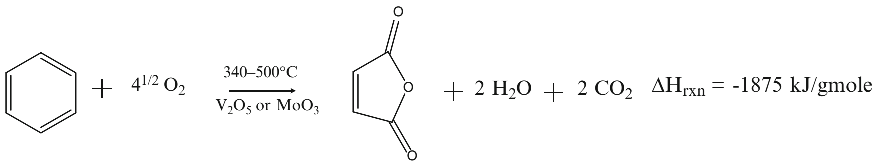
KA Oil: Cyclohexanone/Cyclohexanol [166–168]

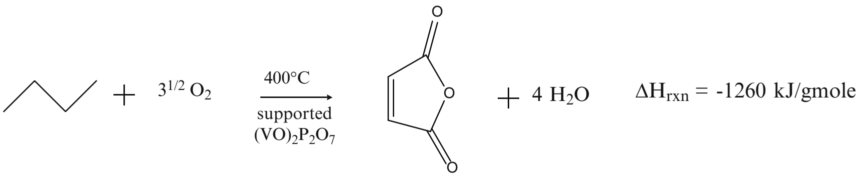
A mixture of cyclohexanone and cyclohexanol known as KA (Ketone-Alcohol) oil is made predominately by the liquid phase cobalt-catalyzed air oxidation of cyclohexane at 140–160°C, 0.8–2.0 MPa. A typical catalyst is the soluble Co(II) naphthenate, which gives a ketone to alcohol ratio of 1:1 to 1:3.5. Conversion must be kept low (about 5–7% per pass) in order to achieve selectivities as high 75–77%. The addition of chromium increases the ketone to alcohol ratio close to 2:1, while promotion with boron compounds such as boric anhydride (B_2O_3), boric acid (H_3BO_3), or metaboric acid (HBO_2) results in a ketone to alcohol ratio as low as 1:10 and increased selectivity to 85–90%. Stoicheometric amounts of slurried borate are required and must be recycled. In an alternative approach, the cyclohexyl hydroperoxide intermediate is formed in an initial uncatalyzed air oxidation step, followed by decomposition in the presence of cobalt or other transition metal catalyst. Yields as high as 84% are achievable with the high hydroperoxide route. In all process variations, the crude KA oil is purified by water extraction for removal of dibasic acids such as adipic, glutaric, and succinic, as well as C_4-C_6 carboxylic and hydroxycarboxylate acids, followed by distillation.

KA oil is an important intermediate for the production of both adipic acid and caprolactam for nylon-6,6 production (see Chap. 22). Purified cyclohexanol and cyclohexanone both find limited uses as solvents.

Maleic Anhydride [169–172]

Prior to 1975, benzene was the predominant raw material for the production of maleic anhydride. Currently worldwide, over 80% of maleic anhydride is made from butane or butane/butene mixtures of high alkane content. About 2% is recovered as a by-product of phthalic anhydride production, and the remainder is derived from benzene. The formation of maleic anhydride from either benzene or butane is highly exothermic.





Nonselective complete combustion of benzene (−3,136 kJ/gmole) and butane (−2,655 kJ/gmole) are much more energetic than the desired reaction.

Benzene is oxidized in air over V_2O_5 or MoO_3 catalysts in a multi-tubular fixed bed reactor at 340–500°C, 0.15–0.25 MPa, with molten salt-steam generation for heat control. The benzene content is kept below the lower explosive limit (LEL). Typically more steam is generated than is needed in the process. Selectivity to maleic anhydride is around 73 mole% at 96% conversion per pass.

Typical conditions for the air oxidation of *n*-butane are a temperature of about 400°C, pressure of 0.15–0.25 MPa, with 80% conversion per pass, at about 70% selectivity to maleic anhydride. Vanadyl pyrophosphate, $(VO)_2P_2O_7$, often bound with TiO_2, SiO_2, or Al_2O_3, (with the latter two supports used for abrasive service in fluid beds) is the only material found thus far with high enough selectivity to be used commercially for butane oxidation. Multi-tubular fixed bed, fluid bed, and transport reactors have all been used for this service. The inlet gas to a multi-tubular reactor is kept below the LEL, about 1.8% butane, while the excellent heat control and thermal mass of the circulating catalyst in the fluid bed design allows for operation in the explosive regime at close to the stoichiometric O_2 to butane ratio (about 5.5% butane). As a result, fluid bed processes have much smaller equipment in the reactor loop, and are favored for new installations.

Two alternative recovery and purification are practiced commercially. The reactor effluent is further cooled by generating lower pressure steam, and then is absorbed into either an organic solvent or water. If water is chosen as the absorbent, a substantial amount of the maleic anhydride may be hydrolyzed to maleic acid. The maleic acid/anhydride mixture must then be converted back. A common method for water removal is heterogeneous azeotropic distillation with an entrainer such as xylene. Maleic acid is prone to thermal isomerization to fumaric acid, a high melting (286°C) by-product. Thus, residence time at the high temperatures in the dehydrating column base must be kept to a minimum. The resulting concentrated maleic anhydride solution is then distilled to remove lights and traces of solvent, and then taken overhead to remove heavies.

When a high-boiling, water-immiscible solvent, such as dibutyl hexahydrophthalate, is used for absorption, much less maleic anhydride is converted into the acid. No separate dehydration step is required, resulting in much lower energy requirements. Maleic anhydride is stripped overhead from the rich solvent stream from the absorber. This crude anhydride is distilled to separate light and heavy impurities. A slipstream of lean recycle solvent is treated to eliminate any heavy by-product that may build up.

The transport process, commercialized in 1996, was revolutionary in its use of separate butane oxidation and catalyst reoxidation zones. The butane was oxidized by the oxygen lattice of the catalyst, which was then regenerated with air in a separate vessel. Air and butane were never mixed. However, an unexpectedly high make rate of fumaric acid in the transport reactor system caused severe operational difficulties and poor plant availability. These problems led to the permanent shut down and dismantling of the one and only commercialized transport reactor process in 2004 [173].

Formaldehyde [174–179]

The production of formaldehyde from methanol illustrates that many different process alternatives, such as choice of catalyst, conversion per pass, separation selection, and safety considerations, can lead to a commercially practiced, economically viable process. With the mixed metal oxide catalyst (Fe/Mo, often with Cr promoter), formaldehyde is produced via the endothermic oxidative dehydrogenation of methanol:

$$CH_3OH + 1/2\ O_2 \longrightarrow H_2C{=}O + H_2O \quad \Delta H_{rxn} = -157 kJ/gmole$$

The additional pathway of equilibrium-limited, endothermic dehydrogenation occurs simultaneously in a ratio of 40% dehydrogenation, 60% oxidative dehydrogenation on silver catalysts:

$$CH_3OH \rightleftharpoons H_2C{=}O + H_2 \qquad \Delta H_{rxn} = +85\ kJ/gmole$$

In a typical silver catalyst process, a mixture of air–steam–methanol is reacted adiabatically over a thin layer (2.5–3.0 cm thick) of silver gauze or crystals at slightly above atmospheric pressure and high temperature. The formaldehyde-rich effluent is cooled initially by steam generation, and formaldehyde is recovered from the gas in a water-fed circulating quench absorber. The hydrogen-rich absorber tail gas is burned for steam generation. Both complete and partial methanol conversion processes are practiced commercially. In the complete conversion process, the feed gas comprises sufficient methanol to be above the upper explosive limit (25–27 mole% methanol), along with 46–54 mole% air, and 20–30 mole% steam. Conversion is greater than 99% at reaction temperature of 600–720°C, at a selectivity of 89–91%. A product of 40–55 wt% formaldehyde with <1 wt% methanol can be produced without further distillation. In the partial conversion process a mixture of 40–45% methanol, 20–25% air, and 20–35% steam is adiabatically converted to formaldehyde at 550–650°C. Methanol conversion is 70–80%, with a selectivity of 90–92%. Distillation is required to concentrate the product formaldehyde and remove unreacted methanol for recycle. Some vendors offer as high as 68 wt% formaldehyde with less than 1 wt% methanol.

The mixed metal oxide process operates below the LEL with 9 mole% methanol in air diluted with recycled off-gas and steam. The methanol is converted in a non-isothermal, multi-tubular fixed bed reactor at 350–390°C, where the heat of reaction is controlled by shell-side generation of high pressure steam. Formaldehyde is recovered by absorption into sufficient water to produce a solution of 50 wt% formaldehyde, with less than 1 wt% methanol without further distillation. Conversion is greater than 99%, at a selectivity of 93–96%. Purged gas does not contain sufficient combustibles and must be incinerated or catalytically oxidized with supplemental fuel. Formaldehyde is susceptible to oligomerization in the liquid phase and must be kept warm if concentrated, and is often stabilized with the addition of a small amount of methanol [180, 181].

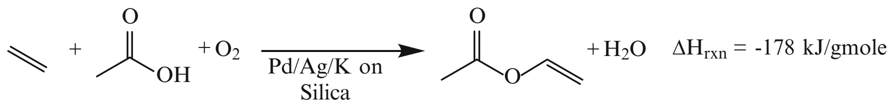
Esters via Oxidative Acetylation of Olefins, VAM [182–186]

Many routes for the production of the important monomer vinyl acetate (VAM), the unsaturated two-carbon ester of acetic, have been developed. These include:

1. Addition of acetic acid to acetylene form VAM directly (once a significant commercial process)
2. Reaction of acetic anhydride with acetaldehyde to form EDA, with subsequent cleavage to VAM and acetic acid (also once a significant commercial approach)
3. Carbonylation of methyl acetate with CO and H_2 to form EDA, with subsequent cleavage (never commercialized fully)
4. A vapor or liquid phase reaction of acetic acid, ethylene, and oxygen to form VAM and water

The latter route, as a gas phase reaction dominates commercial practice. In the typical industrial process acetic acid is vaporized and mixed with a fresh and recycle stream of ethylene. The combined gas is then mixed with high purity oxygen in a highly reinforced and shielded mixing unit. The feed gas to the reactor typically comprises 10–20 mol% acetic acid, 10–30 mol% CO_2, 10 mol% N_2, Ar or other inerts, 50 mol% ethylene, and less than 8 mol% oxygen. This composition is above the upper explosive limit.

The exothermic reaction occurs in a multitubular fixed bed reactor over a silica supported Pd/Ag/K catalyst at 140–180°C, 0.5–1.2 MPa. Steam is generated on the reactor shell side for heat removal and control of the hot spot temperature. By-products are water, carbon dioxide, and small amounts of EDA, ethyl acetate, acetaldehyde, methyl acetate, and glycol acetates.

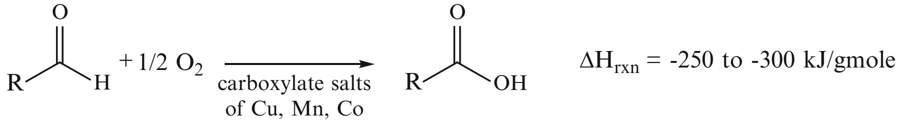


The reactor effluent is interchanged with the incoming feed gases and fed to a predehydration column, where about 50% of the water contained therein is removed by azeotropic distillation using the hot incoming gases for boil-up. The water/VAM azeotrope is taken overhead, decanted and VAM refluxed to the column. The crude VAM in the bottoms stream (20–40% VAM, 6–10 wt% water, the rest acetic acid and by-products) is dehydrated further by a similar water/VAM heterogeneous azeotropic distillation, with the VAM phase refluxed. The aqueous phase is further steam stripped to recover VAM for recycle to the first column and wastewater discharged as underflow. Final purification occurs in a third column, with the product VAM taken overhead, ethyl acetate as a sidedraw, and acetic acid and heavies (EDA, glycol acetates) as the bottoms. The acetic acid is recycled, with a heavy purge taken from the sludge of the acetic acid evaporator.

The ethylene-rich off-gas from the predehydrator is largely recycled, but a fraction is purged to prevent build-up of argon and nitrogen, and another portion is scrubbed with water to remove acetic acid prior to contacting with hot potassium carbonate system for carbon dioxide removal. Per pass conversion of ethylene, acetic acid, and oxygen range from 8–15%, 15–35%, and up to 90%, respectively. Selectivities of 94–96% on ethylene and greater than 98% on acetic acid are achievable.

Oxidation of Aldehydes to Acids [187]

The principle industrial route to C_3-C_{10} carboxylic acids is by the homogeneous liquid phase oxidation of the corresponding aldehyde, due to their wide availability from hydroformylation.



Examples of acids produced at large scale by aldehyde oxidation are propionic, *n*-butyric, isobutyric, *n*-valeric, isovaleric, and 2-ethylhexanoic acid. This chemistry typically is not sufficiently selective for use with unsaturated aldehydes (e.g., acrolein, methacrolein). Other methods are used for these substrates (see above, this section).

The oxidation may be carried out with oxygen or more typically with air, usually without an additional solvent, and with or without a catalyst present. Oxidation of a branched-chain aldehyde, exemplified by 2-ethylhexanal, is extremely sluggish and non-selective without a catalyst, while straight chain moieties can be oxidized successfully without a catalyst. The most effective catalysts are primarily transition metal ions that can transfer only one electron, and have two valence states of equal stability to allow efficient oxidation/reduction, such as Cu^{I}/Cu^{II}, Co^{II}/Co^{III}, and Mn^{II}/Mn^{III}. The metal is most conveniently introduced into the reaction system as the carboxylic acid salt of either acetic acid (commonly available) or as the salt of the acid to be produced. Although these oxidations typically give greater than 90–95% selectivity, common types of by-products include formate esters, ketones, and $n-1$ hydrocarbons and alcohols.

A transition metal-catalyzed oxidation reaction proceeds through a number of identifiable steps: (1) a hydrogen atom is abstracted from an aldehyde molecule by a metal ion to form a carbonyl radical, with reduction of the oxidation state of the metal; (2) dioxygen adds to the carbonyl radical to form a percarbonyl radical; (3) the percarbonyl reacts with another aldehyde molecule to form a peracid and regenerates a carbonyl radical (chain propagation); (4) the peracid reacts with another aldehyde molecule to form two acid molecules; (5) the metal is returned to its original oxidation state via reaction with a peracid, with formation of an acid molecule. The catalyst acts as an initiator for rapid generation of a high flux of carbonyl radicals, but also can inhibit the reaction if at too high a concentration (too many radicals, not enough free aldehyde to accomplish step 3). Typically there is an optimal catalyst concentration that balances radical chain formation and decomposition. Without a catalyst, there is often a lengthy induction period, before sufficient radials are generated to kick off the chain mechanism.

In a generic industrial process illustrated in Fig. 10.25, air, the aldehyde substrate, and the catalyst are fed to a bubble column reactor or CSTR (**1**) at 50–150°C, close to atmospheric pressure. Conversion of aldehyde is 80–95%, depending on the substrate. The off-gas from the reactor is cooled to remove condensable species, separated into vapor and liquid phases (**2**), and off-gases are scrubbed to remove traces of organics before emission to the atmosphere. Temperature, air, and aldehyde flow are adjusted to ensure that the exit gas is outside of the explosive region.

Liquid product from the reactor, containing unreacted aldehyde, product acid, catalyst, and a range of high, medium, and low boiling by-products is distilled (**3**), with materials boiling lighter than the product acid taken overhead and acid, catalyst, heavies as bottoms product. Depending on the by-product slate (some oxidations produce formate esters and ketones boiling closely to the aldehyde), the distillate of the crude column may be further treated in one or two columns to remove lights (**4**), medium boilers (**5**), and to recycle unreacted aldehyde. The crude acid is taken overhead in the product column (**6**), with concentrated catalyst underflow for recycle, less a purge for heavy ends.

Carbonylation

Chemistry

Carbonylation is a catalytic reaction between carbon monoxide (CO) and an organic substrate to form a new carbon–carbon bond with the introduction of an oxygen-containing functional group into the product. Technically any reaction involving CO alone or in tandem with another molecule, e.g., H_2 or H_2O is a carbonylation. However, historically reactions involving the insertion of CO and H_2 commonly have been known as hydroformylation or oxo reactions. This important class of carbonylations is treated separately in the section, “Hydroformylation.” In this section, the types of carbonylation reactions highlighted are:

1. Liquid phase reaction of CO with methanol or methyl acetate catalyzed by homogeneous Rh or Ir complexes to form acetic acid or anhydride
2. Liquid phase reaction of CO with methanol catalyzed by a strong base to form methyl formate
3. Liquid phase reaction of CO and ethylene with methanol over a strong base catalyst to form methyl propionate
4. Liquid phase reaction of CO, water, and olefins catalyzed by a strong acid to form neo acids, also known as Koch carbonylation or hydrocarboxylation

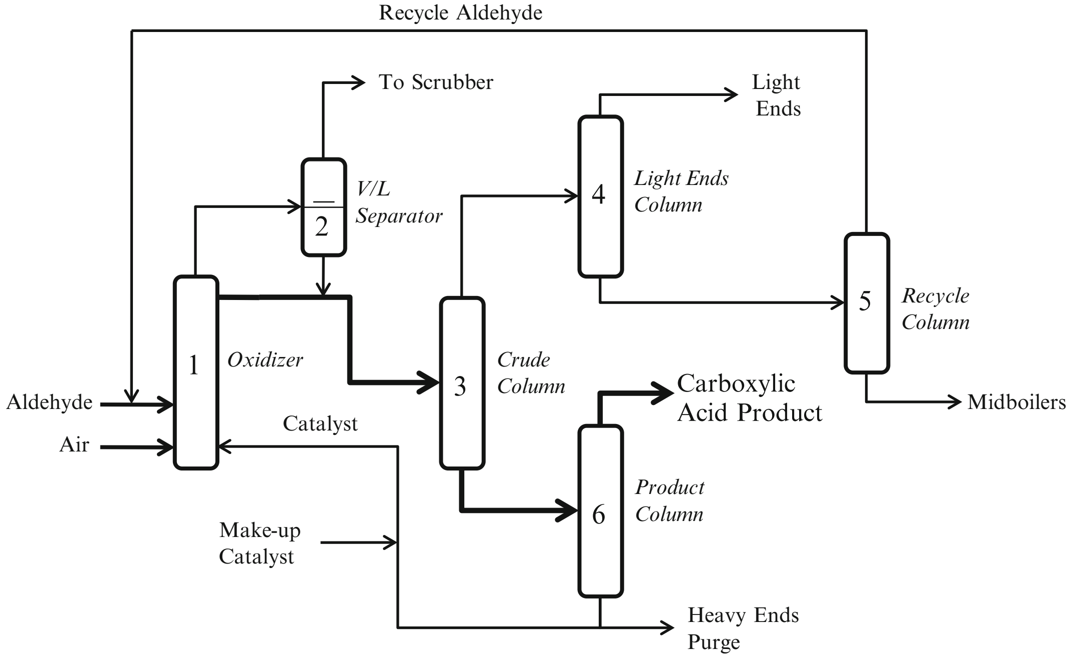
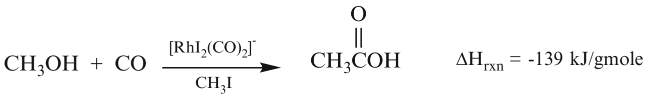


Fig. 10.25 Flowsheet for oxidation of aldehydes to carboxylic acids

Acetic Acid [188–193]

The carbonylation of methanol has largely displaced earlier synthetic routes to acetic acid such as partial oxidation of hydrocarbons, fermentation of carbohydrates, and oxidation of acetaldehyde derived from ethylene. Methanol carbonylation, a net insertion of carbon monoxide into methanol, has been known since 1913 and practiced commercially since the 1950s, initially as a high temperature, high pressure process (250°C, 70 MPa) using cobalt iodide as a catalyst. The breakthrough for methanol carbonylation occurred in the late 1960s with the discovery of rhodium catalysts promoted with methyl iodide, allowing operation at much lower temperatures and pressures and much higher productivity and yield than the original cobalt systems.



$$CH_3OH + CO \xrightarrow[CH_3I]{[RhI_2(CO)_2]^-} CH_3\overset{\overset{\displaystyle O}{\|}}{C}OH \qquad \Delta H_{rxn} = -139\ \text{kJ/gmole}$$

Under reaction conditions, the active catalyst is the $[RhI_2(CO)_2]^-$ anion. Oxidative addition of methyl iodide (formed from HI and methanol) $[RhI_2(CO)_2]^-$ followed by insertion of CO into the Rh-CH_3 bond and coordination of CO yields $[CH_3C(O)RhI_3(CO)_2]^-$. Reductive elimination of acetyl iodide (CH_3COI) liberates the original $[RhI_2(CO)_2]^-$ complex, and rapid hydration of acetyl iodide produces acetic acid and regenerates HI. The reaction is zero order in methanol and CO and first order in methyl iodide and rhodium. The concentration of methyl iodide in the reactor is substantial, around 20–25 wt%. With the original Rh—methyl iodide system, high water concentration (14–15 wt%) is required for high activity and for Rh stability. This high water level also promotes the water-gas shift reaction, resulting in byproduct CO_2 and H_2. Propionic acid is the major heavy byproduct, formed primarily via carbonylation of ethanol resulting from small quantities of acetaldehyde formed by acetic acid reduction or homologation of methanol.

Table 10.30 Commercial methanol carbonylation systems

Active catalyst species	Promoters	Reactor water content (wt%)
$[RhI_2(CO)_2]^-$ complex	CH_3I	14–15
$[RhI_2(CO)_2]^-$ complex	CH_3I, LiI	4–5
$[RhI_2(CO)_2]^-$ complex, bound to polyvinyl pyridine resin	CH_3I	3–7
$[IrI_2(CO)_2]^-$ complex	CH_3I, Ru or Re salts	<8

Many catalytic advances have been made since the 1970s, enhancing selectivity, productivity, and reducing energy usage. The four basic catalytic systems used in practice today are summarized in Table 10.30. The latter three allow high selectivity and productivity at much lower water content in the reactor loop than the original Rh-methyl iodide system. The water-rich end of acetic acid–water vapor–liquid equilibrium is fairly pinched, the heat of vaporization of water is high, and all the water present in the reactor loop must be distilled overhead, so low water content translates into substantial energy and capital savings.

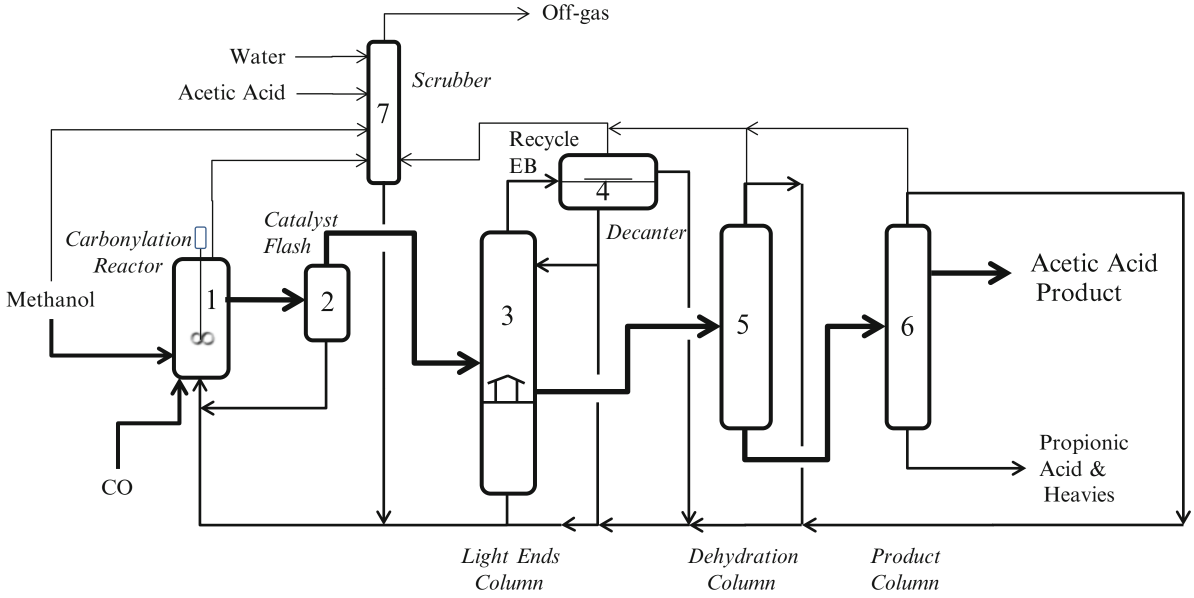


Fig. 10.26 Flowsheet for carbonylation of methanol to acetic acid

The addition of LiI greatly enhances rates by assisting in formation of methyl iodide from methanol and stabilizes the rhodium complex under low water conditions. Use of a heterogeneous rhodium catalyst, addresses many of the drawbacks of the original homogeneous Rh system. The rhodium is immobilized in a highly stable state on a polyvinyl pyridine resin. The bound, catalytically active $[RhI_2(CO)_2]^-$ remains stable at low water levels, and without LiI addition. The reaction is accomplished in a three phase gas lift reactor. By-product propionic acid, CO_2 and H_2 production are reduced.

Iridium-based carbonylation has been commercialized fairly recently by BP (the Cativa process), and is yet a third method to reduce water content. The rate of methyl iodide addition to the active $[IrI_2(CO)_2]^-$ complex is about 150 times faster than to $[RhI_2(CO)_2]^-$. Insertion of CO to form the iridium acetyl species is the slow step, but is accelerated by addition of Re or Ru promoters. Substantial Ir concentrations are required for good rates, but low water operation is possible with lower methyl iodide concentrations and without LiI addition. With Ir, production of propionic acid, CO_2 and H_2 is reduced, but methane formation is higher.

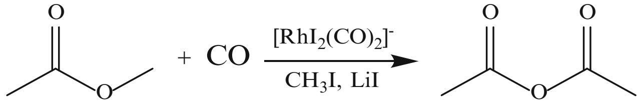
A flow sheet for a typical modern acetic acid plant is shown in Fig. 10.26. High purity carbon monoxide, methanol, catalyst, and promoters (methyl iodide and others), are introduced into a CSTR or gas lift (thermosyphon) reactor system (**1**) operated at (150–200°C, 2–4 MPa). The exothermic heat of reaction is controlled by flashing of the reactor effluent to produce a crude acetic acid vapor product and a catalyst residue for recycle to the reactor (**2**). Noncondensable byproducts, (methane, hydrogen, carbon dioxide) are vented from the reactor loop to control the CO partial pressure. These off-gases must be thoroughly scrubbed, typically by a staged absorption system of methanol, acetic acid, water to ensure that no iodine containing species are released (**7**).

The crude acetic acid from the flash unit is sent to the light ends column (**3**) where methyl iodide, water, and methyl acetate are taken overhead as a two phase mixture (**4**) for recycle to the reactor. Bottoms from this column are recycled to the reactor and wet acetic acid is taken as a vapor side draw above the feed. The wet acid is dehydrated in another column (**5**), with dry acid as underflow, and water/acetic acid distillate for recycle. The product column (**6**) produces glacial acetic acid as a sidedraw high in the column. Overhead material is recycled and by-product propionic acid and heavies are purged as underflow. Overall selectivity on methanol is typically 99% and greater than 90% on carbon monoxide.

Although highly selective and productive, Rh and Ir are very expensive, requiring millions of dollars for catalyst inventory. Moreover, the processes require high metallurgy (many parts Hastelloy) and appropriate systems to handle volatile iodide species.

Acetic Anhydride [194–196]

Acetic anhydride is an important acetylating agent for the production of such materials as cellulose acetate esters, ibuprofen and other pharmaceuticals, and hard-to-esterify esters. The predominant route to acetic anhydride involves the thermal cracking of acetic acid to ketene and water, followed by the reactive absorption of ketene into acetic acid (see the section, “Ketene”). However since 1983, a significant fraction of acetic anhydride (about 20% worldwide, over 50% in the USA) is produced via the carbonylation of methyl acetate:



ΔH$_{rxn}$ = -95 kJ/gmole

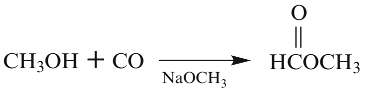
The carbonylation of methyl acetate is similar in many respects to the carbonylation of methanol described above, but differs in some key aspects:

1. Methanol carbonylation is done in an aqueous medium. Anhydride production necessitates anhydrous conditions, with significantly reduced Rh stability/solubility at reaction conditions.
2. Hydrogen addition is required to generate and maintain the active catalyst form. Unlike in the aqueous methanol system, no water is available for in situ hydrogen generation via water-gas shift.
3. LiI promoter is absolutely required to generate the active Rh catalyst species and to activate the alkyl methyl group of methyl acetate under anhydrous conditions.
4. Methyl acetate carbonylation is equilibrium-limited, with conversion of methyl acetate between 50 and 75% vs. essentially 100% in methanol carbonylation.
5. The by-product slate is different. EDA, "tars," acetone, and CO_2 are the primary by-products. Propionic acid is not formed to any extent, nor is methane.
6. Production of "tars," primarily derived from by-product EDA, is significantly higher under methyl acetate carbonylation conditions. These tars must be purged and tend to bind Rh more effectively than acetic acid residue. A much more complex tar purge/Rh recovery system is required.

In the anhydride production process, methyl acetate, recycle methyl iodide, catalyst, CO, a small amount of H_2, and LiI promoter are fed to a back-mixed reactor system at 160–190°C, 2–5 MPa to produce crude anhydride. Unreacted CO and other purge gases are scrubbed after leaving the reactor to recover acetyl and iodide moieties. The liquid draw off from the reactor is expanded and flash distilled under low CO/H_2 pressure to prevent catalyst decomposition. The flash bottoms are largely recycled, but a purge is taken for tar removal and Rhodium recovery. The crude anhydride flash distillate is purified in a three column sequence. In the first, methyl iodide and methyl acetate are distilled for recycle to the reactor. In the second, acetic acid id distilled overhead for recycle to the methyl acetate production process. The methyl acetate feedstock can be produced via reactive distillation, as shown in the section, "Methyl Acetate." The final column produces 99% acetic anhydride product as a distillate, with tars and heavies as the bottoms stream.

Methyl Formate [197]

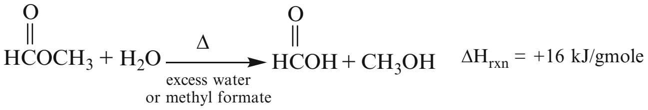
The equilibrium for the direct synthesis of formic acid by the addition of water to carbon monoxide is quite unfavorable and has not been exploited to produce commercial quantities of formic acid. Instead, formic acid is produced by acidification of formate salts (30% of global production) or hydrolysis of methyl formate (70% of global production). The acidification route involves reacting sodium or calcium formate (largely generated as a by-product of aldol/cannizzaro reactions, see the section, "Neopentyl Polyhydric Alcohols") with a strong acid, usually sulfuric acid, to yield dilute free formic acid and sodium or calcium sulfate. The second route begins with carbonylation of methanol to give methyl formate in a liquid phase reaction using a dilute strong base catalyst (e.g., sodium methoxide) at a temperature around 80°C and pressure of 4–5 MPa:



$$CH_3OH + CO \xrightarrow[NaOCH_3]{} H\overset{O}{\overset{\|}{C}}OCH_3 \qquad \Delta H_{rxn} = -29\text{kJ/gmole}$$

This carbonylation is equilibrium-limited, resulting in 30–50% methanol and 95% CO conversion per pass. Sodium methoxide can react with methyl formate to give sodium sulfate and dimethyl ether. Dry conditions are required to limit hydrolysis of methyl formate to sodium formate and methanol, and to prevent leveling of the strong alkoxide base to the catalytically ineffective weaker hydroxide ion. The crude methyl formate is flashed to atmospheric pressure and distilled. The low-boiling methyl formate/methanol azeotrope is taken overhead for hydrolysis and the underflow is recycled to the carbonylation step, with purge for removal of sodium formate by crystallization.

The equilibrium-limited hydrolysis of methyl formate is accomplished in both two-stage (1-1 water-methyl formate molar ratio, followed by addition of excess methyl formate—30% conversion/pass) and one-stage processes (5-1 water-methyl formate molar ratio, 60% conversion/pass):

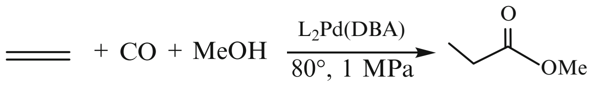


$$H\overset{\overset{O}{\|}}{C}OCH_3 + H_2O \xrightarrow[\text{excess water or methyl formate}]{\Delta} H\overset{\overset{O}{\|}}{C}OH + CH_3OH \qquad \Delta H_{rxn} = +16 \text{ kJ/gmole}$$

Unreacted methyl formate and methanol are distilled from the crude hydrolysis product, leaving a dilute water-formic acid mixture as underflow. Production of concentrated formic acid from this underflow product presents particular problems due to the high-boiling water/formic acid azeotrope. A number of separation schemes are practiced commercially; the most common approach is pressure-swing distillation. Excess water is distilled overhead in a first higher pressure column to produce azeotropic formic acid/water as a bottoms product. This azeotrope composition is distilled in a second vacuum column to produce concentrated formic acid (90–98% purity) as distillate, and the low pressure azeotrope composition is recycled to the first column [198]. Energy usage is improved with the one-stage hydrolysis process if excess water is removed by extraction (with a secondary amide or other solvents), prior to pressure-swing distillation.

Methyl Propionate [199–201]

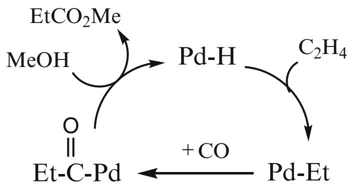
When ethylene and carbon monoxide are reacted in methanol, in the presence of a palladium diphosphine catalyst, either polyketone or methyl propionate may result depending on the identity of the diphosphine ligand. Lucite, formerly a subsidiary of ICI, has developed this carbomethoxylation as a step of their "Alpha Process" for the production of methyl methacrylate. In the Alpha Process, methyl propionate is reacted with formaldehyde to ultimately produce MMA. The rate and selectivity of this carbonylation catalysis are extraordinarily high under relatively mild conditions.



L_2=1,2-bis(di-t-butylphosphinomethyl)benzene
DBA=dibenzylidene acetone

selectivity = 99.8%
TON = 50,000/h

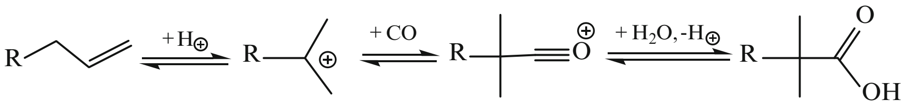
Mechanistic studies have established the catalytic cycle depicted below with the key intermediates being a palladium hydride, which reacts with ethylene to produce a palladium ethyl species that rapidly inserts a carbon monoxide yielding the palladium propionyl complex. This complex is methanolyzed to produce methyl propionate product and regenerate the palladium hydride intermediate.



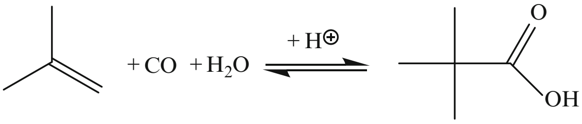
Shell developed very similar catalysis for the production of their Carilon polyketone in which the diphosphine ligand in bis(diphenylphosphino)propane rather than the 1,2-bis(di-t-butylphosphinomethyl)benzene employed by Lucite, illustrating the exquisite sensitivity of this catalysis to the structural details of the palladium diphosphine catalyst. A commercial scale MMA plant (Jurong Island, Singapore) based on the Alpha process has operated at design capacity since 2008.

Koch Acids

Koch carbonylation (also called hydrocarboxylation or the Koch-Haaf reaction) is the synthesis of predominantly tertiary (neo) carboxylic acids from olefins, CO, and water using a strong acid catalyst, typically H_3PO_4/BF_3 or an acidic zeolite. In the initial step of the reaction, the olefin is protonated, generally with isomerization and rearrangement, to form the most stable (i.e., tertiary) carbonium ion. Carbon monoxide then adds to the carbonium ion to form an acylium cation. Further reaction with water and regeneration of the proton results in the formation of the tertiary carboxylic acid.



An example of this chemistry is conversion of mixed butenes, especially isobutylene to pivalic (neopentanoic) acid.

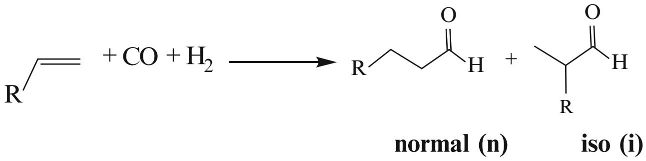


In the typical industrial process, CO is contacted with the olefin and H_3PO_4/BF_3 catalyst in a CSTR at 20–80°C, 2–10 MPa. Conversion of the olefin is generally very high with 80–100% selectivity to the neo acid. Main by-products are neo acids of the oligomerized feedstock olefins. Product neo-acids are washed with water and decanted for removal of the catalyst acids. C_5 to C_{11} neo acids are made industrially form the corresponding C_4-C_{10} olefins. The higher olefinic feedstocks are typically derived from ethylene oligomerization or dehydration of paraffinic fractions.

Hydroformylation [202–207]

Chemistry

Hydroformylation, commonly known as oxo synthesis, is a C–C bond-forming reaction of an olefin with syngas (~1:1 M ratio of CO:H_2) to produce aldehydes. When the olefin is a terminal, or α-olefin, two products may be formed and are referred to as the normal (straight chain) and iso (branched chain) products.



The reaction is employed in practice with C_2 to C_{14} olefins, but hydroformylation of propylene is the primary application. The oxo aldehydes are valuable and reactive intermediates for a number of secondary transformations, including hydrogenation to alcohols, oxidation to acids, and aldol condensation to higher aldehydes and ketones, with subsequent further transformations of these products to esters, anhydrides, ketenes, glycol ethers, and sulfonates to name a few.

Process conditions vary significantly with olefin substrate and catalyst employed, but typically fall in the range of 1–20 MPa and 40–200°C. All current commercial catalysts are based on either Co or Rh transition metal complexes. The first commercial catalyst, $HCo(CO)_4$, discovered by Otto Roelen in Germany in 1938, required high pressures (20–35 MPa) for acceptable operation. Subsequent work demonstrated that phosphine ligands could, allowed practical operation at much lower pressures 1–10 MPa and could further hydrogenate the product aldehydes to alcohols (Shell process). Another quantum leap in productivity and selectivity occurred with the introduction of Rh-based catalyst complexes in the early-1970s, with operating pressures as low as 1–2 MPa. Although Rh is roughly 1,000 times more expensive than Co, it is also several orders of magnitude more active than Co, for a substantial net benefit. Low pressure Rh operation reduces syngas compression energy over fivefold and significantly lowers capital costs over that of the original cobalt hydrido carbonyl catalyst system. The formation of an aldehyde via Rh-catalyzed oxo synthesis involves the following catalytic steps:

1. An olefin coordinates to a coordinatively unsaturated metal hydrido carbonyl complex, exemplified by $HCo(CO)_3$ and $HRh(CO)(PPh_3)_2$, followed by insertion of the olefin into the metal hydride bond to give an alkyl complex.
2. A second migratory insertion of CO converts the alkyl into an acyl group.
3. Oxidative addition of H_2 yielding a metal dihydride.
4. Reductive elimination of the aldehyde product regenerates the coordinatively unsaturated metal hydrido carbonyl complex (back to step 1).

A simplified expression for the rate of aldehyde formation illustrates some of the kinetic tradeoffs: [208, 209]

$$\text{Rate of aldehyde formation} = k\frac{[\text{olefin}][\text{metal (Co or Rh)}]p_{H_2}}{p_{co}}$$

For lower olefins, especially ethylene and propylene, the concentration of the olefin in the liquid reaction phase is proportional to pressure, higher pressure gives higher solubility. Likewise, high hydrogen partial pressures favor aldehyde formation, as does increasing catalyst concentration (for Rh, this approach is expensive). At some point hydrogenation of the olefin to paraffin or loss of selectivity by hydrogenation of the product aldehyde may also occur. However, reducing CO partial pressure enhances the rate, but a minimum CO pressure is needed to maintain stability of the metal catalyst complexes, as well as for product formation. In practice, a complex tradeoff of these effects is required, with low pressure, low catalyst concentration, and about 1:1 H_2/CO molar ratio tending toward optimal.

A key consideration in hydroformylation of α-olefins is the n:i ratio of products for C_3 and higher olefin substrates. For example, with propylene, the C–C bond may form to give a linear (n), or a branched product (i):

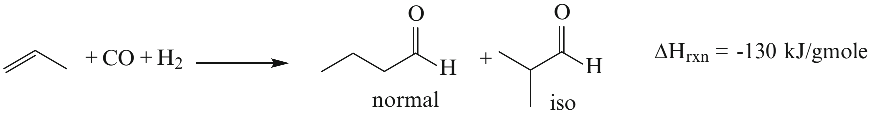


Table 10.31 Hydroformylation systems

Descriptor	Unmodified cobalt	Cobalt ligand	Rh, no ligand	Rh, classic low P	Water soluble	High normal	Higher Iso
Catalyst metal	CO	CO	Rh	Rh	Rh	Rh	Rh
Ligand	No ligand	Triphenyl phosphine	No ligand	Triphenyl phosphine	Triphenylphosphine sodium sulfonate	Bisphosphite	Halophosphite
Catalyst structure (Fig. 10.27)	I	II	III	IV	V	VI	VII
Feedstock	Branched med-long chain,α-olefin	C_7-C_{14} internal/α-olefin	C_6-C_{14} internal/α-olefin	C_3-C_5 α-olefin	Propylene	C_3-C_4 + α/internal olefin	C_2-C_{10} α-olefin
Conditions							
Temperature, °C	110–180	150–190	95–180	85–90	110–130	85–115	80–120
Pressure, Mpa	20–35	4–8	20–30	1.5–2.0	4–6	1.5–2.0	1.5–2.0
Results							
Aldehyde selectivity	Medium	Low	High	High	High	Very high	High
Normal/iso ratio	3:1 to 4:1	6:1 to 7.5:1	1:1 to 1.3:1	10:1 to 12:1	19:1 to 20:1	22:1 to 32:1	1:1 to 3:1
Heavy ends	Some, 4–6%	Little	Little	Little	Little, 0.2–0.8%	Little	Little
Metal cost	Low	Low	High	High	High	High	High
Ligand cost	None	Low	None	Low	High	High	Medium
Practitioner/licenser	No active licensers, but still used	Shell	Johnson Matthey/Davy	Dow/Davy	Ruhrchemie/Rhône-Poulenc	Dow/Davy	Eastman Chemical Co.

Linear aldehydes are more valuable intermediates for production of plasticizers and detergents (linear molecules are more biodegradable). Significant research effort has been expended over the last 80 years developing high n:i ratio catalysts. Formation of straight chain olefins is favored by bulky, electron-rich ligands. The vast majority of companies practicing oxo synthesis use catalyst systems that preferentially produce normal as opposed to iso aldehyde products. A company uses proprietary rhodium—halophosphite catalysts that produce an unusually low n:i product mix.

The chain-length of the feedstock olefin and desired n:i ratio of the aldehyde product largely dictate which metal–ligand catalyst combination is more desirable. Table 10.31 summarizes operating conditions and results for a number of commercial systems, with highlights about each given below. Corresponding catalyst structures are presented in Fig. 10.27.

Catalyst I. $HCo(CO)_4$: The first generation cobalt carbonyl/hydrocarbonyl catalyst, originally developed in the 1930s, required a high pressure, 20–35 MPa, for good activity. It is not used commercially for lower olefins any longer, but finds some applications with higher olefins.

Catalyst II. Cobalt, phospine ligand: The second-generation phosphine-modified cobalt catalyst, developed in the 1960s, significantly reduced the required operating pressure, improved the n:i ratio to between 6:1 to 9:1, and gives mostly terminal hydroformylation, even when fed internal olefins (by promoting double bond migration). This process is used mostly for production of detergent range alcohols, where the greater propensity of cobalt to hydrogenate the oxo aldehyde is not a disadvantage. Trialkyl phosphines are preferred and derivatives of the phobanes depicted in Fig. 10.27 are used commercially by Shell.

Catalyst III. HRh(CO)4: A rhodium carbonyl hydride catalyst without a phosphine or phosphite is currently used for hydroformylation of C_6 to C_{14} linear and branched olefins. Without expensive ligands to recover, a unique adsorption/ion exchange method is used for efficient catalyst recovery. Isononanol and isodecanol are important products.

Catalyst IV. Rhodium, triphenyl phosphine ligand: The phosphine-modified rhodium catalyst was developed in the 1970s and has been successfully employed since, largely for propylene and butene hydroformylation. This system dramatically reduced required operating pressures, while giving several orders of magnitude rate improvement and higher n:i ratios. Several variations are practiced; (1) vapor stripped reactor for product removal and catalyst separation (product taken as vapor directly from reactor, (2) liquid take-off with flash separation of aldehyde and catalyst.

Catalyst V. Rhodium, water soluble phosphine: The phosphine-modified rhodium catalyst was made water-soluble by adding sulfonate groups to the phenyl rings. In this way a simple separation of catalyst for recycle is facilitated by decantation of the catalyst-bearing aqueous phase from the product-laden organic phase. This low pressure process is used commercially for hydroformylation of propylene to high-normal butyraldehyde n;i ratio of about 10:1 to 20:1.

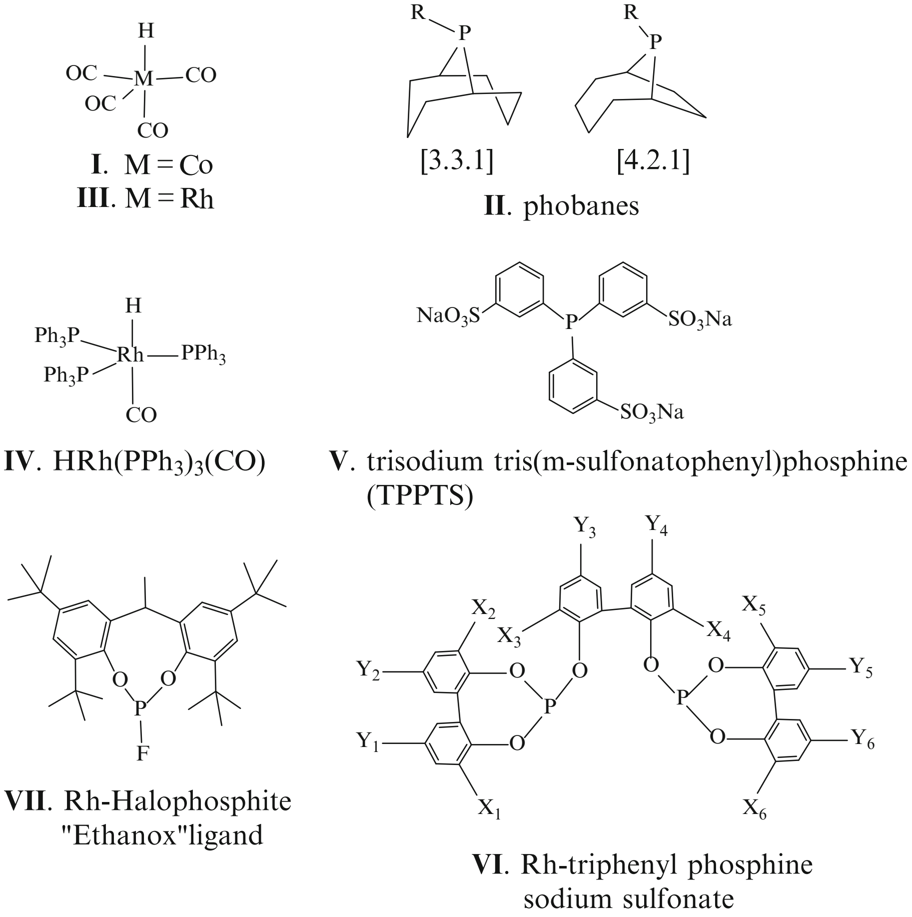


Fig. 10.27 Hydroformylation catalyst structures

I. M = Co
III. M = Rh

[3.3.1] [4.2.1]
II. phobanes

IV. HRh(PPh$_3$)$_3$(CO)

V. trisodium tris(m-sulfonatophenyl)phosphine (TPPTS)

VII. Rh-Halophosphite "Ethanox"ligand

VI. Rh-triphenyl phosphine sodium sulfonate

Catalyst VI. Rhodium, *bisphosphite ligand*: The family of bisphosphite ligands dramatically increases the n:i ratio to about 30:1, the highest level yet achievable. Both propylene and butene feedstocks are usable.

Catalyst VII. Rhodium, *tridentate halophosphite ligands*: Unlike the other six catalyst systems, tridentate halophosphite ligands produce the lowest commercialized n/i ratio, typically less than 3:1. Ethylene and propylene are common feedstocks. With these halophosphite ligand, the n:i butyraldehyde ratio dependson the ligand to rhodium molar ratios, the reaction temperature, and the carbon monoxide partial pressure [210].

More detailed descriptions of typical oxo processes for short and long chain olefins are given below.

Lower C$_2$-C$_5$ Olefins

In a typical commercial process for catalyst types V, VI, and VII, a mixture of 1:1 syngas and lower olefins (can be a mixture) are fed to a stirred tank or bubble column reactor along with fresh and recycle homogeneous catalyst solution, where hydroformylation takes place. Temperature and pressure conditions depend on the metal-ligand catalyst chosen, as well as the chain length of the olefin. Heat is removed by the reactor jacket, an external cooling loop, or alternatively, boiling of the product aldehyde. If a vapor-stripped reactor system is used, the vaporous effluent and unreacted gases from the reactor are cooled, allowed to condense, and separated into gas and liquid streams. If conversion of olefin is low, the off-gases are compressed and recycled, with some amount purged to prevent buildup of inerts (i.e., alkanes in the olefin feed). If the aldehyde is the final end product, the crude aldehyde is distilled to remove water and high boilers (usually alcohols and esters), and possibly further fractionated to separate iso and normal isomers. The phosphine or phosphite ligands tend to degrade at a low, but finite rate, so it is necessary to purge some amount of the catalyst solution. Subsequent recovery of Rh, typically involves distillation, evaporation, decantation, or extraction. Decantation is used, for example, with the water soluble ligand system described in Table 10.31.

Higher Oxo: Detergent Alcohols [211]

In one commercial process, linear and branched C$_6$-C$_{14}$ olefins derived from light olefin oligomerization or paraffin dehydrogenation, are reacted in the liquid phase with 1:1 syngas at 95–180°C, about 25 MPa, using a homogenous rhodium carbonyl catalyst (without additional ligands) to form isomeric oxo aldehydes. Depending on the reactivity of the olefin feedstock (branched < internal olefin < α-olefin), conversion can reach 95%, but is typically limited to 80% for a first pass. The less reactive olefins are removed and sent to a second pass with more severe conditions and higher catalyst loading, with 75% second pass conversion (95% overall conversion for two pass). Mild conditions favor linear, terminal olefin synthesis, while more forcing conditions tend to lead to isomerization of double bonds, yielding higher levels of internal olefins and thus branched aldehydes products.

The pressure of the reactor effluent is let down to remove dissolved CO and hydrogen. This off-gas is recycled, with a purge to prevent build-up of inerts. Since there is no expensive ligand to recover, Rh ions are adsorbed from the degassed olefin/aldehyde liquid product by passing through a series of ion exchange resin beds. The ion exchange bed is periodically incinerated, converting the Rh metal to Rh oxide, which is readily recycled to the Rh salt catalyst precursor. The crude catalyst-free aldehyde is typically sent directly to hydrogenation for conversion into alcohols before distillation. See the section, "Aldehydes and Ketones to Alcohols," for information on the hydrogenation of the higher aldehydes to alcohols.

Chlorination [212, 213]

Chemistry

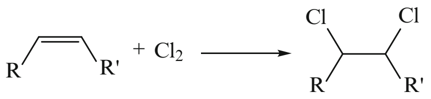
As the name implies, chlorination is a family of reactions in which one or more chlorine atoms are incorporated into a molecule. For large scale commercial synthesis the most common chlorine sources are Cl_2, HCl, and HOCl (hypochlorous acid). Chlorinated molecules are extremely important to many industries, including construction, automotive, toys (vinyl chloride, PVC), silicon products (methyl chloride), dry cleaning and solvent applications, and air conditioning (HFC's). Some rank among the largest volume chemicals produced. Several basic types of chlorination reactions are employed industrially:

1. *Substitution* is the abstraction and replacement of a hydrogen atom by a chlorine atom (from Cl_2 gas), with HCl as the co-product. The reaction involves a free radical chain mechanism initiated by thermal dissociation of Cl_2, with a chloride radical as the chain carrier. Selectivity is generally poor, as all hydrogen atoms on a molecule are more or less available for substitution in the order tertiary > secondary > primary, saturate > unsaturated > aromatic. Further information on relative reactivity of various substrates is available in standard texts [214]. The substitution reaction can be extremely fast and exothermic, with a heat of reaction on the order of −90 to −135 kJ/gmole of Cl added. Less reactive substrates require either high temperature or a Lewis acid catalyst. Alkane and aromatic substrates are used:

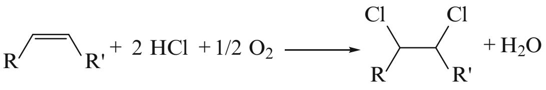
$$C_nH_{2n+2}+mCl_2 \longrightarrow C_nH_{2n+2-m}Cl_m+m\ HCl \qquad 1\leq m\leq 2n+2$$

$$ArH+Cl_2 \longrightarrow ArCl + HCl$$

2. *Addition* involves the insertion of Cl_2 across an unsaturated (double or triple) bond in a hydrocarbon molecule. Mild temperatures, 70–120°C, are required. The addition is highly specific; chlorine atoms are symmetrically distributed, and selectivities above 90% are common. The substitution reaction is exothermic, with a heat of reaction on the order of −180 to −230 kJ/gmole of Cl_2 added. Lewic acids such as $FeCl_3$, $ZnCl_2$, and PCl_3 are common catalysts.



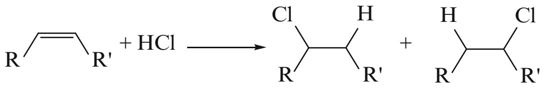
3. *Oxychlorination* is the reaction of HCl and oxygen with a substrate to produce a chlorinated species and water. Reaction temperature can range from 80 to over 200°C, depending on substrate. The reaction is generally selective (>90%). The heat of reaction is very high, around −250 to −300 kJ/gmole. Common catalysts are $CuCl_2$ and $ZnCl_2$.



4. *Hydrochlorination* is the reaction of HCl with a substrate to produce a mono-chlorinated species. If the substrate is an alcohol, then the by-product is water:

$$R\text{-}OH + HCl \longrightarrow R\text{-}Cl + H_2O$$

The reaction with an alcohol is generally very selective (>90%), with the diether as the major by-product. The heat of reaction, around −30 to −50 kJ/gmole, is relatively low compared to other types of chlorinations. HCl may also be added across a double bond:

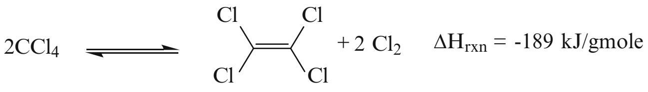


5. *Perchlorination* (sometimes called chlorinolysis) is the high temperature (400–700°C) exhaustive chlorination of a multi-carbon substrate with concomitant breaking of at

least one C–C bond to form smaller fully chlorinated fragments. Both hydrocarbon and partially substituted chlorohydrocarbons are potential feedstocks. At full conversion and complete fragmentation to C_1 pieces, the reaction stoichiometry is:

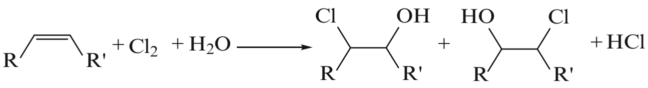
$$C_nH_{2n}Cl_2 + (3n\text{-}1)\ Cl_2 \longrightarrow n\ CCl_4 + 2nHCl \qquad n > 1$$

At high temperature (>600°C), an additional equilibrium exists between carbon tetrachloride and perchloroethylene.



Large excesses of Cl_2 or high pressure are needed to drive to mostly single carbon fragments. Perchlorination is an important component of an integrated chlorination facility, taking in low value chlorinated by-products from other processes, to produce exhaustively chlorinated, lower carbon number materials. Any aliphatic, aromatic, or oxygenated hydrocarbon may be reacted via perchlorination. Oxygenates produce phosgene as a by-product.

6. *Chlorohydrination* is the reaction of an olefin with Cl_2 and water to form a vicinal hydroxy chloride moiety (chlorohydrin), with HCl as a co-product.



A large excess of water and low olefin concentration is required to prevent by-product formation by addition and substitution reactions. When $R \neq R'$, isomers differing in the position of the hydroxyl and Cl may be formed. Chlorohydrins are important intermediates in the synthesis of epoxides, especially propylene oxide.

Hydrochlorination and oxychlorination reactions are often more economical uses of HCl generated by substitution and addition reactions than disposal of excess HCl as a gaseous or dilute aqueous co-product. Examples of some industrially important chlorination reactions are summarized in Table 10.32. The types of reactions discussed above may also be applied to other common halogens, with reactivity in the order F > Cl > Br > I. Fluorine is so reactive and exothermic that fluorinations are not normally done directly, but by indirect methods exemplified by halogen substitution [215].

Chloromethanes [216–218]

The chloromethanes (one-carbon molecules containing one to four chlorines), are primarily used as intermediates for other chemistries: methyl chloride for silicones production; methylene chloride for solvents; chloroform for hydrochlorofluorocarbon production; and carbon tetrachloride for chlorofluorocarbons. The more substituted chloromethanes are made predominately by successive substitutions of methyl chloride or perchlorination of C_1-C_3 chlorinated streams. The use of carbon disulfide as a feedstock for methyl chloride synthesis is essentially obsolete, and thermal substitution of methane is falling out of favor. The flowsheet for a typical integrated chloromethanes process is shown in Fig. 10.28. Methyl chloride is now largely produced via gas phase hydrochlorination of methanol with excess HCl at 280–350°C, 0.3–0.6 MPa, in a multitubular steam-generating fixed bed reactor filled with γ-alumina catalyst (**1**).

$$CH_3\text{-}OH + HCl \xrightarrow[\gamma\text{-alumina}]{} CH_3\text{-}Cl + H_2O \qquad \Delta H_{rxn} = -33\ \text{kJ/gmole}$$

An older process using 70% aqueous zinc chloride/HCl at 130–150°C is not used much anymore. The reactor effluent is cooled by direct contact quench with recycle ~30 wt% aqueous HCl (**2**). Diluted ~20 wt% aqueous HCl exits as underflow from the quench tower, and gaseous crude methyl chloride is scrubbed further with caustic (**3**), and then concentrated (96 wt%) H_2SO_4 to remove dimethyl ether (which reacts to form methyl sulfate) and water (**4**). The diluted H_2SO_4 (~80 wt%) is a by-product. Dry methyl chloride is compressed (**5**) to around 2 MPa, condensed, and distilled overhead as a pure product, or used along with the higher chlorinated by-products as feedstock for further substitution chlorination.

In order to produce a slate of chloromethanes, methyl chloride and by-products from the hydrochlorination step are further thermally substituted with gaseous chlorine at 350–400°C, 0.8–1.5 MPa, in an adiabatic reactor (**6**), usually without catalyst. If the reaction mixture is allowed to reach about 450°C or above, rapid pyrolysis may occur, with the very undesirable formation of carbon soot. The reactor effluent is quenched (**7**), separated from co-product HCl by multistage condensation (**8**), stripped of residual HCl (**9**), and distilled to take mono-, di-, tri- and tetra-chloromethanes overhead in successive columns (**10**), (**11**), (**12**), (**13**) respectively. Depending on the desired product distribution, any partially substituted product may be recycled to the substitution reactor for further chlorination. Overall yield on methanol is at least 95% and greater than 98% on chlorine.

The heavy chlorinated residue from the last column and other chlorination by-products (e.g., 1,2-dichloropropane from propylene oxide production, vinyl chloride residues) may be fed to a perchlorination unit for production of smaller chain fully substituted materials. A typical high pressure tubular perchlorination unit operates adiabatically at up to about 620°C, 10–20 MPa, with a Cl_2/substrate

Table 10.32 Chlorination products

Product	Common industrial method[a]	Feed substrate	Chlorine source	Co-product	Heat of reaction, kJ/gmole
Methyl chloride	H	Methanol	HCl	Water	−33.0
Methyl chloride	S (less common)	Methane	Cl_2	HCl	−103.5
Methylene chloride	S	Methyl chloride	Cl_2	HCl	−102.5
Chloroform	S	Methylene chloride	Cl_2	HCl	−99.2
Carbon tetrachloride	S or P	Chloroform	Cl_2	HCl	−94.8
Carbon tetrachloride	P	Chlorinated C_1-C_3 species	Cl_2	HCl	Exothermic, depends on feed substrate
Chlorobenzene	S	Benzene	Cl_2	HCl	−131.5
o-, *p*-, *m*-Dichlorobenzene	S	Monochlorobenzene	Cl_2	HCl	−124.4
1,2-Dichloroethane	A	Ethylene	Cl_2	None	−180
1,2-Dichloroethane	O	Ethylene	HCl/O_2	H_2O	−240
Vinyl chloride monomer	Cr	1,2-Dichloroethane	Not applicable	Cl_2	+71
Allyl chloride	S	Propylene	Cl_2	HCl	−113
1,3-Dichloropropan-2-ol 2,3-Dichloropropan-1-ol	C	Allyl chloride	HOCl	None	
Propylene chlorohydrin	C	Propylene	Cl_2/H_2O	HCl	

[a] *S* substitution; *A* addition; *O* oxychlorination; *H* hydrochlorination; *P* perchlorination; *C* chlorohydrination; *Cr* cracking

ratio dependent on the desired perchloroethylene/carbon tetrachloride product mix. The effluent, comprising carbon tetrachloride, HCl, Cl_2, perchloroethylene, and small amounts of hexachlorobenzene (recyclable to extinction), is quenched with cold carbon tetrachloride, flashed to remove Cl_2 and HCl, and distilled to produce carbon tetrachloride and perchloroethylene as pure streams. The reactor must be Ni clad to prevent corrosion. With the decline in applications for carbon tetrachloride, the production of most perchlorination units are slated heavily to perchloroethylene by a ratio of roughly 3-1.

Chloroaromatics [219–221]

A wide variety of chlorinated aromatics, with and without additional functional groups on the aromatic ring, may be produced by substitution chlorination, with co-production of HCl. These include mono- and di-chlorobenzenes, chlorotoluene, benzyl chloride. The product mix ultimately obtained is highly dependent on Cl_2/aromatic molar ratio, temperature, catalyst composition, co-catalyst (if any), location of other groups already on the benzene ring, and contacting mode. High temperature and high Cl_2/aromatic ratio favor multiple chlorinations. Backmixed (CSTR) operation leads to increased levels of multi-chlorinated products; batch or reactive distillation approaches tend to increase monochloride content. The heats of reaction for mono- and dichlorobenzenes are summarized in Table 10.32. A typical continuous process for mono and di-chlorination of benzene operates at 20–80°C, with dissolved ferric chloride catalyst/organic sulfide co-catalyst, and with very low water content (25–35 ppm) to prevent catalyst deactivation. The resulting effluent is neutralized with caustic, or in more modern plants distilled directly to separate HCl and Cl_2 from aromatics. Unreacted benzene is recovered overhead first, with monochlorobenzene taken overhead in a second column. Dichlorobenzenes are close boiling and generally must be separated by crystallization to obtain high purity single isomers. Product distribution between mono- and dichlorobenzenes is highly dependent on the ratio of benzene to chlorine fed to the reactor.

1,2-Dichloroethane (Ethylene Dichloride) and Vinyl Chloride Monomer (VCM) [222–225]

Ethylene dichloride (EDC), and its primary end product vinyl chloride monomer (VCM), are produced at truly immense scale, constituting about 8% of global output of chemicals. EDC and VCM were once produced primarily by chlorination of acetylene. Limited application for the by-product HCl posed a major restriction to the growth of EDC and VCM. However, major producers now use a combination of direct chlorination and oxychlorination of ethylene to produce EDC and VCM without significant coproduction of HCl. Figure 10.29 shows such an integrated

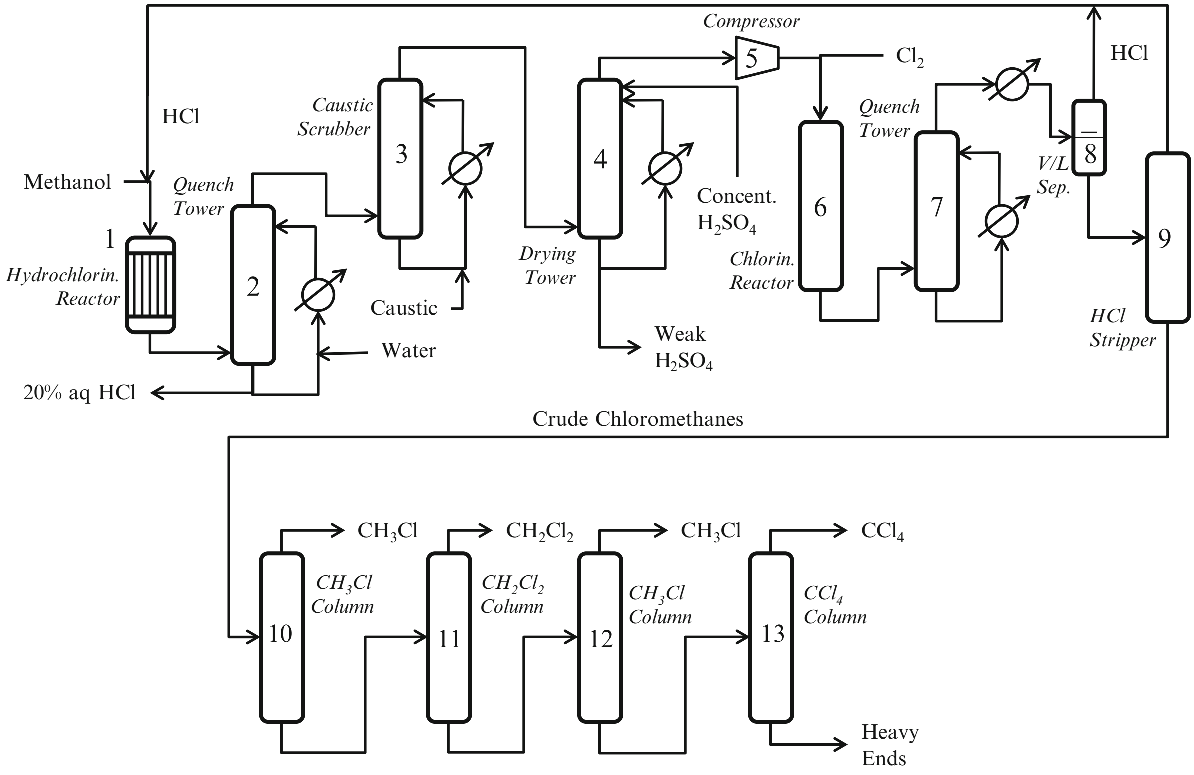
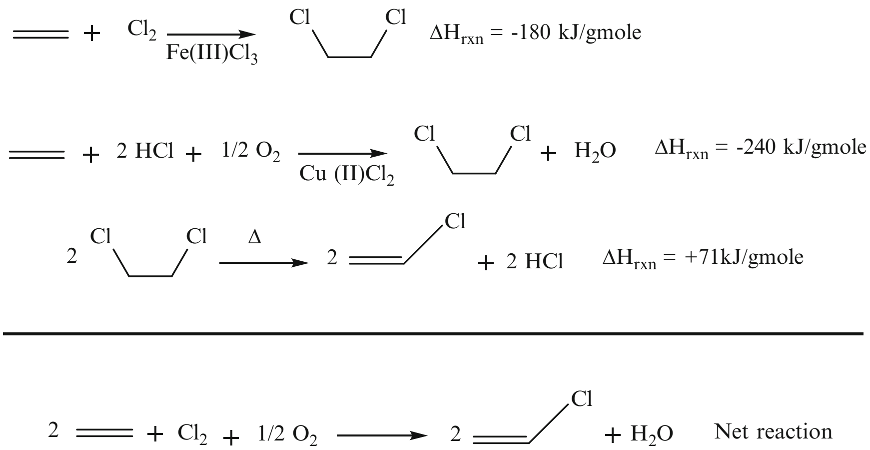


Fig. 10.28 Flowsheet for production of chloromethanes

plant for producing EDC and vinyl chloride from ethylene, chlorine, and air. The chemical reactions are as follows:



The oxychlorination step (**1**), catalyzed by Cu(II)Cl_2-impregnated alumina at 200°C and 0.2–0.5 MPa pressure, uses recycled HCl from the EDC pyrolysis step, with oxygen and ethylene to produce EDC and water. This reaction is even more exothermic (−290 kJ/gmole) than direct chlorination. Both multi-tubular fixed bed and fluidized bed reactor designs with steam generation are used commercially to control the heat release. The fluid bed design, although suffering from lower selectivity than a PFR due to back mixing, can be run advantageously in the explosive region. Conversion is typically 93–97% per pass, with a selectivity on ethylene around 95%. Major by products are heavy trichlorinated species, trichloroacetaldehyde, and oxygenated lights such as CO, carbon dioxide, formic acid.

The addition chlorination of ethylene is done usually in a liquid-phase reactor (**3**), catalyzed by the Lewis acid, iron(III)

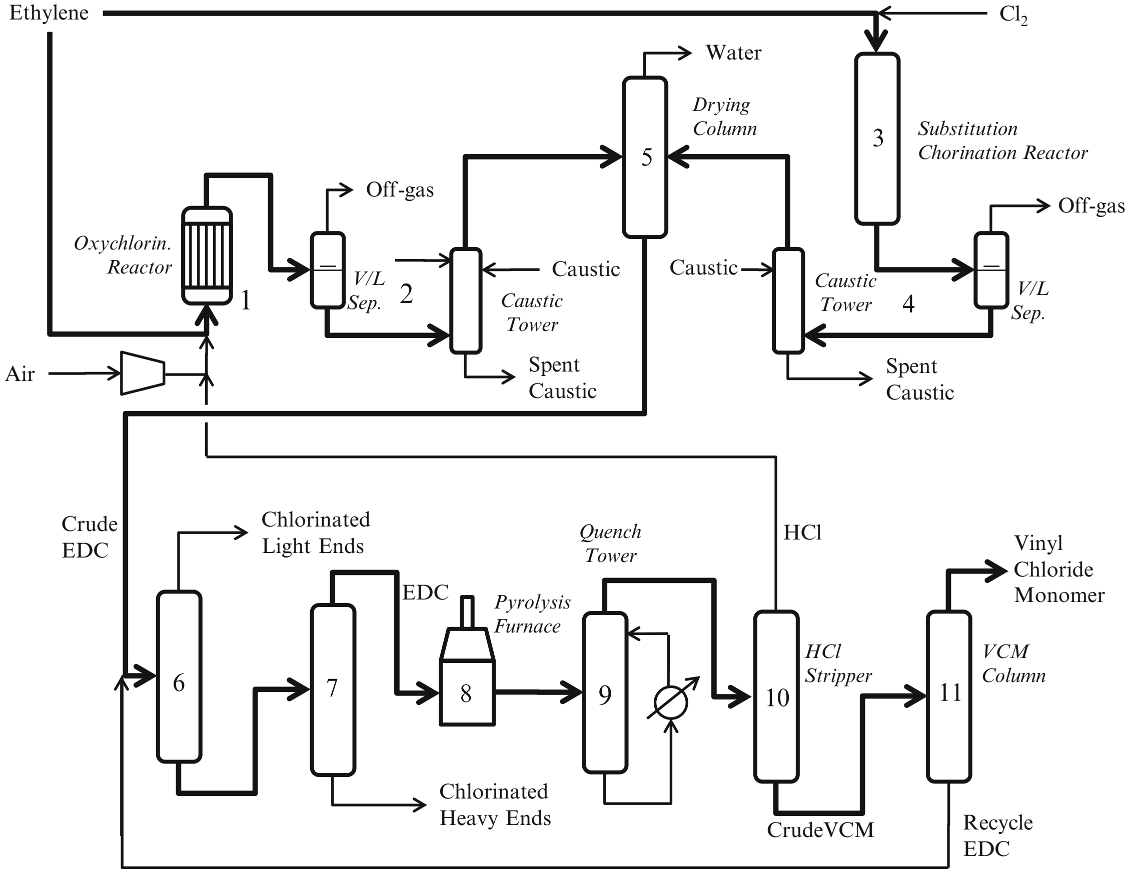


Fig. 10.29 Flowsheet for production of vinyl chloride monomer

chloride at 40–100°C, 0.4–0.5 MPa. Ethylene absorption is the rate limiting step. The heat of reaction (−180 kJ/gmole) can be removed either by heat exchange in the reactor, or more economically by running the reaction above 100°C and using the generated heat to distill the EDC. High-purity ethylene, with low propylene content, is preferred to avoid formation of chlorinated C_3 by-products. Ethylene conversion approaches 100%, with a selectivity of 98% on ethylene and 99% on chlorine.

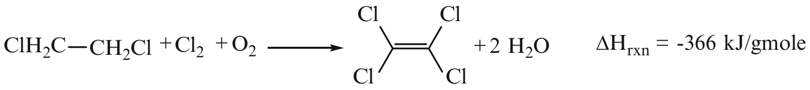
The effluents from reactors (**1**) and (**3**) are disengaged from vapors and scrubbed with caustic for additional removal of HCl (**2**),(**4**). The crude EDC is azeotropically distilled to remove residual water (**5**), and subjected to distillation to separate light (**6**) and heavy (**7**) chlorinated by products. These chlorinated species may be used in chlorinolysis or production of other halogenated species.

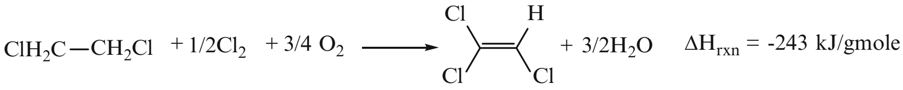
The cracking (pyrolysis) of EDC to VCM typically is carried out at temperatures of 500–550°C without a catalyst (**8**). Selectivity to VCM is greater than 98%. The hot gases are quenched (**9**) and distilled to strip out HCl (**10**) and then VCM (**11**). The unconverted EDC is returned to the EDC purification train.

About 35% of the chlorine produced globally goes to the manufacture of VCM. More than 95% of all VCM is used to produce polyvinyl chloride (PVC), an important polymer for the housing and automotive industries (See Chap. 15). The rest of the VCM goes into the production of chlorinated solvents and ethylenediamine.

Chloroethanes [226–228]

The commercial importance of the chlorinated ethane derivatives, perchloroethylene, 1,1,1-trichloroethane, and trichloroethylene, has declined significantly since the 1980s due to t the declining use of chlorofluorocarbons. Major uses of all three are as precursors for fluorocarbon production, dry cleaning agents, and solvents. With declining use of carbon tetrachloride, perchlorination (see Section 10.7.1) is becoming a more significant route to perchloroethylene. 1,1,1-trichloroethane is produced predominantly by a two-step process starting with vinyl chloride monomer (VCM): (1) hydrochlorination of VCM with HCl produce predominantly 1,1-dichloroethane; (2) thermal or photochemical (more selective) substitution chlorination of 1,1-dichloroethane to 1,1,1-trichloroethane. Perchloroethylene and trichloroethylene are also made by the oxychlorination of 1,2-dichloroethane (VCM intermediate, see above):

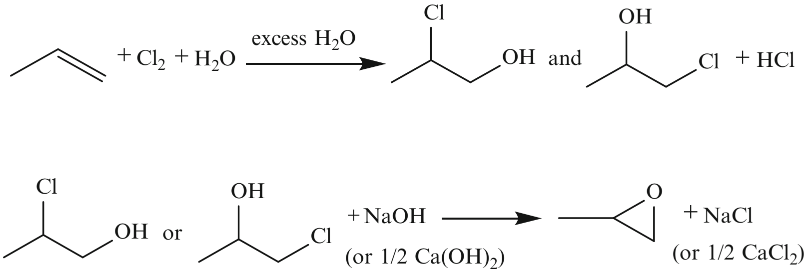




Propylene Oxide via Propylene Chlorohydrin [229, 230]

Although differing from ethylene oxide by only one extra carbon in the backbone, propylene oxide (PO) has not been made successfully at commercial scale by the direct oxidation of propylene. The presence of allylic hydrogens dramatically and adversely affects the efficacy of silver-based catalysts for the selective oxidation of propylene to PO with oxygen [231]. Instead all world scale PO plants are based either on indirect oxidation of propylene by either the co-product hydroperoxide routes (50% of total, see the section, "Hydroperoxide Co-product Processes for Propylene Oxide") or the chlorohydrin route described herein.

The chlorohydrin route entails two main steps: (1) synthesis of propylene chlorohydrins (PCH) by the reaction of propylene, chlorine, and water (chlorohydrination) and (2) synthesis of PO via the reaction of PCH with an alkali or alkaline earth hydroxide (dehydrochlorination).



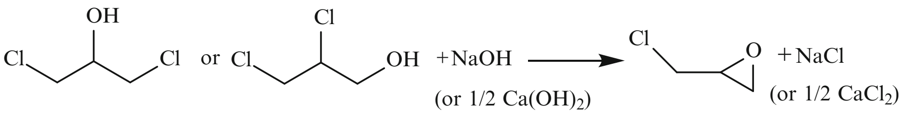
In the first step chemical or polymer grade propylene, gaseous chlorine reacts in equimolar amounts in the presence of excess water at 45–90°C, 0.11–0.19 MPa in bubble column or CSTR, to produce a solution of about 4 wt% PCH isomers, roughly a mole of HCl per mole of PCH, and other chlorinated by-products. Excess water (around a 50:1 mole ratio of water to propylene) is critical to suppress undesired substitution and addition reactions, but 1,2-dichloropropane (DCP), 1,3-dichloropropan-2-ol or 2,3-dichloropropan-1-ol (DCH), 2,2′-dichlorodiisopropyl ether (DICPE) are major by-products. Yields are typically 88–96% PCH isomers, 3–10% DCP, 0.3–1.2% DCH, 0.2–0.8% DCIPE, and 1% of others including allyl chloride. The PCH solution is disengaged from the vent gases (subsequently scrubbed with caustic and thermally oxidized), and conveyed to the dehydrochlorination reactor.

The PCH solution is reacted with slightly more than two equivalents of base (either NaOH or $Ca(OH)_2$) to neutralize the HCl produced in the first step and to saponify PCH to propylene oxide and salt. The reactions are rapid and exothermic. The crude PO is distilled in a two column sequence, with lights taken overhead in the first column and PO distilled overhead in the second column. DCP, DCIPE, and epichlorohydrin are underflowed in the second column. Almost all streams in the chlorhydrin process are extremely corrosive, requiring expensive materials of construction such as Teflon coatings, fiber reinforced plastics, graphite, rubber, brick-linings, Inconel, Hastelloy, or Monel alloys.

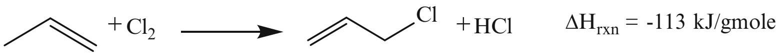
This process produces roughly 2 kg of either NaCl or $CaCl_2$ per kg of PO. Economics dictate that the PO unit be associated with a world-scale chloralkali, mercury, or membrane chlorine processes for regeneration of the chlorine and base. 1,2-Butylene oxide (and 1,2-butanediol upon hydrolysis) is made from 1-butene at a substantially smaller scale by a similar chlorohydration/dehydrochlorination route.

Epichlorohydrin [232, 233]

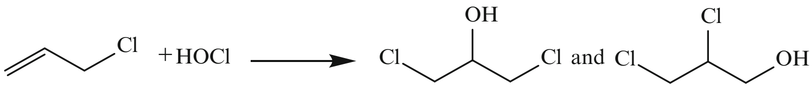
Epichlorhydrin, an important intermediate for coatings and urethane applications, is produced in 90% yield via dehydrochlorination of 1,3-dichloropropan-2-ol or 2,3-dichloropropan-1-ol (DCH).



The DCH may be a by-product of propylene oxide production (see above), or formed on purpose in a two-step process from propylene. In 1936, it was discovered that when propylene is reacted with chlorine at high temperature (500–510°C in practice), free radical substitution of an allyl hydrogen predominates over double bond addition, giving allyl chloride in 80–85% yield.



Excess propylene (4:1 M ratio) is required to ensure essentially complete conversion of chlorine, and for good selectivity. In the second step, allyl chloride is reacted with aqueous hypochlorous acid at 50–60°C to yield a mixture of DCH isomers.



Condensation

In its broadest sense, a condensation reaction involves the combination of two molecules or functional groups of the same molecule to form a new species, typically with the elimination of a small molecule such as water or an alcohol. In the present context, condensation reactions are limited to the coupling of two molecules via the formation of a new carbon–carbon bond, with elimination of water. The discussion to follow covers aldol condensations and the acid-catalyzed condensation of phenols and ketones, as exemplified by the synthesis of bisphenol A (BPA).

Aldol Condensation [234–236]

Chemistry

The aldol reaction is one of the most important industrial means of forming carbon–carbon bonds to directly synthesize complex functionalized molecules. A large number of commercially significant molecules are produced via aldol chemistry as shown in Table 10.33. Many of these compounds are intermediates for further transformations such as hydrogenation or oxidation.

In the aldol addition reaction, an enolizable ketone or aldehyde, i.e., one that has at least one acidic proton alpha to the carbonyl, reacts with another aldehyde or ketone to form a β-hydroxycarbonyl compound. The enolizable carbonyl acts as a nucleophile and other carbonyl as an accepting electrophile. The reaction is called a crossed aldol addition when the nucleophile and electrophile are different and a self aldol dimerization when they are the same.

Aldol reactions are catalyzed most commonly by bases such as alkali earth and alkaline earth hydroxides (e.g., NaOH or $Ca(OH)_2$), and amines (e.g., triethylamine, dimethylamine). Acid catalysis also occurs (and is often a source of by-product formation in acid catalyzed reaction where ketones or aldehydes are present), although is less effective for on-purpose production. The aldol addition reaction may proceed via two fundamentally different mechanisms, depending on whether an acid or base catalyst is used.

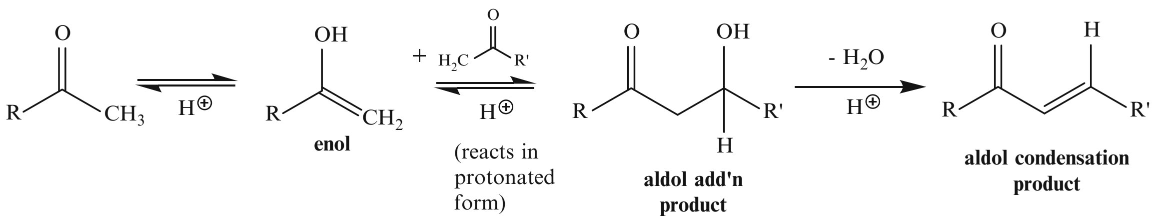
With an acid catalyst, the initial step in the reaction mechanism involves acid-catalyzed tautomerization of the carbonyl compound to the enol. The acid also serves to

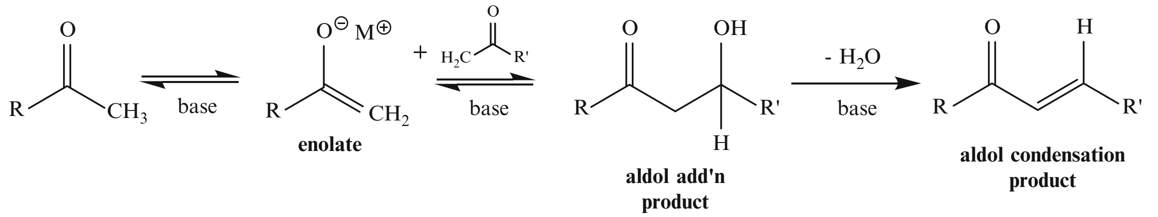
Table 10.33 Aldol reaction products

Enolate-forming carbonyl	Acceptor carbonyl	Chemistry	Product
n-Butyraldehyde	*n*-Butyraldehyde	SA, D, DH	2-Ethylhexaldehyde
n-Butyraldehyde	*n*-Butyraldehyde	SA, D, DH, CH	2-Ethylhexanol
Acetone	Acetone	SA	Diacetone alcohol
Acetone	Acetone	SA, D	Mesityl oxide
Acetone	Acetone	SA, D, DH	4-Methyl-2-pentanone
Acetaldehyde	Formaldehyde	CA, CH or CA, CZ	Pentaerythritol
Propionaldehyde	Formaldehyde	M	Methacrolein
Propionaldehyde	Formaldehyde	CA, CH or CA, CZ	Trimethylolethane (TME)
n-Butyraldehyde	Formaldehyde	CA, CH or CA, CZ	Trimethylolpropane (TMP)
i-Butyraldehyde	Formaldehyde	CA,CH or CA, CZ	Neopentyl glycol (NPG)
Acetaldehyde	Acetaldehyde	SA, D	Crotonaldehyde
Acetaldehyde	Acetaldehyde	SA, D, DH	*n*-Butyraldehyde
Acetaldehyde	Acetaldehyde	SA, D, DH, CH	*n*-Butanol
Acetaldehyde	Acetaldehyde	SA, CH	1,3-Butanediol
Acetaldehyde	Acetaldehyde	T	Ethyl acetate
i-Butyraldehyde	*i*-Butyraldehyde	T	2,2,4-Trimethyl-1,3-pentanediol (TMPD), isobutylisobutyrate, mono-/di-isobutyrate esters of TMPD
n-Valeraldehyde	*n*-Valeraldehyde	SA,D,DH,CH	2-Propylheptanol

SA self-aldol condensation; *CA* crossed-aldol condensation; *D* dehydration; *DH* double bond hydrogenation; *CH* carbonyl hydrogenation; *CZ* Canizzaro reaction; *T* Tishchenko reaction; *M* Mannich base condensation

activate the carbonyl group of another molecule by protonation, rendering it highly electrophilic. The enol is nucleophilic at the α-carbon, allowing it to attack the protonated carbonyl compound, leading to the aldol after deprotonation. The secondary or tertiary hydroxyl thus formed is highly favored to dehydrate to give the α,β-unsaturated carbonyl compound under acidic conditions. The combination of the aldol addition and dehydration steps is sometimes called an aldol condensation reaction.





Base-catalyzed aldol reactions begin as acid–base reactions and are thus extremely fast. An α-proton next to a carbonyl group is weakly acidic (pKa of 12–14). In the presence of a strong base, the α-carbon is deprotonated to a small extent to form an enolate ion. The coupling reaction occurs via nucleophilic attack by the resonance-stabilized enolate on the electrophilic carbonyl group of another molecule, which form a β-hydroxy carbonyl derivative (the aldol product) upon regeneration of the base catalyst. The aldol product may then undergo based-catalyzed dehydration to give the α,β-unsaturated carbonyl compound.

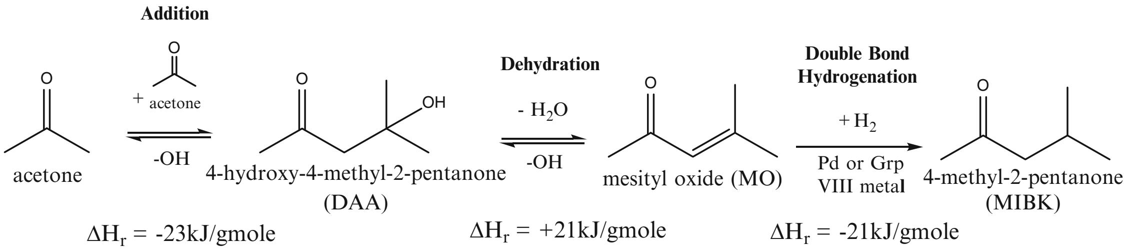
Ketones generally have a slightly lower pKa than aldehydes. Thus in a mixed ketone-aldehyde system, the ketone is favored to form the enolate and the aldehyde to act as the acceptor. Both steps are equilibrium limited, but the extent of the formation of addition and aldol dehydration products dictated by the structure of the reactants. For the following nucleophile-electrophile pairs the order of greatest shift to aldol products is:

$$\text{Aldehyde-Aldehyde} > \text{Ketone-Aldehyde} > \text{Ketone-Ketone}$$

Reactivity also diminishes with steric hindrance and with reduced solubility in the catalyst-containing phase (if the reaction is performed with a separate caustic-rich phase). A selection of industrially significant aldol-based processes is described in more detail below.

MIBK, Diacetone Alcohol (DAA), and Mesityl Oxide (MO) [237–239]

MIBK (4-methyl-2-pentanone), related intermediates, and its derivatives are the third largest end use for acetone behind methyl methacrylate and BPA. The synthesis of MIBK is a classic example of an aldol-based reaction pathway comprising self-aldol addition, dehydration, and hydrogenation:



This chemistry is practiced in both three-step and one-step process implementations. In the three-step process, acetone, dilute caustic is fed to a CSTR controlled at 10–20°C to enhance the equilibrium amount of DAA, but acetone conversion is typically only 8–10%. The reactor effluent is neutralized with phosphoric acid and stripped of unreacted acetone. Additional phosphoric acid is added to acidify the mixture to catalyze dehydration of DAA during distillation. Back-reacted acetone and the Mesityl oxide/water heterogeneous azeotrope are taken overhead and by-product salty water is underflowed. The organic layer from the overhead decanter is dehydrated and stripped of residual acetone in the next column. The dry underflow mesityl oxide (98–99 wt%) is hydrogenated over a Pd or Group VIII metal (e.g., Ni, Cu, Cr) in either a liquid phase reactor or in a low pressure hydrogenation/distillation column. One advantage of this process is that it allows for isolation of intermediates DAA and mesityl oxide.

The one-step process is much less capital intensive and generally allows for 30–50% per pass conversion of acetone. Acetone and hydrogen are fed to a packed bed reactor, slurry reactor, or catalytic distillation, typically at 130–140°C, 3–6 MPa, with an appropriate dual function catalyst combination to perform the aldol addition, dehydration, and hydrogenation reactions simultaneously. Many catalyst combinations have been reported in the open literature, including cation exchange resin impregnated with Pd, physically mixed exchange resin and supported Pd, zirconium phosphate-Pd, and niobium-Pd catalysts. Acetone, MIBK, and many of the reaction intermediates are present and active for further aldol and hydrogenation reactions, resulting in a very large reaction network. Typical yields of MIBK are 90–96% on acetone.

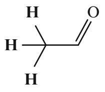
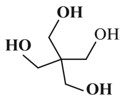
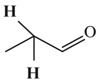
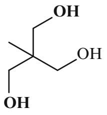
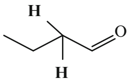
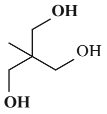
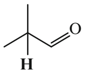
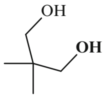
Neopentyl Polyhydric Alcohols [240, 241]

An important family of polyhydric alcohols with a neopentyl structure are pentaerythritol, trimethylolethane (TME), trimethylolpropane (TMP), and neopentyl glycol (NPG). These compounds are formed by the successive aldol addition reactions of formaldehyde with C_2 to C_4 aldehydes followed by reduction of the resulting hydroxy-aldehyde intermediate, as shown in Table 10.34. The aldol addition reactions are rapid, cannot be readily stopped at a lower number of formaldehyde additions, and do not lead to dehydration.

The typical process for pentaerythritol has not changed significantly for several decades. Aqueous formaldehyde, acetaldehyde, and aqueous $Ca(OH)_2$ or NaOH are fed to a CSTR maintained at 15–45°C. Three successive aldol additions of formaldehyde to acetaldehyde is followed by a Cannizzaro reaction (reduction of trimethylol acetlaldehyde by formaldehyde with formation of formic acid as a co-product) to form pentaerythritol. Overall the aldol-Cannizzarro sequence releases about 92 kJ/gmole of heat, necessitating effective reactor cooling. Although stoichiometry dictates the use of 4 moles of formaldehyde per mole of aldehyde, up to 16 moles per mole is added to limit formation of other acetaldehyde aldol and dipentaerythritol by-products.

Following neutralization of excess base with formic acid, unreacted formaldehyde is recovered by pressure distillation. The product polyol is recovered by stepwise vacuum concentration and fractional crystallization. If caustic is used as the base, an additional recrystallization from water is required to produce pentaerythritol free of sodium formate. On the other hand, calcium formate has very low solubility

Table 10.34 Neopentyl polyols

Name	Aldehyde feedstock	Feedstock structure	No. of α-hydrogens	Product structure
Pentaerythritol	Acetaldehyde		3	
Trimethylolethane (TME)	Propionaldehyde		2	
Trimethylolpropane (TMP)	*n*-Butyraldehyde		2	
Neopentyl glycol (NPG)	*i*-Butyraldehyde		1	

in water and can be precipitated prior to pentaerythritol concentration/crystallization. No effective methods for direct hydrogenation of trimethylol acetaldehyde to pentaerythritol have been commercialized. Yields are typically about 90% on acetaldehyde. Finding industrial uses for the formate salts limits the application of this chemistry.

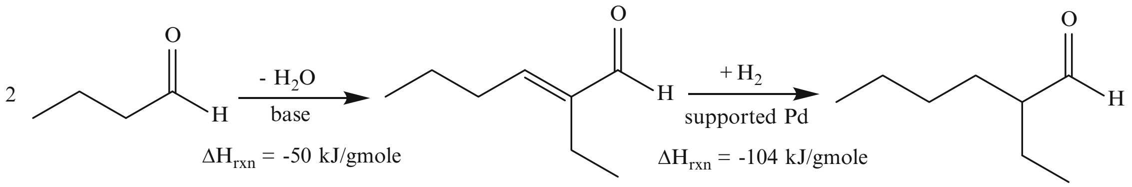
TME and TMP are produced commercially by very similar aldol-Cannazzaro processes, with the exception that the polyol often is extracted from the reaction mixture with a solvent, such as acetate esters, alcohols, or cyclohexanol, then purified by distillation. Yields are similar to pentaerythritol, reaching about 90%.

A significant portion of the neopentyl glycol produced commercially is synthesized by aldol addition with either sodium hydroxide or calcium hydroxide catalyst. The intermediate, hydroxypivaldehyde, can be finished via a Cannizzaro reaction, or more commonly by direct catalytic hydrogenation with Cu, Cu/Cr,Co, or Ni catalysts at 80–200°C, greater than 3.5 MPa pressure. Because the self-aldol of isobutyraldehyde requires a base stronger than hydroxide ion to occur to a significant extent (see section "Tischenko Reactions: Ethyl Acetate and 2,2,4-Trimethyl-1,3-Propanediol Derivatives" below), a relatively small excess of formaldehyde is sufficient for high conversion and a selectivity greater than 90%. Major by-products are 2,2,4-trimethyl-1,3-propanediol, neopentyl glycol isobutyrate, formic acid, formate salts, and NPG-hydroxypivalic acid esters. Purification involves extraction of the hydroxypivaldehyde intermediate with a solvent, followed by a series of distillations.

State-of-the-art NPG plants now use a tertiary amine catalyst in place of hydroxide catalysts. Complete conversion of formaldehyde occurs with a slight excess of isobutyraldehyde. The excess aldehyde is distilled with the amine and the two are recycled together. The reaction is quite selective and the most of acidic by-products produced by formaldehyde reduction are absent. The synthesis is finished with a similar catalytic hydrogenation. In both processes, hydroxypivaldehyde may be isolated and oxidized to the product hydroxypivalic acid.

2-Ethylhexanal [242, 243]

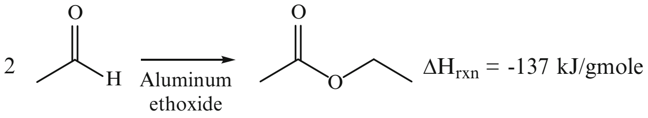
2-ethylhexanal is an extremely important intermediate for solvent and plasticizer applications. It is manufactured from *n*-butyraldehyde (see section "Hydroformylation") by aldol condensation in an alkaline medium or with a basic ion exchange resin at 80–130°C, 0.3–1.0 MPa. This is followed by catalytic hydrogenation of the double bond of the α,β-unsaturated aldehyde under mild conditions, 50–100°C, 0.2–2 MPa, to give 2-ethylhexanal. Supported Pd is a typical catalyst.



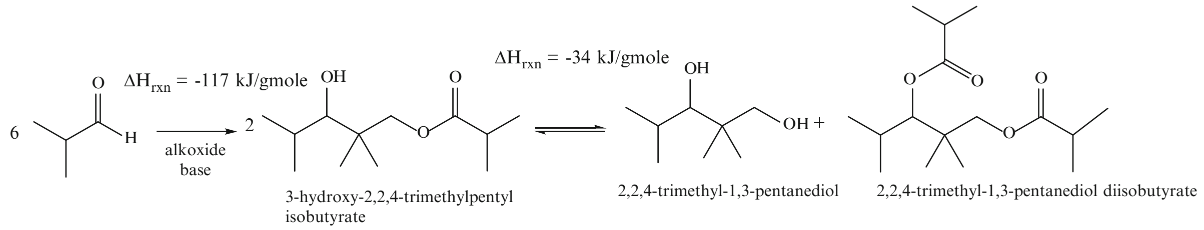
Conversion of *n*-butyraldehyde is high. The yield to 2-ethyhexanal are well above 90%, with major by-products of 2-ethylhexanol, and higher aldolized species. Under more severe conditions the α,β-unsaturated aldehyde intermediate can be converted directly to 2-ethylhexanol (see the section, "Hydrogenation"). Purification of 2-ethylhexanal occurs by distillation.

Tishchenko Reactions: Ethyl Acetate and 2,2,4-Trimethyl-1,3-Pentanediol Derivatives [244–246]

A closely related cousin of the aldol addition is the Tishchenko reaction occurring with aldehydes. In the presence of a base such as an aluminum alkoxide, one aldehyde molecule is reduced and the other is oxidized, with the direct formation of an ester. The largest industrial use of this reaction is in the formation of ethyl acetate from acetaldehyde:



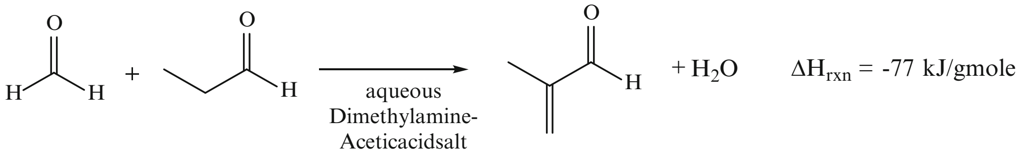
With a stronger alkali metal alkoxide base (e.g., sodium alkoxide), aldol addition is a competitive parallel reaction. An important application of the tandem Aldol-Tishchenko pathway is in the synthesis of 3-hydroxy-2,2,4-trimethylpentyl isobutyrate from isobutyraldehyde in the liquid phase using an alkali metal alkoxide salt as the catalyst.



A parallel Tishchenko reaction pathway produces isobutyl isobutyrate while equilibrium-limited transesterifications lead to 2,2,4-trimethyl-1,3-pentanediol and 2,2,4-trimethyl-1,3-pentanediol isobutyrate co-products. The products are separated by distillation.

Mannich Base Condensation: Methacrolein [247, 248]

The industrial use of a Mannich base (carboxylic or inorganic acid salt of a secondary amine) for aldol-type condensations is illustrated by the synthesis of methacrolein, the initial step in an alternate synthesis of methylacrylic acid. Propionaldehyde is condensed with aqueous formaldehyde in the liquid-phase in the presence of the Mannich base to form methacrolein and water:

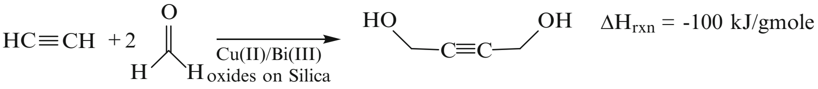


The best-performing base has been found to be a 40% aqueous solution of the dimethylamine-acetic acid salt. The synthesis methacrolein is remarkably selective (greater than 95%) and conversion of both formaldehyde and aldehyde are very high. Methacrolein is recovered by distillation from the reactor effluent as the methacrolein/water heterogeneous azeotrope, and further purified by distillation. The dimethylamine-acetic acid base is recycled to the reactor. The methacrolein thus produced can be oxidized in the same fashion as described in the section, "Acrolein/Acrylic Acid and Methacrolein/Methacrylic Acid."

Condensation of Formaldehyde with Acetylene: 1,4-Butanediol [249, 250]

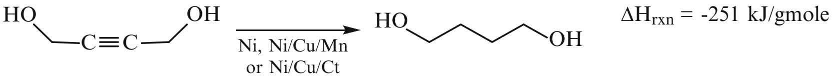
One of the major remaining applications of acetylene as a feedstock is in the synthesis of 2-butyne-1,4-diol, a precursor of 1,4-butanediol. This route still commands about 40% of the capacity for 1,4-butanediol. Acetylene is first reacted with 30–55 wt% formaldehyde in a series of three to five trickle bed columns at 80–110°C, 0.2–2.0 MPa over a silica supported Cu acetylide/Bi-promoted catalyst to produce

2-butyne-1,4-diol. Selectivity is about 90% on acetylene and greater than 90% on formaldehyde.



The CuO is converted to an acetylide under acetylene partial pressure. The reaction is first order in formaldehyde and zero order in acetylene. Back-mixed reactors helps to keep the unconverted formaldehyde level low in order to prevent catalyst inhibition and formation of formals. Conversion of formaldehyde is typically 97–99%. Dilute alkali is added to keep the pH in the 5–8 range. The Bi_2O_3 inhibits the formation of water-insoluble Cu-acetylene oligomers. The oligomers are quite shock sensitive, especially when dry. Deposition of these complexes in the reactor head spaces may lead to the formation of potentially explosive conditions and are to be avoided. Major by-products are propagyl alcohol, and sodium formate, along with unreacted formaldehyde. The butyne-diol is purified by vacuum distillation of the reactor effluent. Propagyl alcohol is removed as an azeotrope.

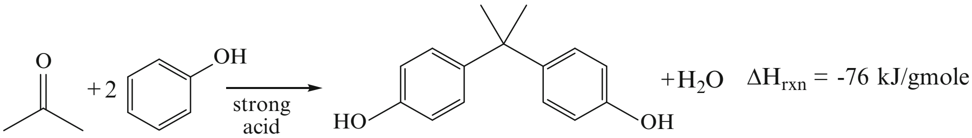
In the second step, the butyne is hydrogenated to 1,4-butanediol in the liquid phase at 70–170°C, 14–30 MPa, if Raney Ni is used, or 180–200°C, 20 MPa with Ni/Cu/Cr catalysts.



Both one- and two-stage fixed bed hydrogenation processes are used. Major by-products are methanol, *n*-propanol, *n*-butanol, hydroxyl-buyraldehyde, 2-methyl-1,4-butanediol, and formals Selectivity to butanediol is about 95%. The product is purified by distillation.

Bis-Hydroxyaryl Alkanes: Bisphenol A [251–253]

The bis-hydroxyaryl alkane structure, comprising two phenol molecules linked by a hydrocarbon chain, is synthesized by the acid-catalyzed liquid phase condensation of phenol with an aldehyde or ketone. Cyclohexanone, formaldehyde, and acetone adducts are produced commercially by similar processes. The acetone adduct, known as BPA, is by far the most important. BPA is produced from acetone and phenol in a fixed liquid-circulating bed or CSTR at 50–90°C, 0.11–0.40 MPa:



A high molar ratio of phenol to acetone, up to 15:1, helps suppress aldol reactions of acetone and multiple acetone/phenol condensations. The new carbon–carbon bonds occur almost exclusively in the 4-position of the phenol molecules. Acetone conversion is essentially complete, with yields greater than 90%. Sulfonated acid ion exchange resins are favored in newer installations over mineral acids, such as HCl or H_2SO_4, due to lower corrosion and ease of separation from the crude BPA.

Many flowsheet variations are practiced, but in one licensed process the reactor effluent is distilled to remove water and any unreacted acetone overhead. The bottoms of this column is distilled further to remove phenol for recycle and to produce a concentrated BPA bottoms stream. Solvent is added (alkane or aromatic) and the 1-1 phenol:BPA adduct is crystallized from the concentrate. Distilled to recover solvent, the remaining mother liquor is recycled to the reactor. A purge of the recycle, along with phenol-laden process water, is treated for removal of heavies and recovery of phenol. The adduct is melted and separated by distillation. Phenol is taken overhead, and the molten BPA underflow is prilled in a final processing step.

Hydration/Hydrolysis/Dehydration/Alcoholysis

Hydration of Olefins: Ethanol, Isopropanol, sec-Butanol, tert-Butanol [254–258]

Strong acid catalyzed hydration of olefins is an important industrial method for the production of alcohols. Ethanol, isopropanol, sec-butanol, and tert-butanol are produced in this fashion from ethylene, propylene, 1-butene, and isobutylene respectively. Both indirect and direct methods continue to be used. For unsymmetical olefins (anything greater than C_2), the addition of water follows Markovnikov's rule. Initial protonation of the olefin occurs so as to give the more stable, most substituted carbocation [259].

In the indirect method the olefin is absorbed into concentrated (about 60 wt%) sulfuric acid resulting in protonation of the olefin and formation of a carbocation. The HSO_4^- ion then adds to the carbocation to form mono- and di-alkyl sulfate intermediates. The alkyl sulfates are hydrolyzed to the corresponding alcohol by sparging with steam, which also serves to azeotropically distill the alcohol overhead and reconcentrate the sulfuric acid solution for recycle. Reactor conditions of 75–80°C, 0.6–3.0 MPa are typical, with selectivities of greater than 90%. The concentrated sulfuric acid solutions employed in the indirect process require the use of expensive corrosion-resistant materials of construction.

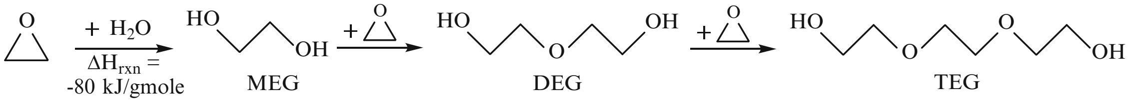
In direct processes, the carbocation intermediate is formed by a heterogeneous strong acid catalyst, such as a sulfonated polystyrene ion exchange resin, tungsten oxide, or supported phosphoric acid catalyst. Water then adds to form the corresponding alcohol in one step. Direct hydrolysis is an equilibrium-limited reaction, favored by low temperature, high pressure, and high water to olefin molar ratios. Both high (260–300°C, 7–20 MPa) and low (130–160°C, 8–10 MPa) temperature fixed bed processes are employed, with 20–75% conversion of the olefin per pass and selectivities of greater than 94%. Due to more mild conditions and lower corrosivity, direct hydrolysis has become the favored method.

In both direct and indirect methods, higher alcohols resulting from olefin oligomerization/hydrolysis, as well as the corresponding dialkyl ether of the product alcohol are typical by-products. All of the C_2-C_4 alkanols form low-boiling azeotropes with water. Thus, the recovery of an anhydrous product cannot be accomplished in a simple single-feed distillation step. Rather, dehydration of the alcohol is accomplished by a two or three column azeotropic distillation/decantation sequence. If dilute in the reactor effluent, the crude alcohol is first concentrated to close to its azeotropic concentration with water. An entrainer is added that forms an appropriate low-boiling heterogeneous azeotrope with water. Water is removed overhead as the heterogeneous water-entrainer azeotrope in a second distillation. This overhead mixture is decanted, with the water phase steam stripped to recover the entrainer and any alcohol, and the organic phase refluxed to the column. Anhydrous alcohol is recovered as the bottoms of the azeo column.

A convenient entrainer for isopropanol is diisopropylether made as a by-product in the hydration reaction. No additional entrainer is required for sec-butanol, nor for tert-butanol dehydration, as these form heterogeneous azeotropes with water. Diethyl ether can be used for ethanol, as well as other compounds such as *n*-hexane. A hybrid distillation/adsorption process, utilizing small pore molecular sieves for final removal of water is becoming more common, especially for dehydration of ethanol.

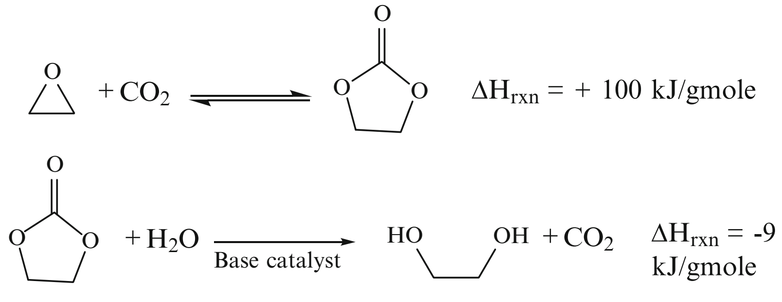
Hydrolysis of Epoxides and Carbonates: Ethylene Glycol, Propylene Glycol, and Higher Analogues [260, 261]

The ring opening of epoxide rings by addition of water is an extremely important industrial reaction, resulting in the formation of glycols and oligomeric ether glycols. Hydrolysis of ethylene oxide is the major industrial route to the production of ethylene glycol. In the typical industrial process, ethylene oxide is reacted thermally at 200°C with a high excess of water (20-1 mole ratio) to produce mono-, di-, and tri-ethylene glycols at a molar selectivity of 90-9-1 respectively.



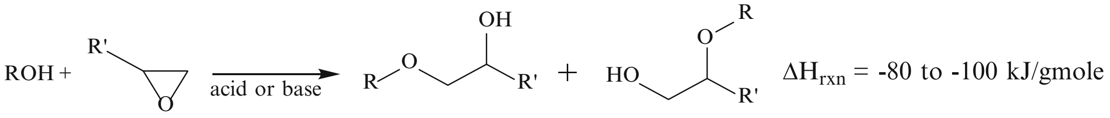
The hydrolyzed mixture is purified in a series of distillation columns with increasing vacuum level, to successively remove excess water, monethylene glycol, diethylene glycol, triethylene glycol overhead, and higher glycol oligomers as the bottoms of the final column. The purification train is quite energy intensive.

A recently commercialized process involves the hydrolysis of ethylene carbonate, with very high selectivity (>98%) to monoethylene glycol. In this process, ethylene oxide is reacted with carbon dioxide to form ethylene carbonate. Base-catalyzed hydrolysis of ethylene carbonate lead to predominantly monoethylene glycol.



Alcoholysis of Epoxides: Glymes and Ether Alcohols [262–266]

Another important use of epoxides is the formation of glycol ethers by alcoholysis:



When R´ is not hydrogen, such as with propylene oxide, two isomers may be formed. One isomer has a primary hydroxyl group and secondary ether group, the other has the opposite.

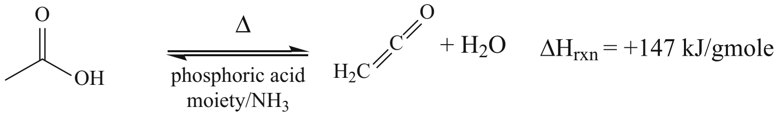
The reaction may be catalyzed by both acids and bases, exhibiting a catalyst-dependent product isomer mixture. The epoxide ring may open at either the primary or secondary C–O bond. With anionic (basic) catalysts, the epoxide ring opens preferentially at the least sterically hindered position, typically resulting in 95% or more secondary alcohol product (i.e., 1-alkoxy-2-alkanol). Acid catalysis affords a mixture of the 1-alkoxy-2-alkanol and 2-alkoxy-1-alkanol, with the relative ratio affected by the particular acid catalyst chosen and steric effects related to the size and structure of the reactant alcohol.

Large quantities of glycol ethers derived from C_1-C_4, C_6, and C_8 alcohols and ethylene oxide or propylene oxide are produced worldwide, mostly for solvent applications. In order to minimize the further reaction of the formed glycol ether with additional epoxide, a large excess of alcohol is used, at least 5:1 moles of alcohol per mole of epoxide. Alcoholysis of an epoxide is exothermic, with an average heat of reaction of about -80 to -100 kJ/gmole. Reaction conditions depend on the alcohol and epoxide, but are typically 170–220°C, 1.0–1.5 MPa. Conversion of the epoxide is essentially complete, with 80–90% molar selectivity to the mono glycol ether. The product is distilled to recover unreacted alcohol overhead for recycle, and the product glycol ether is distilled under vacuum in a second column, with higher glycol ethers as underflow. If markets dictate, diglycol and higher ethers may be recovered by further distillation of the bottoms.

The reaction of ethylene and propylene oxides with C_{10}-C_{14} detergent range alcohols affords ethoxylates and propoxylates, which are commonly employed as nonionic surfactants. The solubility of the product alkoxylate can be varied according to the number and type of epoxide molecules incorporated in to the molecule, as well as the chain length of the fatty alcohol. Longer chain groups reduce the solubility in water. Alcohols may be derived from natural fatty acids or higher oxo alcohols (either via ethylene oligomers or kerosene-based olefins).

Dehydration of Acetic Acid: Ketene [267–271]

Ketene, a highly reactive and useful intermediate, has been produced commercially since the 1920s by the high temperature, low pressure, equilibrium-limited, thermal decomposition of acetic acid, with concomitant generation of water (700–750°C, 0.005–0.02 MPa):



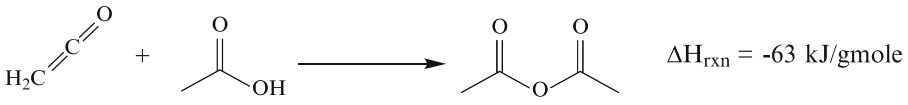
The equilibrium is shifted toward ketene production as the temperature is increased. The reaction occurs without a catalyst, but the higher temperatures needed to achieve reasonable rates and high equilibrium conversion without a catalyst lead to excessive decomposition to methane, CO, and CO_2. Commercial processes universally are catalyzed, typically with a derivative of phosphoric acid. Phosphoric acid itself, trialkyl phosphates, especially triethyl phosphate, and many other similar compounds are reported in the literature. Heterogeneous catalysis has been extensively studied, but has not been successfully commercialized [272].

In order to efficiently provide the large endothermic heat of reaction, the synthesis is carried out in a plug flow tubular reactor inside a fired furnace. Residence time is a matter of seconds, with achievable conversion of acetic acid of 70–80% per pass at the exit of the tubes. The temperature must be reduced quickly, with effective and rapid separation of liquid water and acetic acid from ketene vapor to prevent extensive back reaction of ketene with water. This quenching is usually accomplished by a series of low residence time, low pressure drop condensers. Overall conversion ends up at 40–70% per pass, depending on condenser design. Ammonia is added, either to the feed or to the furnace effluent to neutralize the phosphoric moieties and help prevent catalysis of the back reaction.

The condensed liquids, with concentration of 30–45% acetic acid in water, are processed, typically by extraction and/or azeotropic distillation to recycle low-water acetic acid to the furnace feed. The ketene thus produced is an efficient acylating agent, and is employed in a variety of downstream processes, such as direct production of esters via reaction with alcohols, synthesis of diketene, and predominantly for the formation of acetic anhydride. In the 1950s and early 1960s the major route to acrylic acid involved the pyrolysis of β-propriolactone, formed by the reaction of formaldehyde with ketene. This route has been completely supplanted by propylene oxidation.

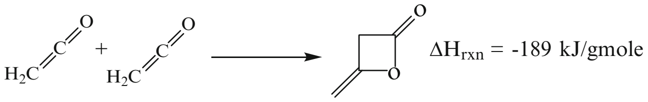
Ketene may also be produced by thermal decomposition of acetone above about 550–600°C (ketene + methane), or from acetic anhydride above about 600–650°C (ketene + acetic acid). The anhydride route is used to some extent to make ketene derivatives in India.

For acetic anhydride production, crude ketene effluent from the condensing train is scrubbed from the vapor phase by acetic acid in an absorber at 45–55°C. Rapid reaction with acetic acid produces acetic anhydride at essentially complete conversion and greater than 95% selectivity.



Heat is removed by a cooled circulation loop at the bottom of the absorber. The crude acetic anhydride, typically 80–90% purity, containing acetic acid and high boiling tars, is distilled in a two column sequence. Acetic acid and lights are recovered as the first distillate and largely recycled to the furnace, while acetic anhydride is recovered overhead in the second column, with tars removed as underflow.

Diketene is produced in a similar fashion by absorption of ketene into crude diketene in an absorber. The [2+2] cycloaddition product of ketene with itself is the four-member lactone, diketene:



Diketene is an extremely reactive and energetic molecule, but can be purified by distillation at reduced pressure and temperature. Diketene is used to produce a variety of specialty acetoacetates and arylides.

Esterification and Related Reactions [273–277]

Chemistry

Esterification is the reaction of a carboxylic acid with an alcohol to produce the corresponding ester and water:

In general, esterifications follow an $A_{Ac}2$ reaction pathway. A strong acid catalyst normally is necessary to achieve industrially viable rates. Typical catalysts are sulfuric acid, sulfonic acids, such as toluene sulfonic and methanesulfonic acids, or sulfonated polystyrene ion exchange resins in the hydrogen form. Although less expensive than sulfonic acids, sulfuric acid tends to promote dehydration of the alcohols to ethers and olefins more so than sulfonic acids. This is especially true for secondary and tertiary alcohols, which also suffer from significantly lower rates of reaction and equilibrium amounts of ester than primary alcohols.

The equilibrium constant for esterification, K_{eq}, can be expressed in terms of liquid concentrations or mole fractions and activity coefficients,

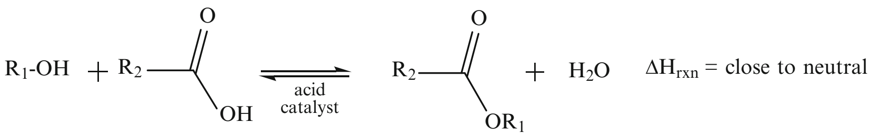
$$K_{eq} = \frac{[C_{ester}][C_{water}]}{[C_{acid}][C_{alcohol}]} = K_x K_\gamma = \frac{[x_{ester}][x_{water}]}{[x_{acid}][xC_{alcohol}]} \frac{[\gamma_{ester}][\gamma_{water}]}{[\gamma_{acid}][\gamma_{alcohol}]}$$

The equilibrium constant is usually on the order of magnitude of unity and the heat of reaction is close to neutral, as shown in Table 10.35 for a selection of esters. Thus, high conversion would require large excesses of alcohol or acid and corresponding large recycles of unreacted materials. To overcome these limitations, a common commercial approach

Table 10.35 Equilibrium constants and heat of reactions for acetate esters [277, 322, 323]

Carboxylic acid	Alcohol	K_{eq}	Standard heat of reaction, ΔH_r, kJ/gmole[a] [324]	Alcohol/acid molar ratio required for 50% conversion	Alcohol/acid molar ratio required for 90% conversion	Alcohol/acid molar ratio required for 99% conversion
Formic acid	Methanol	7.1	−16.0	0.57	2.0	14.8
Acetic acid	Methanol	5.2	−8.0	0.60	2.5	19.8
Acetic acid	Ethanol	4.0	−3.2	0.63	2.9	25.5
Acetic acid	*n*-Propanol	4.1	−4.9	0.62	2.9	24.9
Acetic acid	*n*-Butanol	4.2	−4.1	0.61	2.8	24.3
Acetic acid	2-Propanol	2.4	−3.3	0.71	4.3	41.8
Acetic acid	2-Butanol	2.1	−3.7	0.74	4.8	47.7
Acetic acid	3-pentanol	2.0	−3.9	0.75	5.0	50.0
Acetic acid	Tert-butanol	0.005	+3.6	93	1,400	18,000
Formic acid	Isobutanol	3.2	−1.5	0.66	3.4	31.6
Acetic acid	Isobutanol	4.3	−2.7	0.62	2.8	23.8
Butyric acid	Isobutanol	5.2	−14.2	0.60	2.5	19.9
Benzoic acid	Isobutanol	7.0	+5.0	0.57	2.1	15.0

[a] If $\Delta H_r > 0$, then endothermic, if $\Delta H_r < 0$, then exothermic



is to use reactive distillation with the aid of azeotropes and two-phase liquid-liquid formation to allow for high conversion at or near stoichiometric feeds. Ester formation is an excellent illustration of LeChatlier's principle; an equilibrium reaction can be driven to completion by removal of one or more products of the reaction. Several flowsheets are possible, depending on the relative boiling points and azeotropes formed by the acid-water-ester-alcohol system.

Low Boiling Esters

When the ester forms a heterogeneous binary low-boiling azeotrope with water or a heterogeneous tertiary water-ester-alcohol azeotrope, the continuous flow sheet illustrated in Fig. 10.30 can be used. Large volume esters made by this process include ethyl acetate, *n*-propyl acetate, isopropyl acetate, and *n*-butyl acetate. In this process, the acid and alcohol are continuously fed to the reboiler-reactor (**1**) of the esterification column (**2**), along with a homogeneous strong acid catalyst, such as sulfuric acid, methanesulfonic acid, or toluene sulfonic acid. The low-boiling water-laden azeotrope is taken as the distillate product, which decants into two phases (**3**). The water layer is steam-stripped (**4**), with organic distillates returned to the ester column and underflow water discarded. If the water content of the ester-water azeotrope is more than the amount of water created in the esterification reaction, then some of this water may be recycled to the ester column. A portion of the organic layer from the esterification column decanter is refluxed to the ester column, and the rest is distilled in a low boiler column (**5**). The low boiler column distillate, containing alcohol, water, and ester are returned to the reactor. The dry, alcohol-free ester underflow from the low boiler column is fed to a refining column (**6**) for final purification from heavies. Heavies, sludge, and spent catalyst is removed as the bottoms of the esterification column and discarded. Sometimes extra water is added near the top of the esterification column to aid in azeotroping the ester overhead.

High Boiling Esters

With high boiling esters that cannot be conveniently taken overhead, or those that boil too closely or form azeotropes

Fig. 10.30 Flowsheet for production of low-boiling esters

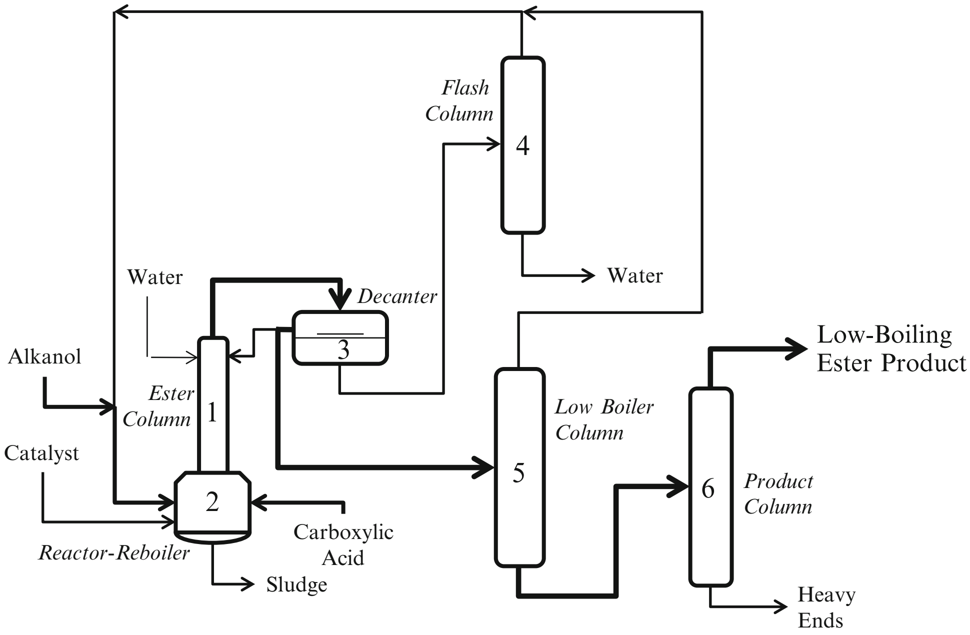
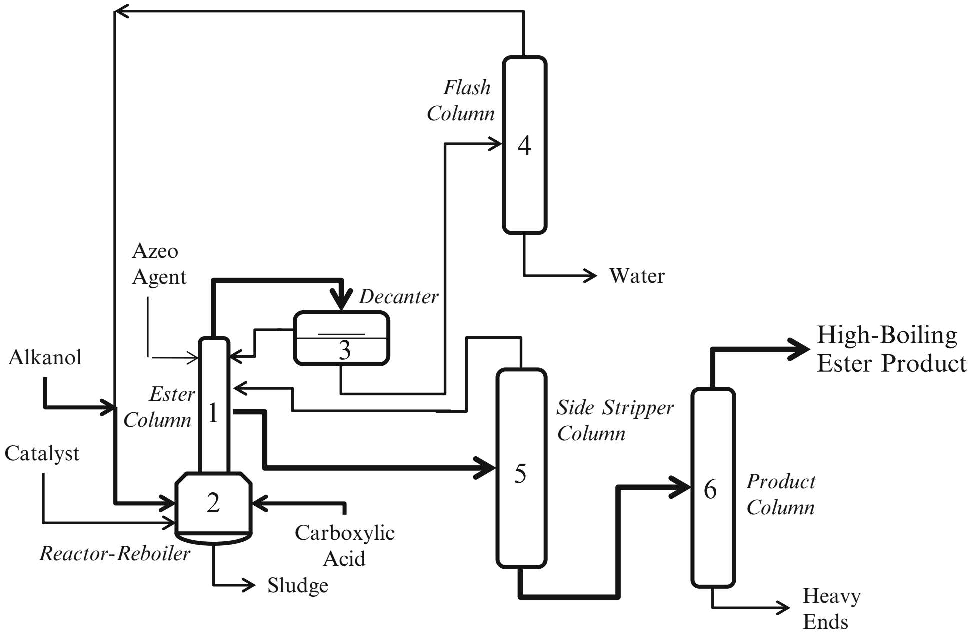


Fig. 10.31 Flowsheet for production of high-boiling esters



with their raw material acids and alcohols to allow for easy separation, the reaction still can be driven to completion by using an azeotropic distillation-sidedraw scheme as shown in Fig. 10.31. The acid and alcohol are continuously fed to the reboiler-reactor (**1**) of the esterification column (**2**), along with a homogeneous strong acid catalyst, such as sulfuric acid, methanesulfonic acid, or toluene sulfonic acid. Water is removed overhead by azeotroping out as any convenient heterogenous acid-water-ester-alcohol azeotrope or by adding an inert species such as cyclohexane or toluene that forms a low-boiling heterogeneous azeotrope with water. The azeotrope is decanted (**3**), with the organic layer refluxed to the ester column. The water layer is steam-stripped (**4**), with organic distillates returned to the ester column and underflow water discarded. The crude product ester is removed low down in the ester column as a sidedraw product with further purification by a sidedraw stripper (**5**). The ester is taken overhead in the product column (**6**), with heavy residue as underflow. Many glycol ether esters and higher boiling esters such as butyl butyrate can be made with such a flow sheet.

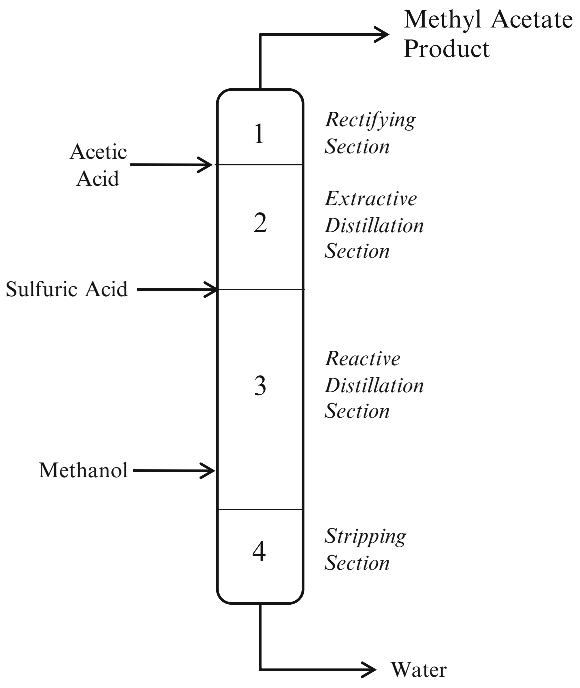


Fig. 10.32 Flowsheet for production of methyl acetate by reactive distillation

Methyl Acetate [278]

Methyl acetate cannot be produced in high purity using the simple esterification schemes outlined above due to unfavorable vapor–liquid and liquid-liquid equilibrium behavior:

1. The methyl acetate-methanol–water system does not form a ternary azeotrope.
2. The methyl acetate-methanol azeotrope is lower boiling than the water-methyl acetate azeotrope.
3. There is a distillation boundary between the two azeotropes.
4. The liquid-liquid region does not include either the methyl acetate-water or the methyl acetate-methanol azeotrope, nor does it cross the distillation boundary.

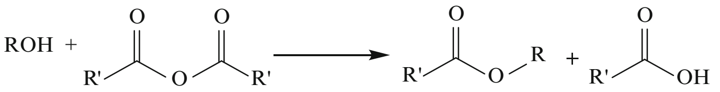
Consequently the water generated by esterification cannot be removed effectively, nor can the methyl acetate azeotropes be broken readily. The reactive distillation process (see Fig. 10.32) overcomes both unfavorable reaction and poor phase equilibriums to produce high purity methyl acetate and water in one column from near stoichiometric acetic acid-methanol feed ratios. Conceptually this column can be thought of as four heat-integrated distillation columns and a reactor stacked on top of each other. Reaction occurs below the sulfuric acid feed in a series of countercurrent high hold-up trays. This countercurrent separation, along with the low feed point of methanol results in high local excesses of at each end of the reactive section (**3**) in spite of the overall stoichiometric feed. The acetic acid feed acts as an extractive distillation agent above the reaction zone (**2**) to break the methyl acetate-methanol azeotrope. The upper rectification section (**1**) knocks back any acetic acid from the methyl acetate product, and the lowermost stripping section (**4**) removes methanol from the water bottoms. High purity methyl acetate is the distillate product, and clean water underflows.

Plasticizer Esters

One of the primary uses of higher boiling and low volatility esters is as plasticizers. A plasticizer increases the flexibility and ductility of a brittle thermoplastic polymer by reducing the glass transition of the polymer. PVC is the most commonly plasticized polymer. A large number of plasticizers based on phthalic anhydride, TPA, adipic acid, TMA, benzoic acid with 2-ethylhexanol, or C_9 and C_{10} iso-alcohols are made commercially, by anhydride reactions with alcohols (e.g., di-2-ethylhexylphthalate) or high boiling acid-alcohol reactions driven to completion by introduction of excess alcohol and removal of water generated [279].

Esterifications with Anhydrides

For hindered, unreactive alcohols, those susceptible to dehydration, and those producing very high boiling esters, it is often more favorable to produce the ester via reaction with an anhydride. The basic reaction of a symmetric anhydride with an alcohol is given by:



ΔH_{rxn} = -60 to - 70 kJ/gmole

The equilibrium is quite favorable and lies essentially completely toward the ester. Since the reaction is first order in both alcohol and anhydride, a small excess (less than 5 mole%) of either the alcohol or anhydride is added in practice to ensure complete conversion in reasonable reaction times. Typically a catalyst is not necessary, but strong acids or pyridine derivatives can be used to accelerate the rate.

Many carboxylic acid anhydrides can be made via reactive distillation in which a low boiling anhydride is reacted with a higher boiling carboxylic acid. This reaction, occurring stepwise, is equilibrium-limited and can be driven to complete formation of the symmetric anhydride of the high boiling acid by distillative removal of the generated low-boiling acid. The most common low boiling anhydride used in such a process is acetic anhydride.

Hydrogenation [280–283]

Chemistry

Hydrogenation involves the addition of molecular hydrogen (H_2) to a functional group to reduce its oxidation state. Typical substrates and products of hydrogenation reactions include: alkynes, alkenes, and aromatic rings to alkanes, aldehydes to primary alcohols, ketones to secondary alcohols, esters and carboxylic acid to alcohols, nitriles to imines and amines, and amides to amines. A closely related type of reaction, hydrogenolysis, involves hydrogen addition with concomitant breaking of other carbon–carbon or carbon-heteroatom (oxygen, nitrogen or halogen) bonds. Hydrogenation differs from protonation or hydride addition; in a hydrogenation the products have the same charge as the reactants. The ease of hydrogenation of various functional groups generally follows the decreasing order: [284]

$$\text{alkynes} > \text{alkenes} > \text{nitriles, nitros} > \text{aldehydes} > \text{ketones} \\ > \text{aromatic rings} > \text{anhydrides} > \text{esters} > \text{carboxylic acids}$$

Steric factors and substitution patterns further dictate the ease of hydrogenation, decreasing in order of straight chain > branched far from the functional group to be hydrogenated > branched adjacent to the functional group to be hydrogenated. Bulky substituents also tend decrease reactivity. Selective hydrogenation of one functional group in the presence of another functional group (e.g., hydrogenation of the double bond of an α β-unsaturated ketone, while preserving the ketone group) is a fairly common issue, often solved by judicious selection of catalyst metal as well as reaction conditions.

A hydrogenation reactor must be designed for effective heat and mass transfer. Most hydrogenation reactions are highly exothermic, with heats of reaction at least above −50 kJ/gmole and high adiabatic temperature rises. Typical heats of reaction for various functional group hydrogenations are given in Table 10.36. Moreover, hydrogen gas must get to the surface of the catalyst through vapor or liquid films. Common reactor formats include CSTR's with internal cooling coils or external exchanger loops, Buss loop reactors, and trickle beds with cooling loops. These formats provide good mixing for mass and heat transfer, but as backmixed systems require large reactor volumes for high, i.e., greater than 95%, conversion. Often a small plug flow polishing fixed bed is included after the main reactor to get the conversion up to above 99%. See for example, the hydrogenation of benzene to cyclohexane. Examples of commercially significant hydrogenation reactions are given in the subsections which follow.

Table 10.36 Typical heats of reaction for hydrogenation of various functional groups

Functional group transformation	Heat of reaction, kJ/gmole
Saturation of double bond	−90 to −130
Saturation of triple bond	−300
Aromatic ring saturation	−200 to −215
Carbonyl to alcohol	−60 to −65
Ester to alcohols	−25 to −100
Acid to alcohol	−25 to −75
Nitro to amine	−550

Methanol [285–289]

Methanol is produced by the hydrogenation of carbon monoxide and carbon dioxide. The first industrial production of methanol from CO, CO_2, and H_2, i.e., synthesis gas, was carried out over a zinc oxide/chromium oxide catalyst at high pressure (25–35 MPa) and temperature (300–450°C). Developments during the late 1960s and early 1970s resulted in the first modern Cu/ZnO/catalysts capable of operation at much lower temperatures (200–310°C) and pressures (4–10 MPa). The low pressure process results in lower investment and production costs, higher yields, improved reliability, and much larger potential plant size. Plant capacity has increased dramatically from less than 1,000 metric tons per day for old plants, to more typically greater than 2,500 tons per day, with single train capacities of 5,000 metric tons per day or more now in operation.

The synthesis of methanol can be thought of as the hydrogenation of carbon monoxide and carbon dioxide. A typical feed gas contains somewhat higher than a 2–1 M ratio of H_2 to CO, and 2–12 mole% CO_2. The feed gas composition often is characterized by the stoichiometric number, SN, defined as:

$$\mathrm{SN} = \frac{(\mathrm{H_2} - \mathrm{CO_2})}{(\mathrm{CO} + \mathrm{CO_2})}$$

The value of SN should be about 2.05 for optimal performance. Syngas for methanol production can (and does) originate from a variety of sources and processes

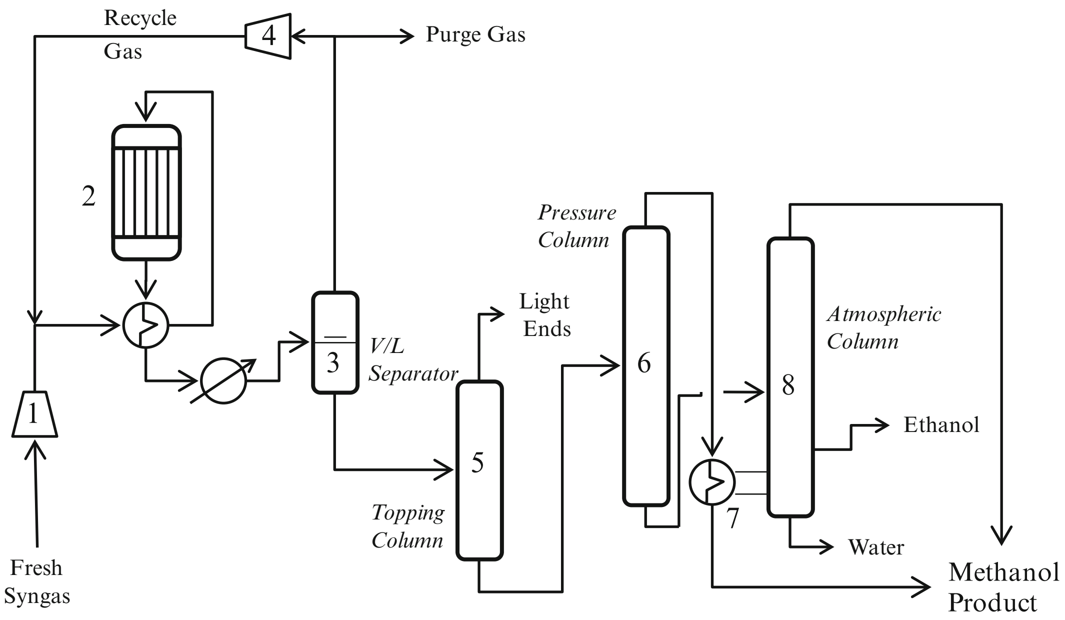


Fig. 10.33 Flowsheet for production of methanol

including natural gas via SMR, partial oxidation (POX), or autothermal reforming (ATR), as well as gasification of coal. ATR generally produces the appropriate SN for methanol directly. SMR needs CO_2 addition; coal gasification and POX require additional water-gas shift and subsequent CO_2 separation to enhance hydrogen sufficiently to give the desired SN of 2.05. See Section "Synthesis Gas", and Chaps. 19, 20, and 22 for further details on syngas production.

The net stoichiometry of the reaction is:

$$CO + 2H_2 \xrightleftharpoons{Cu/ZnO/Al_2O_3} CH_3OH \qquad \Delta H_{rxn} = -91\ \text{kJ/gmole}$$

In reality, the main source of the carbon incorporated into the methanol molecule is actually derived from carbon dioxide via reaction with hydrogen:

$$CO_2 + 3H_2 \rightleftharpoons CH_3OH + H_2O \qquad \Delta H_{rxn} = -49\ \text{kJ/gmole}$$

As the carbon dioxide in the feed is consumed, the reverse water-gas shift reaction kicks in to supply more:

$$CO_2 + H_2 \rightleftharpoons CO + H_2O \qquad \Delta H_{rxn} = +42\ \text{kJ/gmole}$$

The overall methanol synthesis network is equilibrium-limited, exothermic, and results in a reduction of volume upon reaction. Low temperatures and high reaction pressures favor conversion to methanol.

The modern low pressure methanol process is relatively simple as shown in Fig. 10.33. A syngas with SN = 2.05, less than 0.05 ppm sulfur (as H_2S), is compressed (**1**), if needed, to the reactor loop pressure of 4–10 MPa,[2] interchanged with the reactor effluent, combined with recycle gas, and fed to the fixed bead reactor containing the $Cu/ZnO/Al_2O_3$ catalyst (**2**). Both multistage intercooled adiabatic, and multi-tubular steam cooled reactors are used commercially. Some commercial processes use reactors in series at different conditions for higher per pass conversion. The reactor temperature is usually maintained below 260°C. Conversion of CO is about 50% per pass. Major by-products are ethanol (150–200 ppm); higher alcohols (300–400 ppm); acetone, MEK, and other ketones (less than 10 ppm); and depending on the generation of catalyst, 20–100 ppm dimethyl ether. Yield from CO is quite high, usually more than 99%.

The methanol–rich reactor effluent is cooled, and crude methanol is separated from non-condensables before further purification (**3**). A small purge is taken from the separator off-gas to prevent build-up of inerts such as argon, nitrogen, methane entering with the inlet syngas. The remainder of the gas (typically at a SN of 3 to about 10) is boosted to loop pressure and recycled (**4**).

The crude methanol is typically purified in a two or three column distillation sequence. The first column (**5**) removes light ends overhead (dissolved gases and other species). The product methanol is then distilled overhead, with water taken as underflow, and intermediate boiling-impurities

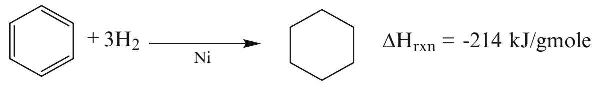
---

[2] Coal gasification-derived gas is usually high enough pressure, SMR, ATR, POX gas needs compression.

(ethanol, higher alcohols, ketones) as a sidedraw, if high purity is needed. The product distillation is often done as a dual-column pressure swing setup to reduce energy consumption. In this scheme, steam is fed only to the reboiler of first high pressure methanol product column (**6**). A portion of the methanol is distilled overhead, with methanol–water as bottoms. The vaporous product methanol from the first column is not condensed, but used to drive the reboiler (**7**) of the second lower pressure column (**8**). The remainder of the methanol is distilled in the second column, with water taken as bottoms product, and impurities as a sidedraw if needed.

Aromatic Ring Saturation: Cyclohexane [290–292]

Cyclohexane, via the hydrogenation of benzene, is an important intermediate in the production of adipic acid and caprolactam, monomers for nylon 6,6.



Both vapor and liquid phase processes are practiced commercially. In one common liquid phase process, benzene is hydrogenated isothermally over a fine slurry of suspended Raney nickel at 180–200°C, 5 MPa. Soluble catalysts are also used. The heat of reaction is removed by a combination of an external steam generating heat exchanger and by vaporization of part of the reactor contents. Vapors from the first stage reactor are passed to a fixed-bed polishing reactor for final conversion. Conversion is about 95% in the first stage, with essentially complete conversion in the polishing bed. The vapor effluent is flashed, condensed and separated from hydrogen and light gases for recycle and purging. If the hydrogen purity is low, then the cyclohexane is fed to a small stabilizer column to remove light ends. Low temperature operation significantly reduces the equilibrium isomerization of cyclohexane to methylcyclopentane. Selectivity is greater than 99.8% to benzene with less than 100 ppm methylcyclopentane.

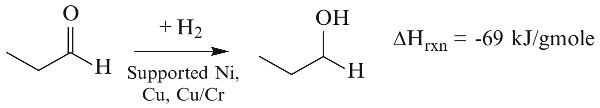
α,β-Unsaturated Carbonyls to Saturated Ketones and Aldehydes [293, 294]

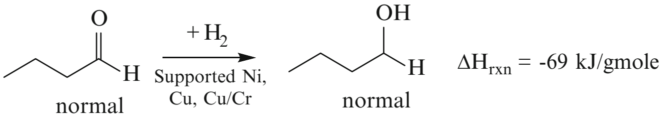
An important aspect of the production of ketones and aldehydes via aldolization (see section "Aldol Condensation") is the selective hydrogenation of the double bond of the α,β-unsaturated carbonyl intermediate without formation of alcohols. Pd on alumina, silica, titania, or carbon have been found to be very effective catalysts for this transformation, with 100/1. Ni, Cu, Cu/Cr, and Pt catalysts normally are not selective for double bond hydrogenation without excessive alcohol formation.

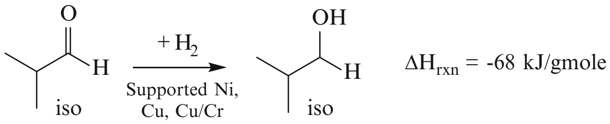
Much research effort has been expended on developing selective catalysts for reduction of the α,β-unsaturated carbonyl to an alcohol, while retaining the double bond functionality.

Aldehydes and Ketones to Alcohols [295, 296]

Aldehydes are readily hydrogenated to their corresponding alcohols under relatively mild conditions. Three alcohols of significant volume produced by the hydrogenation of Oxo-derived aldehydes are *n*-propanol, *n*-butanol, and *i*-butanol:

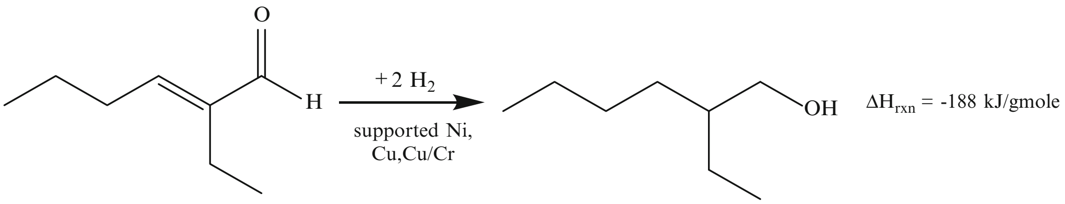






The hydrogenation may be conducted in either the gas or liquid phase with supported Ni, Cu, or Cu/Cr, or Cu/Zn catalysts at about 115–160°C, 0.2–8 MPa. Yields are typically above 98%. By-products consist of minor amounts of acetals, ethers, and higher boiling condensation products. Detergent range alcohols are also produced in this fashion from higher Oxo aldehydes.

The important plasticizer alcohol, 2-ethyl-hexanol, may be synthesized by a similar hydrogenation of 2-ethylhexaldehyde, or alternatively directly from the aldol-derived α β-unsaturated aldehyde by the simultaneous hydrogenation of the double bond and carbonyl.



Both vapor and liquid phase processes are utilized with supported Ni, Ni/Cu, Cu, or Cu/Cr catalysts. Vapor phase processes are commonly run in two stages, with about 90% conversion in the first stage. The first stage is operated at 100–170 or 200–250°C, 0.5–5.0 MPa. The final polishing stage is operated at 200–250°C, 0.5–5.0 MPa.

Carbinols, the common name for ketone-derived secondary alcohols, can be produced in very similar vapor or liquid hydrogenation processes under slightly more forcing temperature and pressure conditions than aldehyde conversion. The conversion of more hindered ketones, such as MIBK (4-methyl-2-pentanone) to methyl isobutyl carbinol (MIBC) require somewhat longer residence times, temperature, and/or hydrogen pressure, but yields are essentially quantitative. The equilibrium for formation of ketols from ketones and product alcohols is much less favorable than for aldehyde-alcohol acetal formation.

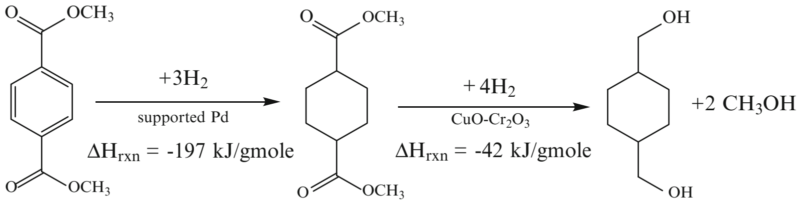
Supported heterogeneous Pt and Ru catalysts are successfully employed for the reduction of highly hindered or sensitive carbonyls, as exemplified by the conversion of the ring-strained 2,2,4,4-tetramethyl-1,3-cyclobutadione to 2,2,4,4-tetramethyl-1,3-cyclobutadiol [297], but are not normally used for more commodity applications.

Esters to Alcohols: Dimethyl Maleate to γ-Butyrolactone, THF, and 1,4-Butanediol [298–301]

In one commercial process, maleic anhydride is esterified with methanol in a reactive distillation column to form the intermediate dimethyl maleate. The reaction is driven to completion by removal of water and methanol from the top of the column. Methanol and water are separated by distillation, with the methanol recycled. The maleate ester underflow from the first column is vaporized in excess hydrogen and is hydrogenated in a fixed bed adiabatic reactor. The reactor effluent is cooled, condensed, and the crude BDO product is separated from hydrogen (which is recompressed and recycled). The crude BDO is separated in a three-column distillation sequence. THF and other lights are removed overhead in the first column. Methanol is distilled overhead in second column for recycle, and high purity product BDO is the distillate of the third column. Unreacted maleate ester and γ–butyrolactone by-product in the underflow of the column are recycled for further conversion. The molar yield of BDO from maleic anhydride is about 97%. Conditions can be modified to produce all THF or to recover GBL. In another commercial process maleic anhydride is hydrogenated directly without first making the diester.

Esters to Alcohols: Dimethyl Terephthalate to 1,4-Cyclohexanedimethanol [302]

An example of the large-scale hydrogenation of an ester to alcohols is the production of 1.4-cyclohexanedimethanol (CHDM) from DMT. DMT in a mixture of methanol and CHDM is hydrogenated in the liquid phase over a supported Pd catalyst at 110–180°C, 30–40 MPa to give a mixture of cis/trans-cyclohexane-1,4-dicarboxylic acid dimethyl ester. Further hydrogenation over copper chromite at more severe conditions yields CHDM.



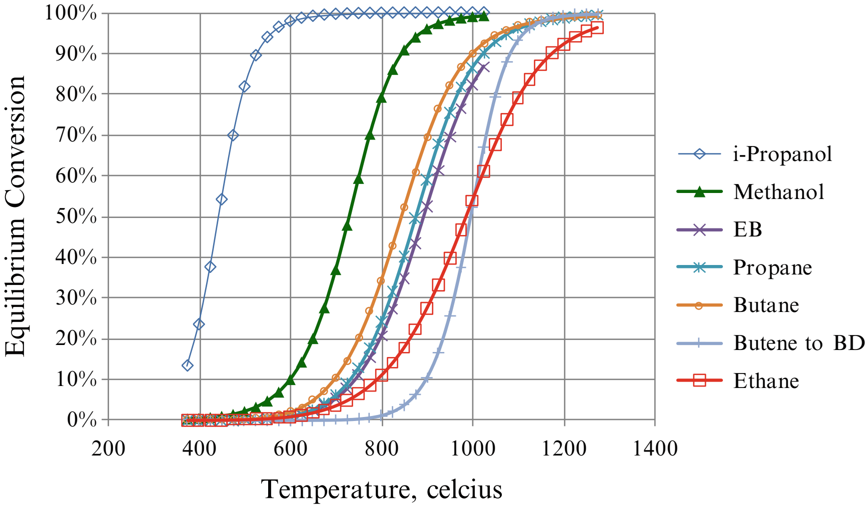
Co-polyesters of TPA, ethylene glycol, and CHDM have enhanced toughness, better injection molding and sheet extrusion properties than simple polyethyleneterephthalate polymers.

Dehydrogenation [303–305]

Chemistry

Conceptually, direct dehydrogenation, a reaction producing hydrogen as a by-product, is simply the reverse of hydrogenation. Direct dehydrogenation reactions find commercial use primarily in the conversion of saturated alkanes and alkyl aromatics into olefinic and diolefinic compounds (butanes to butenes, butenes to butadiene, ethylbenzene to styrene, long chain *n*-paraffins to olefins), as well as the conversion of secondary alcohols into ketones (isopropanol to acetone, sec-butanol to 2-butanone, methanol to formaldehyde). Oxidative dehydrogenation, wherein the hydrogen is converted in situ to water, is also practiced for accomplishing the same ends as direct dehydrogenation. See for example the discussion on formaldehyde earlier in this chapter.

Fig. 10.34 Effect of temperature on equilibrium conversion for several dehydrogenation reactions, pressure of 0.1 MPa



Dehydrogenation reactions are endothermic and generally severely equilibrium-limited. Reasonable rates and conversion are favored by high temperature and low pressure, as the number of moles increases with generation of hydrogen. Figure 10.34 presents the effect of temperature on equilibrium conversion for several important dehydrogenation systems (at 0.1 MPa total pressure, with no diluents added). Note that very high temperatures are required with most hydrocarbon systems for a reasonable conversion of 50%. This graph illustrates at least one reason ethane to ethylene dehydrogenation has not been commercialized. To get conversions of ethane comparable to steam cracking, one has to operate at temperatures where thermal cracking is significant anyway. High temperature operation leads to other collateral problems:

1. Significant high temperature (i.e., expensive) energy is needed for heating feeds and reaction mixtures.
2. The thermal cracking of molecules has a high activation energy and tends to compete with dehydrogenation above about 650°C.
3. Consecutive dehydrogenations of the same molecule (often energetically favorable) tend to lead to rapid coke formation and catalyst fouling, with frequent regenerations required.
4. Metal sintering and permanent deactivation is accelerated at high temperatures.

Reactor designs reflect these factors. Common designs include, (a) direct-fired furnace feed heaters with adiabatic fixed beds in series (often with hot shots of feed between beds), (b) isothermal fixed beds (reactor in a furnace), (c) moving beds to allow for regeneration, and (d) fluid beds for heat control, with separate continuous regenerators (much like FCC units).

Manipulation of the partial pressure of products generated is another approach to allow lower temperature operation. Common methods include:

1. Operation at subatmospheric pressure.
2. Addition of steam or other inerts to lower hydrogen and product partial pressures.
3. React away hydrogen as it is formed or form water instead (i.e., oxidative dehydrogenation).

The effect of the magnitude of the equilibrium constant, inert addition, feed hydrogen, and total system pressure on equilibrium conversion can be calculated readily by the relationship given below [306]:

$$K_{\mathrm{p}} = \frac{XP(H+X)}{(1-X)(1+I+H+X)}$$

where

K_{p} = temperature-dependent equilibrium constant
X = molar conversion of reactant
H = moles of H_2 in feed to moles of reactant
P = total system pressure
I = moles of inerts in feed to moles of reactant

The process designer has many degrees of freedom, with tradeoff on temperature, partial pressure, and diluents addition, as well as reactor format. Thus, many flowsheet alternatives can accomplish the same end result, and many versions of dehydrogenation processes, particularly for alkanes to olefins, have been successfully commercialized.

Catalysts used for hydrogenations, such as precious metals and Group VIII transition metals, generally also are useful for dehydrogenation reactions. Supported Pt—Sn and Cr-based catalysts are the most common used commercially. Although inactive itself for dehydrogenation, the addition of Sn to Pt catalysts acts to suppress hydrogenolysis, moderates coking rate (promotes migration of coke from the active catalyst metal to the support), and reduces sintering. Table 10.37 summarizes commercial conditions and

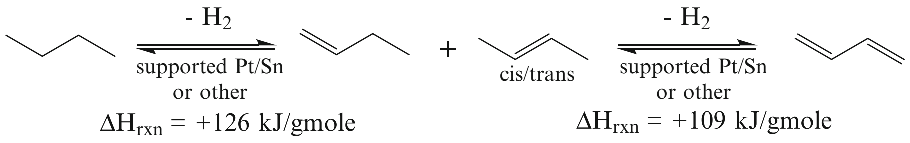
Table 10.37 Important dehydrogenation reactions

Reactant	Product	Catalyst	Conditions	Heat of reaction, kJ/gmole
C_3-C_4 alkanes [325]	C_3-C_4 mono-olefins	Pt/Sn on γ-Al_2O_3, zirconia, or Zn/Ca-aluminate; chromia on γ-Al_2O_3	530–650 °C; 0.03–0.5 MPa; some processes steam co-feed	+126 to +137
C_{10}-C_{16} alkanes	C_{10}-C_{16} internal alkenes	Promoted Pt/Sn on γ-Al_2O_3	300–550 °C; 0.1–0.3 MPa; H_2/alkane (5–9)/1	+120 to +135
Butenes [307]	1,3-Butadiene	Chromia on γ-Al_2O_3; cromium oxide-Ca-Ni-phosphate; iron oxide	575–700 °C; 0.014–0.03 MPa; some with steam co-feed; some processes use oxydehydrogenation	+109
Ethylbenzene	Styrene	K_2O/Fe oxide	600–720 °C; 0.1 MPa; steam co-feed, (1–2)/1 wt ratio	+125
sec-Butanol	Methyl ethyl ketone	Cu oxide, Cu chromite, Cu/K/SiO_2, ZnO, Pt/Al_2O_3	240–420 °C; 0.1–0.6 MPa	+59
Isopropanol	Acetone	Cu oxide, Cu chromite, Cu/Na/SiO_2, ZnO	220–400 °C; 0.1–0.3 MPa	+67
Methanol	Formaldehyde	Silver gauze or crystals	600–720 °C; 0.1 MPa; air co-feed	+85

catalysts for a number of important industrial dehydrogenation reactions. More detailed discussions of a few key processes are given below.

Butenes and 1,3-Butadiene [307–309]

A large number of similar commercial processes have been developed for the dehydrogenation of butanes to butenes and butanes/butenes to 1,3-butadiene, using heterogeneous catalysts (e.g., Pt/Sn on γ-Al_2O_3; zirconia, or Zn/Ca-aluminate; Chromia on γ-Al_2O_3; Cromium oxide-Ca-Ni-phosphate; iron oxide) at 530–700°C, 0.03–0.5 MPa.



Butadiene is particularly susceptible to polymerization and coking, so high steam to feed ratios are often used to reduce coke formation and to lower the partial pressure of hydrogen to shift the equilibrium toward butadiene. Even with these steps, coking occurs rapidly. Multiple reactors are used to allow regeneration/operation to occur in cycles lasting on the order of an hour or less.

In an alternative approach, oxidative dehydrogenation methods have also been employed to overcome the unfavorable equilibrium. In one version, butanes and/or butenes, steam, and air are fed to a fixed bed reactor at a relatively low inlet temperature of about 370°C, where combustion of produced hydrogen supplies the necessary heat of reaction for the dehydrogenation. The outlet temperature rises to about 480–600°C due this combustion. Conversion over the ferrite/Zn or ferrite/Mn catalyst is greater than 60%, at up to about 90–93% selectivity.

Internal Olefins from Higher *n*-Alkanes [310, 311]

The dehydrogenation of C_{10}-C_{14} *n*-paraffinic materials derived from kerosene fuels is an important source of internal olefins for linear alkylbenzene production (see Section 10.3 above). In a typical commercial process, a narrow-boiling kerosene cut is first treated by shape-selective adsorption to recover a stream high in *n*-paraffins. This material is dehydrogenated in the vapor phase over Pt/Sn on alumina (often with In and Li promoters) at 300–550°C, 0.1–0.3 MPa, with H:C ratio of (5–9):1. The higher alkanes are particularly susceptible to thermal cracking reactions, and the hydrogen helps prevent this decomposition pathway. Conversion is 10–15% per pass at selectivities of 90–94%. Unreacted *n*-paraffins for recycle are separated from olefins by adsorption. This is another example of a separation exploiting functional group differences to effect the separation in a single step, whereas separation by volatility differences (distillation) would be prohibitively complicated.

Styrene from Ethylbenzene [312–315]

Direct dehydrogenation of ethylbenzene to styrene accounts for about 85% of commercial production. The remaining styrene is produced via coproduction with propylene oxide (see section 1 of this chapter). The major reaction is the endothermic (+125 kJ/gmole), high temperature, vapor phase conversion of ethylbenzene to styrene and hydrogen.

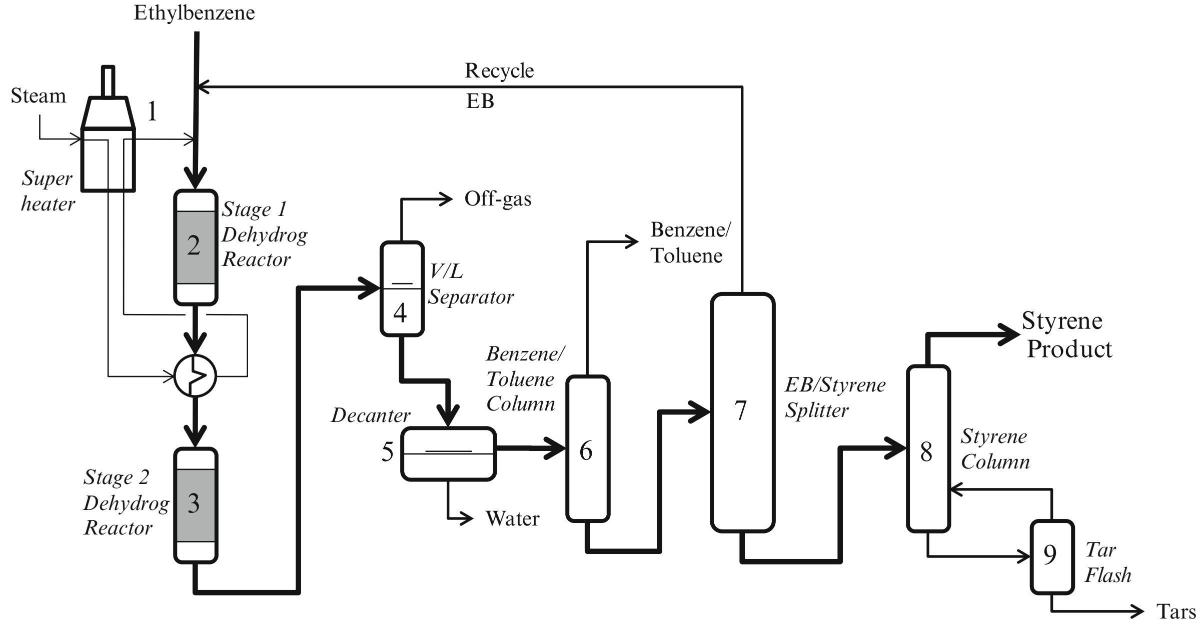


Fig. 10.35 Flowsheet for dehydration of ethylbenzene to styrene

Side reactions include the thermal degradation of ethylbenzene to benzene and ethylene, and catalyzed formation of toluene and methane from styrene and hydrogen.

In a typical commercial process as shown in Fig. 10.35, superheated steam (**1**) and ethylbenzene (1:1 weight ratio, 6:1 M ratio) are fed at about 620°C over potassium promoted iron oxide catalyst in a series of adiabatic fixed bed reactors (**2**), (**3**), under vacuum (around 0.06 MPa). The effluent from each stage must be reheated by superheated steam injection or indirect fired heating to keep equilibrium conversion and rates sufficiently high. Overall conversion is typically 60–70% per pass at up to 97% selectivity to styrene. Steam plays an important role in the process. It:

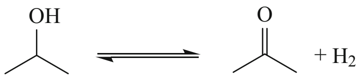
1. Brings in heat to get the reaction mixture up to temperature
2. Reduces the partial pressures of products allowing a tradeoff between temperature and conversion
3. Removes coke by steam–carbon reaction (catalyzed by potassium)
4. Minimizes EB cracking
5. Keeps the iron catalyst in the correct oxidation state (H_2 will reduce iron to its catalytically inactive elemental state)
6. Redistributes potassium on the working catalyst, enhancing lifetime

The reactor effluent is cooled, separated from off-gases (**4**) and allowed to phase separate to remove the bulk of the water (**5**). Purification is accomplished by a straight-forward series of distillations, but time at high temperatures must be limited to minimize losses to styrene polymerization. Toluene and benzene are distilled overhead first (6), followed by distillation of ethylbenzene for recycle (**7**), and finally the styrene product is taken overhead from heavies (**8**). Additional styrene is recovered from the tar stream in a flash step (**9**). Styrene purities range from 99.85 to 99.95%. Ethylbenzene and styrene are close boiling, with a relative volatility of about 1.3. Thus, 70–100 fractionation stages are required, depending on purity requirements. Most modern plants use high efficiency packing to reduce pressure drop, increase stage efficiency for the same height, and to increase throughput for a given column diameter. Styrene is quite prone to polymerization.

An oxidative reheat process is also commercially practiced, typically for retrofits to increase capacity of bottlenecked plants. Air is introduced between the dehydrogenation stages to convert some of the hydrogen into water to overcome the equilibrium limitations and to generate heat. Advantages include increasing the ethylbenzene conversion to about 75% per pass, unloading the costly ethylbenzene/styrene fractionation column, reducing interstage heating requirements, and lowering superheated steam consumption.

Dehydrogenation of Alcohols: MEK, Acetone, and Formaldehyde [316–320]

Aldehydes and ketones can be made by the dehydrogenation of alcohols, liberating hydrogen as a by-product, as illustrated for acetone and MEK:



$\Delta H_{rxn} = +67$ kJ/gmole

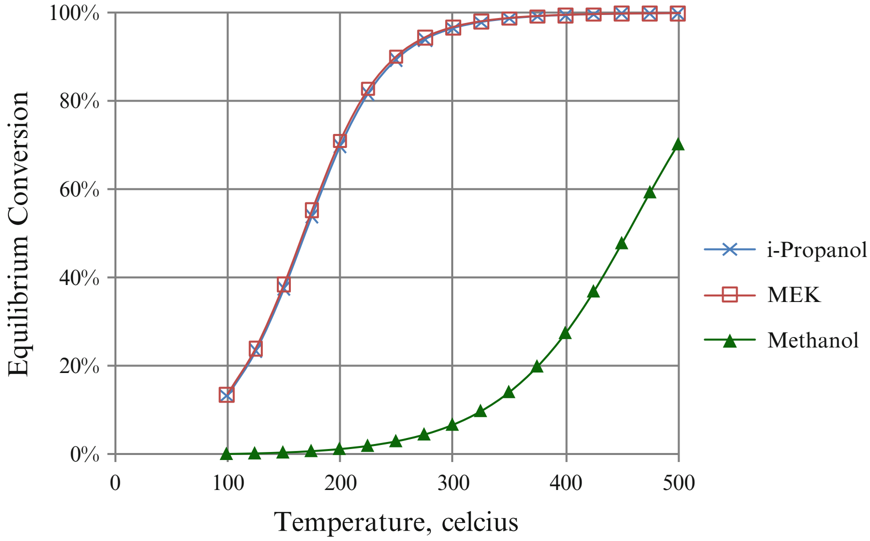
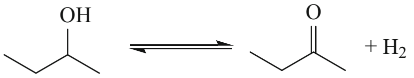


Fig. 10.36 Effect of temperature on equilibrium conversion for alcohol dehydrogenations



$\Delta H_{rxn} = +59$ kJ/gmole

These reactions are equilibrium-limited, modestly endothermic, and are favored by low pressure and high temperatures. Required temperatures are considerably lower than for alkane dehydrogenation. Compare Figs. 10.34 and 10.36. Both 2-butanone (MEK) and acetone are produced commercially by similar high temperature vapor phase catalytic dehydrogenations, from sec-butanol and isopropanol respectively. Typical conditions are 220–450°C, 0.1–0.6 MPa. A multitubular fixed bed reactor with hot oil circulation provides the endothermic heat of reaction. Copper catalysts are highly active for these dehydrogenation reactions and give high selectivities, typically greater than 90%. The main side products are the alkene (butenes or propylene) formed by dehydration of the parent alcohol, as well as small amounts of aldol condensation products. The high temperatures and the formation of unsaturated by-products lead to relatively rapid coking and deactivation. The catalyst must be burned off roughly every 3–6 months for optimal productivity.

Low temperature (150°C) liquid phase processes using Raney nickel or copper chromite are also practiced for the production of MEK, especially in Europe. Selectivity is somewhat better than the vapor phase process. Dehydrogenation of isopropanol to acetone is currently a relatively minor contributor to acetone production worldwide, with most acetone is derived as a by-product of phenol production (see section "Phenol/Acetone or MEK and Hydroquinone/Acetone" for a discussion of the phenol/acetone process).

A similar direct dehydrogenation of methanol to formaldehyde occurs to a limited extent with commercial silver catalysts for the conversion of methanol to formaldehyde, but is combined with oxidative dehydrogenation for more efficient operation (see the section, "Formaldehyde," for further details).

Acknowledgements The author would like to thank Robert T. Hembre and Steven L. Perri, both of Eastman Chemical Co., Kingsport, for taking their time to read and edit the chapter. The author would also like to acknowledge Robert Hembre's helpful revisions to the sections, "Carbonylation" and "Hydroformylation," in particular, as well as his addition of the section, "Metathesis."

References

1. James JL, Shore D (1993) Chemicals from petroleum. In: McKetta JJ (eds) Chemical processing handbook. Marcel Dekker, New York, p 23 and calculations by the author
2. American Chemistry Council (2011) Global business of chemistry. http://www.americanchemistry.com/s_acc/sec_directory.asp?CID=292&DID=747
3. Rase HF (2000) Handbook of commercial catalysts. CRC Press, Boca Raton
4. Leach BE (1983) Applied industrial catalysis, vol 1–3. Academic, New York
5. Rase HF (1990) Fixed bed reactor design and diagnostics. Butterworth, Stoneham
6. Satterfield (1977) Chemical reactor design for process plants, vols 1 & 2. Wiley, New York
7. Rase HF (1977) Chemical reactor design for process plants, vol 1 & 2. Wiley, New York
8. Doherty MF, Fidkowski ZT, Malone MF, Taylor R (2008) Distillation (section 13). In: Green DW (ed) Perry's chemical engineers' handbook, 8th edn. McGraw Hill, New York
9. Barnicki SD, Siirola JJ (1997) Enhanced distillation, In: Perry RH, Green DW, Maloney JO (eds) Perry's chemical engineers' handbook, 7th edn. McGraw Hill, New York, pp. 13–54, 13–85
10. Barnicki SD, Hoyme CA, Siirola JJ (2006) Separations process synthesis, In: Seidel A (ed) Kirk-othmer encyclopedia of chemical technology, vol 10, 6th edn. Wiley, Hoboken, pp. 297–339
11. Dimian AC, Bildea CS (2008) Chemical process design. Wiley-VCH Verlag, Weinheim

12. Sundmacher K, Kienle A, Seidel-Morgenstern A (2005) Integrated chemical processes: synthesis, operation, analysis, and control. Wiley-VCH Verlag GmbH & Co, Weinheim
13. Stichlmair JG, Fair JR (1998) Distillation principles and practices. Wiley-VCH Verlag, New York
14. Rousseau RW (ed) (1987) Handbook of separations process technology. Wiley, New York
15. Chemical economics handbook. Marketing research reports. SRI Consulting, Menlo Park, CA, (2006–2010)
16. Chemical profiles. ICIS chemical business, http://www.isis.com, (2006–2010)
17. Manufacturing energy consumption surveys (MECS). U.S. energy information agency, department of energy, Washington, DC, USA 1998, 2002, 2006. http://www.eia.doe.gov/emeu/mecs/
18. Annual energy review 2009, Report No. DOE/EIA-0384(2009) U. S. energy information agency, department of energy, Washington, DC, USA. 19 Aug 2010
19. Petroleum and fuel oil, data for all years including revisions, State Energy Data System (SEDS). U.S. Energy Information Agency, Department of Energy, Washington, DC, USA (2009) http://www.eia.gov/emeu/states/_seds_updates.html
20. Davis S (2008) Petrochemical industry overview. In: CEH Marketing Research Report, SRI Consulting, Menlo Park, CA, USA, p 7
21. Sundaham KM, Shreehan MM, Olszewski EF (2005) Ethylene. In: Seidel A (ed) Kirk-Othmer encyclopedia of chemical technology, vol 10, 5th edn. Wiley, Hoboken, pp 598–632
22. Devanney MT (2009) Ethylene In: CEH Marketing Research Report, SRI Consulting, Menlo Park, CA, USA, pp 6–17
23. Devanney MT (2008) Propylene. In: CEH Marketing Research Report, SRI Consulting, Menlo Park, CA, USA, pp 6–14
24. Ethylene. (2010) In: Hydrocarbon processing's petrochemical processes 2010, Gulf Publishing Company, Houston. p. 121–130
25. Chauvel A, Lefebvre G (1989) Petrochemical processes, 1. Synthesis-gas derivatives and major hydrocarbons. Gulf Publishing, Houston, pp 117–155
26. Propylene. (2010) In: Hydrocarbon processing's petrochemical processes 2010. Gulf Publishing Company, Houston. p. 245–248
27. Propylene. (2010) In: Hydrocarbon processing's petrochemical processes 2010. Gulf Publishing Company, Houston, TX. p. 249–250,252
28. MacDougall LV (1991) Methanol to fuels routes—the achievements and remaining problems. Catal Today 8:337–369
29. Park T-Y, Froment GF (2001) Kinetic modeling of the methanol to olefins process. 1. model formulation. Ind Eng Chem Res 40:4172–4186
30. Park T-Y, Froment GF (2004) Analysis of fundamental reaction rates in the methanol-to-olefins process on ZSM-5 as a basis for reactor design and operation. Ind Eng Chem Res 43:682–689
31. Alwahabi SM, Froment GF (2004) Single event kinetic modeling of the methanol to olefins process on SAPO-34. Ind Eng Chem Res 43:5098–5111
32. Propylene. (2010) In: Hydrocarbon processing's petrochemical processes 2010, Gulf Publishing Company, Houston, TX. p. 251,253
33. Liebner W (2005) Lurgi MTP® technology. In: Meyers RA (ed) Handboook of petrochemicals production processes. McGraw-Hill, New York, pp 10.3–10.14
34. Pujado PR, Andersen JM (2005) UOP/Hydro MTO process. In: Meyers RA (ed) Handboook of petrochemicals production processes, McGraw-Hill, New York, pp 10.15–10.26
35. World's largest ethylene-from-ethanol plant in Brazil. Plastics today, 1 Oct 2010. http://www.plasticstoday.com/articles/worlds-largest-ethylene-ethanol-plant-brazil
36. Dow chemical explores methods for producing its key feedstocks from renewable resources. ICIS chemical business, 28 Feb 2011. http://www.icis.com/Articles/2011/02/28/9438198/dow-studies-bio-based-propylene-routes.html
37. Wells GM (1995) Ethylene from ethyl alcohol by dehydration, Handbook of petrochemicals and processes. Gower Publishing Company, Brookfield, pp 173–174
38. Aromatization, Benzene, Benzene and Toluene, and BTX Aromatics. (2010) In: Hydrocarbon processing's petrochemical processes 2010, Gulf Publishing Company, Houston, TX, pp. 70, 72–76, 79–83
39. Franck H.-G, Stadelhofer AW (1998) Industrial aromatic chemistry. Springer, Heidelberg, pp. 27–98, 120–31
40. Folkins HO (2003) Benzene. In: Ullmann's encyclopedia of chemical technology, vol 4, 6th edn, Wiley-VCH Verlag GmbH, Weinheim, pp.720–729
41. Fruscella W (2004) Benzene. In: Seidel A (ed) Kirk-Othmer encyclopedia of chemical technology, vol 5, 3rd edn. Wiley, Hoboken, pp 596–624
42. Chauvel A, Lefebvre G (1989) Petrochemical processes, 1. Synthesis-gas derivatives and major hydrocarbons. Gulf Publishing, Houston, pp 235–300
43. Paraxylene and Xylene Isomerization. In: Hydrocarbon processing's petrochemical processes 2010, Gulf Publishing Company, Houston, TX. 2010, pp. 192–196, 284–288
44. Benzene/Toluene, PERP 06/07-6, Chemsystems PERP Program, Nexant, White Plains, New York, June, 2007
45. Aromatics, Transalkylation and Benzene, Ethylbenzene Dealkylation, and BTX Aromatics. In: Hydrocarbon processing's petrochemical processes 2010, Gulf Publishing Company, Houston, TX. 2010, pp. 71, 77
46. Stern DL, Brown SH, Beck JS (2008) Isomerization and transalkylation of alkylaromatics. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis, vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, pp 3168–3194
47. Bradley TW (2005) ExxonMobil PxMax^{SM} p-Xylene from toluene, chapter 13.1 and ExxonMobil XyMax^{SM} xylene isomerization, chapter 13.2. In: Meyers RA (ed) Handboook of petrochemicals production processes, McGraw-Hill, New York, pp. 13.3–13.21
48. Aromatics Extraction and Aromatics Extractive Distillation. In: Hydrocarbon processing's petrochemical processes 2010, Gulf Publishing Company, Houston, TX, 2010, pp. 65–69
49. Franck H-G, Stadelhofer AW (1998) Industrial aromatic chemistry. Springer, Heidelberg, pp 105–121
50. Folkins HO (2003) Benzene. In: Ullmann's encyclopedia of chemical technology. vol 4, 6th edn. Wiley-VCH Verlag GmbH, Weinheim, pp. 730–738
51. Paraxylene and Paraxylene, Crystallization. In Hydrocarbon processing's petrochemical processes 2010, Gulf Publishing Company, Houston, TX, 2010, pp. 192, 195–200
52. Chauvel A, Lefebvre G (1989) Petrochemical processes, 1. Synthesis-gas derivatives and major hydrocarbons. Gulf Publishing, Houston, pp 238–273
53. Commissaris SE (2005) UOP Parex^{TM} process for p-Xylene production, chapter 13.3. In: Meyers RA (Ed) Handboook of petrochemicals production processes. McGraw-Hill, New York, pp 13.23–3.30
54. Ormonde E, Yoneyama M (2009) Butanes. In: CEH Marketing Research Report, SRI Consulting, Menlo Park, CA, pp. 10–13
55. Butadiene/Butylenes. PERP 09/10-5, Chemsystems PERP Program, Nexant, White Plains, New York. (2010)
56. Davis S, Yoneyama M (2008) Butylenes. In: CEH Marketing Research Report, SRI Consulting, Menlo Park, CA, pp 7–20
57. Calamur N, Carrera ME, Wilsak RA (2004) Butylenes. In: Seidel A (ed) Kirk-othmer encyclopedia of chemical technology, vol 4, 5th edn. Wiley, Hoboken, pp 402–433
58. Chauvel A, Lefebvre G (1989) Petrochemical processes, 1. Synthesis-Gas Derivatives and Major Hydrocarbons. Gulf Publishing, Houston, TX, pp 195–226

59. Butene-1, Butene-1, Polymerization Grade, Butenes (Extraction from Butanes/Butenes), and Isomerization. In: Hydrocarbon processing's petrochemical processes 2010, Gulf Publishing Company, Houston, TX. 2010, pp 91, 93, 94, 154
60. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 67–74
61. Brummer R (2005) BASF butadiene extraction technology, Chapter 3.1. In: Meyers RA (ed) Handboook of petrochemicals production processes. McGraw-Hill, New York, New York, USA, p 3.3–3.9
62. Butadiene extraction, BASF NMP process. http://www.lurgi.com/website/Butadiene.55.0.html?&L=1
63. Chauvel A, Lefebvre G (1989) Petrochemical processes, 1. Synthesis-gas derivatives and major hydrocarbons. Gulf Publishing, Houston, TX, USA, p 199–208
64. Butadiene. In: Hydrocarbon processing's petrochemical processes 2010, Gulf Publishing Company, Houston, TX. 2010, pp. 87–89
65. Krupa S, Foley T, McColl S (2005) UOP KLP 1,3-butadiene from acetylene process, Chapter 3.2. In: Meyers RA (ed) Handboook of petrochemicals production processes. McGraw-Hill, New York, New York, USA, p 3.11–3.13
66. Jeanneret JJ, Mowry JR (1997) UOP sorbex family of technologies, Chapter 10.3. Cusher, NA. UOP IsoSiv Process—Chapter 10.5. Sohn SW. Kerosene IsoSiv process for production of normal paraffins, Chapter 10.6. Sohn SW. UOP molex process for production of normal paraffins—Chapter 10.7. Sohn SW. UOP olex process for olefin recovery in handboook of petrochemicals production processes. 2nd Ed. Meyers RA. Editor-in-chief, McGraw-Hill, New York, NY. p.10.45-10.52, 10.61-10.81
67. Vora BP, Peterson GA, Sohn, SW, Riley MG (2009) Detergent alkylate and detergent olefins production. In: Zoller U (ed) Handbook of Detergents, Part F: Production. CRC Press, Taylor & Francis, LLC, Boca Raton, pp 45–47
68. Modler R, Kalin T, Gubler R, Inoguchi Y (2009) Normal paraffins (C_9-C_{17}). In: CEH Marketing Research Report. SRI Consulting, Menlo Park, CA, USA, p 5–16
69. Higman C, van der Burgt M (2003) Gasification. Gulf professional publishing, Elsevier science, USA, pp. 41–146, 298–328
70. Hiller H, et al. (2003) Gas production. In: Ullmann's encyclopedia of chemical technology. vol 21, 6th edn. Wiley-VCH Verlag GmbH, Weinheim, pp. 305–470
71. Shadle LJ, Berry DA, Syamlal M (2004) Coal Gasification. In: Seidel A (ed) Kirk-Othmer Encyclopedia of Chemical Technology, Vol. 4, 5th Edn., John Wiley & Sons, Inc., Hoboken, NJ, p. 771–832
72. Bartholomew CH, Farrauto RJ (2006) Production of hydrogen and synthesis gas via steam reforming (section 6.2). In: Fundamentals of industrial catalytic processes, 2nd edn. Wiley, Hoboken, pp. 342–371
73. Aasberg-Petersen K et al (2001) Technologies for large-scale gas. Appl Catal Gen 221:379–387
74. Jones RM, Shilling NZ (2004) The impact of gas turbine fuel flexibility on IGCC growth. IChemE 6th European Gasification Conference. pp.1–10
75. Heaven DL (1996) Gasification converts a variety of problematic feedstocks and wastes. Oil Gas J 94:49–54
76. Chauvel A, Lefebvre G (1989) Petrochemical processes, 1. Synthesis-Gas Derivatives and Major Hydrocarbons. Gulf Publishing, Houston, pp 19–64
77. Kohl AL (1997) Sulfur recovery processes, chapter 8. Liquid phase oxidation processes for hydrogen sulfide removal, chapter 9. Thermal and catalytic convesion of gas impurities, chapter 13. Physical solvents for acid gas removal. Gas purification, 5th edn. Gulf Publishing Company, Houston, pp. 670–885, 1136–1237
78. Wender I (1996) Reactions of synthesis gas. Fuel Process Technol 48:189–297
79. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 91–98
80. Wells GM (1995) Acetylene, Handbook of petrochemicals and processes. Gower Publishing Company, Brookfield, pp 31–38
81. Chauvel A, Lefebvre G (1989) Petrochemical processes. 1. Synthesis-Gas Derivatives and Major Hydrocarbons. Gulf Publishing, Houston, pp 180–184
82. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 75–77
83. Greiner EOC, Inoguchi Y (2010) Linear alpha-olefins. In: CEH Marketing Research Report, SRI Consulting, Menlo Park, CA, pp. 10–17
84. Calderazzo F, Catellani M, Chiusoli GP (2006) Carbon-carbon bond formation(chapter 5). In: Chiusoli GP, Maitlis PM (eds) Metal-catalysis in industrial organic processes. RSC Publishing, Cambridge, pp 187–192
85. Mol JC (2004) Industrial applications of olefin metathesis. J Mol Catal 213:39–45
86. Banks RL (1984) Olefin metathesis: technology and application. Appl Ind Catal 3:215–239
87. Newman TH, Rand CL, Burdett KA, Maughon RR, Morrison DL, Wasserman, EP (2006) Dow chemical. US. Patent 7,119,216
88. Bielawski CW, Grubbs RH (2007) Living ring-opening metathesis polymerization. Prog Polym Sci 32:1–29
89. Michelotti FW (1967) Interchemical Corp. US. Pat. 3,336,275, 15Aug 1967
90. Michelotti FW, Keaveney WP (1965) Coordinated polymerization of the bicyclo-[2.2.1]-heptene-2 ring system (Norbornene) in polar media. J Polym Sci A 3:895–905
91. Vogt D (1998) SHOP process. In: Cornils B, Herrmann W (eds) Aqueous-phase organometallic catalysis. Wiley-VCH Verlag, Weinheim, pp 541–547
92. Clark MC, Smith CM, Stern DL, Beck JS (2008) Alkylation of aromatics. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis. vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, pp 3153–3168
93. Stefanidakis G, Gwyn JE (1993) Alkylation. In: McKetta JJ (ed) Chemical processing handbook. Marcel Dekker, New York, pp 110–124
94. Franck H-G, Stadelhofer AW (1998) Industrial aromatic chemistry. Springer, Heidelberg, pp 133–137
95. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 335–339
96. Clark MC, Smith CM, Stern DL, Beck JS (2008) Alkylation of aromatics. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis. vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, pp 3156–3157
97. Ethylbenzene. In: Hydrocarbon Processing's Petrochemical Processes 2010, Gulf Publishing Company, Houston, TX, 2010. pp 118–120
98. Ethylbenzene, Part 5. In Meyers RA(Ed) Handboook of petrochemicals production processes. . McGraw-Hill, New York, NY, 2005. pp 5-1 to 5–35
99. Clark MC, Smith CM, Stern DL, Beck JS (2008) Alkylation of aromatics. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis, vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, Germany, p 3157–3160
100. Cumene. In: Hydrocarbon Processing's Petrochemical Processes 2010, Gulf Publishing Company , Houston, TX, 2010. p. 102
101. Franck H-G, Stadelhofer AW (1998) Industrial aromatic chemistry. Springer, Heidelberg, pp 146–148
102. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, New York, USA, p 342–343
103. Cumene, Part 4. In: Meyers RA (ed) Handboook of petrochemicals production processes, McGraw-Hill, New York, New York, USA, 2005, p 4-1–4-19

104. Clark MC, Smith CM, Stern DL, Beck JS (2008) Alkylation of aromatics. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis, vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, Germany, p 3161
105. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 343–345
106. Vora BP, M.C, Peterson GA, Sohn SW, Riley MG (2009) Detergent alkylate and detergent olefins production. In: Zoller U (ed) Handbook of detergents, part F: production. CRC Press, Taylor & Francis Group, LLC, Boca Raton, pp. 39–45
107. Linear Alkylbenzenes (LAB) (2009) PERP 07/08S7, Chemsystems PERP Program, Nexant, White Plains, New York. Jan 2009
108. Haber J et al (2008) Selective oxidations (section 14.11). In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis, vol 8, 2nd edn. Weinheim, Wiley-VCH Verlag GmbH, pp 3359–3547
109. Hodnett BK (2000) Heterogeneous catalytic oxidation. Wiley, Chichester
110. Bartholomew CH, Farrauto RJ (2006) Catalytic oxidations of inorganic and organic compounds (chapter 8). In: Fundamentals of industrial catalytic processes, 2nd edn. Wiley, Hoboken. pp. 560–562, 578–584
111. Bielanski A, Haber J (1991) Oxygen in catalysis. Marcel Dekker, New York
112. Haber J (2003) Selective oxidation—heterogeneous. In: Horvath LT (ed) Encyclopedia of catalysis, vol 6. Wiley, Hoboken, pp 141–189
113. Sheldon RA, Ten Brink GJ, Arends IWCE (2003) Selective oxidation—homogeneous. In: Horvath LT (ed) Encyclopedia of catalysis, vol 6. Wiley, Hoboken, pp 189–239
114. Centi G, Perathoner S (2003) Selective oxidation—industrial. In: Horvath LT (ed) Encyclopedia of catalysis, vol 6. Wiley, Hoboken, pp 239–299
115. Clerici MG, Ricci M, Strukul G (2006) Formation of C–O bonds by oxidation (chapter 2). In: Chiusoli GP, Maitlis PM (eds) Metal-catalysis in industrial organic processes. RSC Publishing, Cambridge, pp 23–78
116. Doraiswamy LK (1984) Heterogeneous reactions: analysis, examples, and reactor designs, vol 1–2. Wiley, Hoboken
117. Rase HF (1977) Chemical reactor design for process plants, vol 1–2. Wiley Interscience, New York
118. Suresh AK, Sharma MM, Sridhar T (2000) Engineering aspects of industrial liquid-phase air oxidation of hydrocarbons. Ind Eng Chem Res 39:3958–3997
119. Coward HF, Jones GW (1952) Limits of flammability of gases and vapors, bulletin 503 bureau of mines. United States Government Printing Office, Washington, DC, USA
120. Stull DR (1977) Fundamentals of Fire and Explosion, AIChE Monograph Series, AIChE, Vol 73(10), New York, NY
121. Medard LA (1989) Ignition temperatures of explosive gaseous mixtures, Chapter 7. Flammability limits of explosive gaseous mixtures, Chapter 8. Flammability limits: safety applications, Chapter 10. In: Accidental explosions, vol 1. Wiley, Hoboken. pp 147–218
122. Lees FP (1996) Loss prevention in the process industries. vols 1–3, 2nd Edn., Elsevier, New York, NY, 1996.
123. Bodurtha FT (1980) Industrial explosion prevention and protection. McGraw-Hill, New York
124. Sheehan RJ (2003) Terephthalic acid, Dimethyl terephthalate, and Isophthalic acid. in Ullmann's encyclopedia of chemical technology. Vol 35, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p.639–651
125. Castillo-Walter F (2005) E PTA: The Lurgi/Eastman/SK Process, Chapter 12.1. In: Meyers RA (editor-in-chief) Handboook of petrochemicals production processes. McGraw-Hill, New York, NY. p.12.3–12.12
126. Rosowski F, Storck S, Zuhlke J (2008) Oxyfunctionalization of alkyl aromatics. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis, vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, pp 3425–3433
127. Partenheimer W (1995) Methodology and scope of metal/bromide autoxidation of hydrocarbons. Catal Today 23:69–158
128. Maki T, Takeda K (2003) Benzoic acid and derivatives. in Ullmann's encyclopedia of chemical technolog. Vol 5, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p.59–72
129. Bizzari SN, Blagoev M, Kumamoto T (2010). Benzoic acid. CEH marketing research report, SRI consulting, Menlo Park, CA. 2010, p. 10–11
130. Phthalic acid and derivatives. Industrial organic chemicals. Vol 7, Wiley-VCH Verlag GmbH, Weinheim. 1999, p. 3868–3877
131. Bizzari SN (2010) Phthalic anhydride. CEH Marketing Research Report SRI Consulting, Menlo Park, CA, USA, pp 13–14
132. Phthalic anhydride. ICIS chemical business. Feb 8–14, 2010, p. 35
133. Centi G, Perathoner S (2003) Selective oxidation—industrial. In: Horvath LT (ed) Encyclopedia of catalysis, vol 6. John Wiley & Sons, Inc, Hoboken, NJ, pp 294–296
134. Shu C-M, Wen P-J, Chang R-H (2002) Investigations on flammability models and zones for o-xylene under various initial pressures, temperatures, and oxygen concentrations. Thermochim Acta 392–393:271–287
135. Ghirardini M, Tampieri, Polimeri Europa Cumene-Phenol Processes, Chapter 9.1. Schmidt, R.J., "Sunoco/UOP Phenol Process, Chapter 9.2"; Moore, A.; Birkhoff, R., "KBR Phenol Process, Chapter 9.3", in *Handboook of Petrochemicals Production Processes,* Meyers, R.A., Editor-in-chief, McGraw-Hill, New York, New York, USA, 2005, pp.9.3 to 9.12, 9.13 to 9.29, 9.31 to 9.50.
136. Phenol. In Hydrocarbon processing's petrochemical processes 2010. Gulf Publishing Company, Houston, TX. 2010, p. 203–205
137. Sifniades S, Levy AB (2003) Acetone. in Ullmann's encyclopedia of chemical technology, Vol 1, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p.197–200
138. Phenol/Acetone/Cumene. PERP 05/06-4, Chemsystems PERP program, Nexant, White Plains, NY. Mar 2007
139. Jordan W, van Barneveld H, Gerlich O, Kleine-Boymann M, Ullrich J (2003) Phenol. in Ullmann's encyclopedia of chemical technology, Vol 25, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p.589–604
140. Schmidt RJ (2005) Industrial catalytic processes—phenol production. Appl Catal Gen 280:89–103
141. Clerici MG, Ricci M, Strukul G (2006) Formation of C–O bonds by oxidation, Chapter 2. In: Chiusoli GP, Maitlis PM (eds) Metal-catalysis in industrial organic processes. RSC Publishing, Cambridge, pp 39–40
142. James DH, Castor WM (2003) Styrene. in Ullmann's encyclopedia of chemical technology. Vol 34, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p. 390
143. Kahlich D, Wiechern U, Lindner J (2003) Propylene oxide. in Ullmann's encyclopedia of chemical technology. Vol 30, 6th Ed. Wiley-VCH Verlag GmbH, Weinheim. p.279–303
144. Clerici MG, Ricci M, Strukul G (2006) Formation of C–O bonds by oxidation, Chapter 2. In: Chiusoli GP, Maitlis PM (eds) Metal-catalysis in industrial organic processes. RSC Publishing, Cambridge, pp 35–39
145. Kobe JM, Evans WE, June RL, Lemanski MF (2003) Epoxidation-industrial. In: Horvath LT (ed) Encyclopedia of catalysis, vol 3. Wiley, Hoboken, NJ, USA, p 239–257
146. Rebsdat S, Mayer D (2003) Ethylene oxide. in Ullmann's encyclopedia of chemical technology, Vol 12, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p. 609–635

147. Ethylene oxide/EG. PERP 08/09-2. Chemsystems PERP Program, Nexant, White Plains, New York. Feb 2010.
148. Bartholomew CH, Farrauto RJ (2006) Ethylene to ethylene oxide, Section 8.4.3. In: Fundamentals of industrial catalytic processes, 2nd edn. Wiley, Hoboken, NJ, USA, p 597–604
149. Linic S, Barteau MA (2008) Heterogeneous catalysis of alkene epoxidation. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogeneous catalysis, vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, Germany, p 3449–3456
150. Schroeder V, Baumeier A, Franzen S, Buettgen F (2007) Explosion limits of decomposition of mixtures of ethylene oxide, propylene oxide, and nitrogen. Chemie Ing Tech 79(8):1241–1245
151. Pekalski AA, Zevenbergen JF, Braithwaite M, Lemkowitz SM, Pasman HJ (2005) Explosive decomposition of ethylene oxide at elevated conditions: effect of ignition energy, nitrogen dilution, and turbulence. J Haz Mat 118(1–3):19–34
152. Kohl AL, Nielsen RB (1997) Alkaline salt solutions for acid gas removal. In: Gas purification, 5th edn. Gulf Publishing Company, Houston, TX, USA, p 330–414
153. Monnier JR, Stavinoha JL, MInga RL (2004) Stability and distribution of cesium in Cs-promoted silver catalysts used for butadiene epoxidation. J Catal 226(2):401–409
154. Fleischmann G, Jira R, Bolt HM, Golka K (2003) Acetaldehyde. in Ullmann's encyclopedia of chemical technology, Vol 1, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p. 131–148
155. Jira R (2002) Oxidation of olefins to carbonyl compounds (Wacker process). Section 2.4.1. In: Cornils B, Herrmann WA (eds) Applied homogenous catalysis with organometallic compounds, vol 1, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, pp 386–412
156. Miller SA (1969) Ethylene and its industrial derivatives. Ernest Benn Limitied, London, England, pp 480–530
157. Centi G, Perathoner S (2003) Selective oxidation—industrial. In: Horvath LT (ed) Encyclopedia of catalysis, vol 6. John Wiley & Sons, Inc, Hoboken, NJ, pp 265–269
158. Wells GM (1995) Acetaldehyde. Handbook of petrochemicals and processes. Gower Publishing Company, Brookfield, Vermont, pp 1–8
159. Goe, G.L (1982) Pyridine and pyridine derivatives. In: Mark H (editor-in-chief) Kirk-othmer encyclopedia of chemical technology. Vol. 19, 3rd Edn., John Wiley & Sons, Inc., USA, p. 454–483
160. Grate JH, Hamm DR, Mahajan S (1993) Palladium and phosphomolybdovanadate catalyzed olefin oxidation to carbonyls. Molecular Engineering 3(1–3):205–229
161. Lambert A, Derouane EG, Kozhevnikov IV (2002) Kinetics of one-stage wacker-type oxidation of C2-C4 olefins by an aqueous PdCl2 heteropoly-anion system. J Catalysis 211(2):445–450
162. Ohara T, Sato T, Shimizo N, Prescher G, Schwind H, Weiberg O, Marten K (2003) Acrylic acid and derivatives. in Ullmann's encyclopedia of chemical technology. Vol 1, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p. 289–306
163. Acrylic Acid. PERP 08/09-3. Chemsystems PERP Program, Nexant, White Plains, New York. Jul 2010, p. 18–33
164. Bauer W (2003) Methacrylic acid and derivatives. in Ullmann's encyclopedia of chemical technology. Vol 21, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p. 585–597
165. Wells GM (1995) Acrolein, Acrylic acid. Handbook of petrochemicals and processes. Gower Publishing Company, Brookfield, Vermont, pp 39–47
166. Adipic acid. PERP 08/09-2, Chemsystems PERP Program, Nexant, White Plains, New York. Feb 2010
167. Clerici MG, Ricci M, Strukul G (2006) Formation of C–O bonds by oxidation, Chapter 2. In: Chiusoli GP, Maitlis PM (eds) Metal-catalysis in industrial organic processes. RSC Publishing, Cambridge, pp 29–30
168. Malveda M.P, Mori H (2009) Cyclohexanol and cyclohexanone. CEH Marketing research report. SRI Consulting, Menlo Park, CA. p. 5, 8
169. Lohbeck K, Haferkorn H, Fuhrmann W, Fedtke N (2003) Maleic and fumaric acids. in Ullmann's encyclopedia of chemical technology. Vol 20, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p. 463–473
170. Greiner EOC, Funada C (2009) Maleic anhydride. CEH marketing research report. SRI Consulting, Menlo Park, CA, pp 7–11
171. Bartholomew CH, Farrauto RJ (2006) n-Butane to Maleic Anhydride, Section 8.4.5. in Fundamentals of industrial catalytic processes. 2nd Edn. John Wiley & Sons, Hoboken, NJ. p. 610–618
172. Maleic anhydride. in Hydrocarbon Processing's Petrochemical Processes 2010. Gulf Publishing Company, Houston, TX. 2010, p. 158–160
173. Maleic anhydride. PERP 07/08-8, Chemsystems PERP Program, Nexant, White Plains, New York. Feb 2009 p. 12.
174. Reuss G, Disteldorf W, Gamer AO, Hilt A (2003) Formaldehyde. in Ullmann's encyclopedia of chemical technology, Vol 15, 6th Ed. Wiley-VCH Verlag GmbH, Weinheim. p.1–34
175. Bizzari SN (2010) Formaldehyde. CEH Marketing Research Report. SRI Consulting, Menlo Park, CA. p. 5, 14–15
176. Formaldehyde. ICIS chemical business. March 31-April 6, 2008, p. 56
177. Formaldehyde and derivatives. PERP 04/05S10, Chemsystems PERP Program, Nexant, White Plains, NY. 2006
178. Davies P, Donald RT, Harbord NH (1989) Methanol oxidation, Section 10.3. In: Twigg MV (ed) Catalyst Handbook, 2nd edn. Wolfe Publishing, Ltd., England, pp 490–503
179. Bartholomew CH, Farrauto RJ (2006) Methanol to formaldehyde, Section 8.4.2, in Fundamentals of industrial catalytic processes. 2nd Edn. John Wiley & Sons, Hoboken, NJ. p. 584–597
180. Maurer G (1986) Vapor–liquid equilibrium of formaldehyde and water containing multicomponent mixtures. AIChE J 32 (6):932–948
181. Ott M, Fischer HH, Maiwald M, Albert K, Hasse H (2005) Kinetics of oligmoerization reactions of formaldehyde solutions. Chem Eng Process 44:653–660
182. Roscher G (2003) Vinyl esters. in Ullmann's encyclopedia of chemical technology, Vol 38, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p.59–75
183. Haber J (2003) Selective oxidation—heterogeneous. In: Horvath LT (ed) Encyclopedia of catalysis, vol 6. John Wiley & Sons, Inc, Hoboken, NJ, pp 283–284
184. Wells GM (1995) Vinyl acetate. Handbook of petrochemicals and processes. Gower Publishing Company, Brookfield, Vermont, pp 363–367
185. Clerici MG, Ricci M, Strukul G (2006) Formation of C–O bonds by oxidation, chapter 2. In: Chiusoli GP, Maitlis PM (eds) Metal-catalysis in industrial organic processes. RSC Publishing, Cambridge, pp 70–73
186. Schunk SA, de Oliveria AL (2008) Acetoxylation of ethylene. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis, vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, pp 3464–3478
187. Koch F (2002) Aliphatic carboxylic acids via aliphatic aldehydes. In: Cornils B, Herrmann WA (eds) Applied homogenous catalysis with organometallic compounds, vol 1, 2nd edn. Weinheim, Wiley-VCH Verlag GmbH, pp 427–432
188. Cheung H, Tanke RS, Torrence GP (2003) Acetic acid. in Ullmann's encyclopedia of chemical technology, Vol 1, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p.149–178
189. Howard MJ, Jones MD, Roberts MS, Taylor SA (1993) C_1 to acetyls: catalysis and process. Catal Today 18:325–354

190. Yoneda N, Kusano S, Yasui M, Pujado P, Wilcher S (2001) Recent advances in processes and catalysts for production of acetic acid. Appl Catal Gen 221:253–265
191. Zoeller JR (1993) Manufacture via methanol carbonylation. In: Agreda VH, Zoeller JR (eds) Acetic acid and its derivatives, handbook, 7th edn. Marcel Dekker, Inc., New York, NY, pp 35–51
192. Yoneda N, Hosono Y (2004) Acetic acid process catalyzed by ionically immobilized rhodium complex to solid resin support". J ChE Japan 37(4):536–545
193. Acetic acid. in Hydrocarbon processing's petrochemical processes 2010. Gulf Publishing Company, Houston, TX. 2010, p. 43
194. Held H, Rengstl A, Mayer D (2003) Acetic anhydrides and mixed fatty acid anhydrides. in Ullmann's encyclopedia of chemical technology, Vol 1, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p.179–194
195. Zoeller JR, Agreda VH, Cook SL, Lafferty NL, Polichnowski SW, Pond DM (1992) Eastman chemical company acetic anhydride process. Cataly Today 13:73–91
196. Malveda MP, Funada C (2010) Acetic anhydride, CEH marketing research report. SRI Consulting, Menlo Park, CA, pp 4–9
197. Reutemann W, Kieczka H (2003) Formic Acid. in Ullmann's encyclopedia of chemical technology, Vol 15, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p.49–71
198. Horsley LH (1973) Azeotrope data for formic acid-water, Azeotrope Data –III, Adv. In Chemistry Sereies 116, ACS, Washington, DC. p.13
199. Clegg William, Eastham GR, Elsegood MRJ, Tooze RP, Wang XL, Whitson K. J Chem Soc, Chem Commun. 1877 (1999)
200. Eastham GR, Heaton BT, Iggo JA, Tooze RT, Whyman R, Zacchini S. J Chem Soc, Chem Commun. 609 (2000)
201. Tooze RP, Eastham GR, Whitson K, Wang XL. J Chem Soc, U.S. Pat. 6,348,621 (Feb. 19, 2002)
202. Kohlpaintner C (2003) Hydroformylation—industrial. In: Horvath LT (ed) Encyclopedia of catalysis, vol 3. John Wiley & Sons, Inc, Hoboken, NJ, pp 787–808
203. Bahrmann H, Bach H (2003) Oxo synthesis. in Ullmann's encyclopedia of chemical technology, Vol 24, 6th Edn.Wiley-VCH Verlag GmbH, Weinheim. p.553–560
204. Oxo alcohols. PEP Report 21E. SRI Consulting, The Woodlands, TX. Sep 2010
205. Oxo alcohols. PERP 06/07-8, Chemsystems PERP Program. Nexant, White Plains, NY. 2007
206. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, NY, pp 129–133
207. Maitlis PM, Haynes A (2006) Syntheses based on carbon monoxide, chapter 4. In: Chiusoli GP, Maitlis PM (eds) Metal-catalysis in industrial organic processes. RSC Publishing, Cambridge, pp 141–150
208. Natta G, Ercoli R, Castellano S, Barbieri FH (1954) The influence of hdrogen and carbon monoxide partial pressures on the rate of the hydroformylation reaction. JACS 76:4049–4050
209. Yang C, Mao Z-S, Wang Y, Chen J (2002) Kinetics of hydroformylation of propylene using $RhCl(CO)(TPPTS)_2$/TPPTS complex catalyst in aqueous system. Catal Today 74:111–119
210. Puckette TA (2007) Halophosphite ligands for the rhodium-catalyzed low-pressure hydroformylation reaction. In: Schmidt SR (ed) Catalysis of organic reactions. CRC Press, LLC, Boca Raton, FL, pp 31–38
211. Butcher, J.; Reynolds, G (2005) Johnson matthey oxo alcohols process™. In: Meyers RA (editor-in-chief) Handbook of petrochemicals production processes. McGraw-Hill, New York, NY. p.8.3–8.13
212. Lee R, Lee AO (1993) Chlorination, liquid phase and vapor phase. In: McKetta JJ (ed) Chemical processing handbook. Marcel Dekker, Inc., New York, NY, pp 295–352
213. McBee ET, Hass HB, Bordenca C (1943) Chlorinolysis of chloroparaffins. Ind Eng Chem 35(3):317–320
214. Smith MB, March J (2001) March's advanced organic chemistry, 5th edn. John Wiley & Sons, Inc, New York, NY, pp 894–911
215. Smart, B.E (1980) Fluorine Compounds, Organic. In: Mark HF (editor-in-chief). Kirk-Othmer Encyclopedia of Chemical Technology. Vol. 10, 3rd Edn. John Wiley & Sons, Inc, Hoboken, NJ. p. 829–870
216. DeForest EM (1993) Chloromethanes. In: McKetta JJ (ed) Chemical processing handbook. Marcel Dekker, Inc., New York, NY, pp 427–482
217. Rossberg M, Lendle W, Pfleiderer G, Togel A, Dreher E.-L, Rassaerts H, Kleinschmidt P, Strack H, Cook R, Beck U, Lipper K.-A, Torkelson TR, Loser E, Beutel KK (2003) Chloronated hydrocarbons, 3.Chloroethanes. in Ullmann's encyclopedia of chemical technology. Vol 8, 6th Ed. Wiley-VCH Verlag GmbH, Weinheim. p.3–25
218. Wells, G.M (1995) Carbon Tetrachloride, Chloroform, Methyl Chloride, Methylene Chloride. Handbook of petrochemicals and processes. Gower Publishing Company, Brookfield, Vermont. p. 115–119, 124–128, 246–254.
219. Kurtz BE, Smalley EW (1993) Chlorobenzene and dichlorobenzene. In: McKetta JJ (ed) Chemical processing handbook. Marcel Dekker, Inc., New York, NY, pp 483–500
220. Rossberg M, Lendle W, Pfleiderer G, Togel A, Dreher E.-L, Rassaerts H, Kleinschmidt P, Strack H, Cook R, Beck U, Lipper K.-A, Torkelson TR, Loser E, Beutel KK (2003) Chloronated hydrocarbons, 1.Chloroethanes. in Ullmann's encyclopedia of chemical technology. Vol 8, 6th Ed. Wiley-VCH Verlag GmbH, Weinheim. p.107–116
221. Wells GM (1995) Chlorobenzene. Handbook of petrochemicals and processes. Gower Publishing Company, Brookfield, Vermont, pp 120–123
222. Rossberg M, Lendle W, Pfleiderer G, Togel A, Dreher E.-L, Rassaerts H, Kleinschmidt P, Strack H, Cook R, Beck U, Lipper K.-A, Torkelson TR, Loser E, Beutel KK (2003) Chloronated hydrocarbons, 3.Chloroethanes. in Ullmann's encyclopedia of chemical technology. Vol 8, 6th Ed. Wiley-VCH Verlag GmbH, Weinheim. p.26–32
223. Vinyl Chloride/EDC. PERP 08/09-4, Chemsystems PERP Program. Nexant, White Plains, NY, Sep 2009
224. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, NY, pp 217–221
225. Wells GM (1995) Ethylene dichloride. Handbook of petrochemicals and Processes. Gower Publishing Company, Brookfield, Vermont, pp 175–182
226. Jordan JI (1993) Chloroethanes. In: McKetta JJ (ed) Chemical processing handbook. Marcel Dekker, Inc., New York, NY, pp 356–381
227. Rossberg M, Lendle W, Pfleiderer G, Togel A, Dreher E.-L, Rassaerts H, Kleinschmidt P, Strack H, Cook R, Beck U, Lipper K.-A, Torkelson TR, Loser E, Beutel KK (2003) Chloronated hydrocarbons, 2.Chloroethanes. in Ullmann's encyclopedia of chemical technology, Vol 8, 6th Edn. Wiley-VCH Verlag GmbH, Weinheim. p.30–50
228. Glauser J, Ishikawa Y (2009) C_2 Chlorinated Solvents. CEH Marketing Research Report. SRI Consulting, Menlo Park, CA, pp 12–17
229. Kahlich D, Wiechern U, Lindner J (2003) Propylene oxide. in Ullmann's encyclopedia of chemical technology. 6th Edn., Vol 30. Wiley-VCH Verlag GmbH, Weinheim. p.279–303
230. Trent DL (2006) Propylene oxide. In: Seidel A (editor-in-Chief). Kirk-othmer encyclopedia of chemical technology. Vol. 20, 5th Edn. John Wiley & Sons, Inc, Hoboken, NJ. p. 790–822
231. Clerici MG, Ricci M, Strukul G (2006) Is the epoxidation of olefins other than ethylene feasible on silver catalysts, section 2.4.3., in Formation of C-O bonds by oxidation, Chapter 2. In: Chiusoli GP, Maitlis PM (eds) Metal-catalysis in industrial organic processes. RSC Publishing, Cambridge, pp 39–40

232. Rowe CE (1993) Chlorohydrins. In: McKetta JJ (ed) Chemical processing handbook. Marcel Dekker, Inc, New York, NY, pp 382–426
233. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, NY, pp 295–296
234. Braun M (2003) Fundamentals and transition-state models. Aldol additions of group 1 and 2 enolates. In: Mahrwald R (ed) Modern aldol reaction, vol 1. Weinheim, Wiley-VCH Verlag GmbH, pp 1–4
235. Parker SP (editor) McGraw-Hill encyclopedia of chemistry. 2nd Edn. Macmillan Publishing Co., Inc, USA. 1993, p.245–246
236. Streitwieser A, Heathcock CH (1981) Introduction to organic chemistry, 2nd edn. McGraw-Hill, Inc, New York, NY, pp 392–396
237. Muthusamy D, Fisher I (2006) Methyl isobutyl ketone. In: Seidel A (editor-in-chief) Kirk-othmer encyclopedia of chemical technology. Vol. 16, 5th Edn. John Wiley & Sons, Inc, Hoboken, NJ. p. 329–355
238. Simons RM (1989) Methyl isobutyl ketone. In: McKetta JJ, Cunningham WA (eds) Encyclopedia of chemical processing and design, vol 34. Marcel Dekker, New York, NY USA, pp 50–82
239. O'Keefe WK, Ng FTT, Rempel GL (2007) Experimental studies on the syntheses of mesityl oxide and methyl isobutyl ketone via catalytic distillation. Ind Eng Chem Res 46:716–725
240. Werle P, Morawietz M (2003) Alcohols, Polyhydric. in Ullmann's encyclopedia of chemical technology. 6th Edn., Vol 2. Wiley-VCH Verlag GmbH, Weinheim. p.47–64
241. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, NY, pp 210–213
242. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, NY, 138
243. Gautreaux MF, Davis WT, Travis ED (1978) Alcohols, higher aliphatic (Synthetic), in Kirk-othmer encyclopedia of chemical technology. Vol. 1, 3rd Edn. John Wiley & Sons, Inc, New York. p.751
244. Smith MB, March J (2001) The tishchenko reaction, Section 19-61. in March's advanced organic chemistry. 5th Edn. John Wiley & Sons, Inc, New York. p. 1565–66
245. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, NY, pp 186–187
246. Mahrwald R (2003) The aldol-tishchenko reaction. In: Mahrwald R (ed) Modern aldol reactions, vol 1. Wiley-VCH Verlag GmbH, Weinheim Germany, pp 327–329
247. Merger F, Foerster H-J (1983) Preparation of alpha-alkylacroleins. U.S. Patent 4,408,079, 4 Oct 1983
248. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, USA, p 285
249. Grafje H et al (2003) Butanediols, Butenediol, and Butynediol, in *Ullmann's Encyclopedia of Chemical Technology, 6th Ed., Volume 5*, Wiley-VCH Verlag GmbH, Weinheim, Germany, p.703-711.
250. Weissermel K, Arpe H-J (1997) Industrial Organic Chemistry, 3rd edn. VCH Publishers, New York, USA, p 99
251. Franck H-G, Stadelhofer AW (1998) Industrial aromatic chemistry. Springer, Heidelberg, Germany, pp 158–160
252. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, USA, pp 358–359
253. "Phenol Derivatives", *Industrial Organic Chemicals, Volume* *, Wiley-VCH Verlag GmbH, Weinheim, Germany, 1999, pp. 3768–377.
254. Propanols (1999) Industrial organic chemicals, vol 7. Wiley-VCH Verlag GmbH, Weinheim, Germany, pp 4067–4070
255. Sherman PW, Kavasmaneck PR (1980) Ethanol. In: Kirk-Othmer (ed) Encyclopedia of chemical technology, vol 9, 3rd edn. Wiley, New York, USA, p 342–350
256. Linak E, Janshekar H, Inoguchi Y (2009) Ethanol. In: CEH Marketing Research Report, SRI Consulting, Menlo Park, CA, USA p 15–16.
257. Neier W, Strehlke G (2003) 2-Butanone. In: Ullmann's encyclopedia of chemical technology, vol 5, 6th edn, Wiley-VCH Verlag GmbH, Weinheim, Germany, p 726
258. Wells GM (1995) Ethyl alcohol, isopropyl alcohol, Handbook of petrochemicals and processes. Gower Publishing Company, Brookfield, Vermont, USA, 1995, p 154–159, 225–229
259. Streitwieser A, Heathcock CH (1981) Introduction to organic chemistry, 2nd edn. Macmillan, New York, p 306
260. Rebstat S, Mayer D (2003) Ethylene glycol. In: Ullmann's encyclopedia of chemical technology, vol 12, 6th edn. Wiley-VCH Verlag GmbH, Weinheim, Germany, pp 593–608
261. "Ethylene Oxide/Ethylene Glycol", PERP 08/09-8, Chemsystems PERP Program, Nexant, White Plains, New York, USA, December, 2009.
262. Chitwood HC, Freure BT (1946) The reaction of propylene oxide with alcohols. JACS 68(4):680–683
263. Parker RE, Isaacs NS (1959) Mechanisms of epoxide reactions. Chem Rev 4:737–799
264. Groggins PH (1958) Unit processes in organic synthesis, 5th edn. McGraw Hill, New York, pp 843–844
265. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 160–161
266. Levinson MI (2009) Surfactant production: present realities and future perspectives. In: Zoller U (ed) Handbook of detergents, Part F: production, CRC Press, Taylor & Francis Group, LLC, Boca Raton, FL, USA, pp 26–29
267. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 181–183
268. Held H, Rengstl A, Mayer D (2003) Acetic anhydride and mixed fatty acid anhydrides. In: Ullmann's encyclopedia of chemical technology, vol 1, 6th edn. Wiley-VCH Verlag GmbH, Weinheim, Germany, p. 179–194
269. Raynolds PW (1993) Ketene (chapter 10). In: Agreda VH, Zoeller JR (eds) Acetic acid and its derivatives. Marcel Dekker, New York, pp 161–171
270. Wells GM (1995) Handbook of petrochemicals and processes. Gower Publishing Company, Vermont, pp 18–19
271. Clemens RJ, Witzeman JS (1993) Diketene and acetoacetates (chapter 11). In: Agreda VH, Zoeller JR (eds) Acetic acid and its derivatives. Marcel Dekker, New York, pp 173–224
272. Martinez R, Huff MC, Barteau MA (2004) Ketonization of acetic acid on titania-functionalized silica monoliths. J Catal 222(2):404–409
273. Bender ML (1960) Mechanisms of catalysis of nucleophilic reactions of carboxylic acid derivatives. Chem Rev 60(1):53–113
274. Witzeman JS, Agreda VH (1993) Alcohol acetates (chapter 14). In: Agreda VH, Zoeller JR (eds) Acetic acid and its derivatives. Marcel Dekker, New York, pp 257–284
275. Simons RM (1993) Esterification. In: McKetta JJ (ed) Chemical processing handbook. Marcel Dekker, New York, pp 537–658
276. Barnicki SD, Siirola JJ (1997) Enhanced distillation. In: Perry RH, Green DW, Maloney JO (eds) Perry's chemical engineers' handbook, 7th edn. McGraw Hill, New York, pp. 13–83 to 13–85.
277. Groggins PH (1958) Esterification. In: Unit processes in organic synthesis, 5th edn. McGraw Hill Book Company, New York, New York, USA, p 694–749
278. Agreda VH, Partin LR, Heise WH (1990) High-purity methyl acetate via reactive distillation. Chem Eng Prog 86(2):40–46
279. Cadogan DF (1999) Plasticizers. In: Kirk-Othmer (ed) Concise encyclopedia of chemical technology, 4th Edn. Wiley, New York, pp 1577–1578
280. Bartholomew CH, Farrauto RJ (2006) Hydrogenation and dehydrogenation of organic compounds, section 7. In: Fundamentals of

industrial catalytic processes, 2nd edn, Wiley, Hoboken, pp 487–532

281. Gallezot P (2003) Hydrogenation—heterogeneous. In: Horvath LT (ed) Encyclopedia of catalysis, vol 4. Wiley, Hoboken, pp 17–55
282. Rylander PN (2003) Hydrogenation and dehydrogenation. In: Ullmann's encyclopedia of chemical technology, vol 17, 6th edn. Wiley-VCH Verlag GmbH, Weinheim, pp 241–252
283. Chen B, Dingerdissen U, Krauter JGE, Lansink Rotgerink HGJ, Mobus K, Ostgard DJ, Panster P, Riermeier TH, Seebald S, Tacke T, Traauthwein H (2005) New developments in hydrogenation catalysis particularly in synthesis of fine and intermediate chemicals. Appl Catal Gen 280:17–46
284. Barnicki SD (1991) Separation system synthesis: a knowledge-based approach. PhD Dissertation, The University of Texas at Austin, pp. 99–100
285. Fiedler E, Grossmann G, Kiersebohm DB, Weiss G, Witte C (2003) Methanol. In: Ullmann's encyclopedia of chemical technology, vol 1, 6th edn. Wiley-VCH Verlag GmbH, Weinheim, pp 611–635
286. "Methanol Synthesis Catalyst Product Bulletin", Sud-Chemie, 2007, pp 1–16
287. Vanden Bussche KM, Froment GF (1996) A steady-state kinetic model for methanol synthesis and the water gas shift reaction on a commercial Cu/Zn/AL_2O_3 catalyst. J Catal 161:1–10
288. Bridger GW, Spencer MS (1989) Methanol syntheis (chapter 9). In: Twigg MV (ed) Catalyst handbook, 2nd edn. Wolfe Publishing, England, pp 441–468
289. "Methanol", in Hydrocarbon Processing's petrochemical processes 2010, Gulf Publishing Company , Houston, 2010, pp 163–172
290. "Cyclohexane", In: Hydrocarbon processing's petrochemical processes 2010, Gulf Publishing Company, Houston, TX, USA, 2010, p. 103
291. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 345–347
292. Wells GM (1995) Handbook of petrochemicals and processes. Gower Publishing Company, Brookfield, pp 133–136
293. Rylander PN (1985) Hydrogenation methods. Academic, Orlando, pp 66–71
294. Claus P, Onal Y (2008) Regioselective hydrogenations. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis, vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, pp 3313–3326
295. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 134–139
296. Bartholomew CH, Farrauto RJ (2006) Hydrogenation of carbonyl groups (section 7.3.5). In: Fundamentals of industrial catalytic processes, 2nd edn. Wiley, Hoboken, pp. 531–532
297. McCusker-Orth JE, Stavinoha JL, Messina AD, Perri ST, Liu Z, Heidt PC, Tennant, BA (2008) Process for the preparation of tetraalkylcyclobutane-1,3-diols from the corresponding butanediones, using a Ru-promoted Co-based catalyst. U.S. Patent Appl. Publ. 20080154069, June 26 2008
298. Haas T, Jaeger B, Weber R, Mitchell SF, King CF (2005) New diol processes: 1,3-propandiol and 1,4-butanediol. Appl Catal Gen 280:83–88
299. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 370–372
300. Chen B, Dingerdissen U, Krauter JGE, Lansink Rotgerink HGJ, Mobus K, Ostgard DJ, Panster P, Riermeier TH, Seebald S, Tacke T, Traauthwein H (2005) Section 1.2.6. Esters, anhydrides, and carboxylic acids. In: New developments in hydrogenation catalysis particularly in synthesis of fine and intermediate chemicals, Appl Cat A: Gen., vol 280, pp. 22–23.
301. "1,4-Butanediol/THF" PERP 06/07-4, Chemsystems PERP Program, Nexant, White Plains, March, 2008.
302. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 403–404
303. Voge HH (1993) Dehydrogenation. In: McKetta JJ (ed) Chemical processing handbook. Marcel Dekker, New York, pp 541–555
304. Caspary KJ, Gehrke H, Heinritz-Adrian M, Schwefer M (2008) Dehydrogenation of alkanes. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis, vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, pp 3206–3229
305. Bartholomew CH, Farrauto RJ (2006) Dehydrogenation: reaction chemistry; Catalyst and reactor technologies (section 7.5.1). In: Fundamentals of industrial catalytic processes, 2nd edn. Wiley, Hoboken, pp. 533–536
306. Resasco DE (2003) Dehydrogenation—Heterogeneous. In: Horvath LT (ed) Encyclopedia of catalysis. Wiley, Hoboken, pp 49–79
307. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 109–112
308. Caspary KJ, Gehrke H, Heinritz-Adrian M, Schwefer M (2008) Dehydrogenation of alkanes. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis, vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, Germany, p 3221–3224
309. Bartholomew CH, Farrauto RJ (2006) Dehydrogenation of C_3 and C_4 alkanes to alkenes (section 7.5.1). In: Fundamentals of industrial catalytic processes, 2nd edn. Wiley, Hoboken, pp 536–542
310. Caspary KJ, Gehrke H, Heinritz-Adrian M, Schwefer M (2008) Dehydrogenation of alkanes. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis, vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, p. 3225
311. Bartholomew CH, Farrauto RJ (2006) Dehydrogenation of C_6-C_{15} alkanes to alkenes (section 7.5.1). In: Fundamentals of industrial catalytic processes, 2nd edn. Wiley, Hoboken, pp 542–543
312. Styrene (2010) In: Hydrocarbon processing's petrochemical processes 2010, Gulf Publishing Company, Houston, pp. 264–265
313. James DH, Castor WM (2003) Styrene. In: Ullmann's encyclopedia of chemical technology, vol 34, 6th edn. Wiley-VCH Verlag GmbH, Weinheim, pp 386–390
314. Kochloefl K, Muhler M (2008) Dehydrogenation of ethylbenzene. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis, vol 8, 2nd edn. Wiley-VCH Verlag GmbH, Weinheim, pp 3229–3238
315. Bartholomew CH, Farrauto RJ (2006) Dehydrogenation of ethylbenzene (section 7.5.2). In: Fundamentals of industrial catalytic processes, 2nd edn. Wiley, Hoboken, pp. 543–545
316. Neier W, Strehlke G (2003) 2-Butanone. In: Ullmann's encyclopedia of chemical technology, vol 5, 6th edn. Wiley-VCH Verlag GmbH, Weinheim, pp.725–732
317. Papa AJ, Sherman PD (1981) Ketones. In: Mark HF (ed) Kirk-Othmer encyclopedia of chemical technology, vol 13, vol 13, 3rd edn. Wiley, Hoboken, pp 894–941
318. Sifniades S, Levy AB (2003) Acetone. In: Ullmann's encyclopedia of chemical technology, vol 1, 6th edn. Wiley-VCH Verlag GmbH, Weinheim, pp. 200–201
319. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York, pp 278–279
320. Wells GM (1995) Methyl ethyl ketone, Handbook of petrochemicals and processes. Gower Publishing Company, Brookfield, pp 255–258
321. Burridge E (2008) TDI, Chemical profile, ICIS Chemical Business, 21–27 Jan 2008, p. 38.
322. Simons RM (1983) Esterification. In: McKetta JJ, Cunningham WA (eds) Encyclopedia of chemical processing and design, vol 19. Marcel Dekker, New York, pp 381–402

323. R. L. Rowley, W. V. Wilding, J. L. Oscarson, Y. Yang, R. J. Rowley, T. E. Daubert, R. P. Danner, *DIPPR® Data Compilation of Pure Compound Properties*, dippr@BYU.edu, Design Institute for Physical Property Data, AIChE®, New York, NY, USA, 2001.
324. RL Rowley, W. V. Wilding, J. L. Oscarson, Y. Yang, R. J. Rowley, T. E. Daubert, R. P. Danner, *DIPPR® Data Compilation of Pure Compound Properties*, dippr@BYU.edu, Design Institute for Physical Property Data, AIChE®, New York, NY, USA, 2001.
325. Caspary KJ, Gehrke H, Heinritz-Adrian M, Schwefer M (2008) Dehydrogenation of alkanes. In: Ertl G, Knozinger H, Schuth F, Weitkamp J (eds) Handbook of heterogenous catalysis, vol 8, 2nd edn. Weinheim, Wiley-VCH Verlag GmbH, p 3222

Chemistry in the Pharmaceutical Industry

11

John F. Kadow, Nicholas A. Meanwell, Kyle J. Eastman,
Kap-Sun Yeung, and Joseph Payack

Introduction

This chapter discusses the role of chemistry within the pharmaceutical industry [1–3]. Although the focus is upon the industry within the United States, much of the discussion is equally relevant to pharmaceutical companies based in other first-world nations such as Japan and those in Europe. The primary objective of the pharmaceutical industry is the discovery, development, and marketing of safe and efficacious drugs for the treatment of human disease. However, drug companies do not exist as altruistic, charitable organizations. As with other shareholder-owned corporations within a capitalistic society, drug companies must earn profits in order to remain viable. Profits from the enterprise finance the essential research and development that leads to new drugs designed to address unmet medical needs. Thus, there exists a tension between the dual goals of enhancing the quality and duration of human life and that of increasing stockholder equity. Much has been written and spoken in the lay media about the high prices of prescription drugs and the hardships these place upon the elderly and others of limited income. Consequently, some consumer advocate groups support governmental imposition of price controls on ethical pharmaceuticals in the United States, such as those that exist in a number of other countries. However the out-of-pocket dollars spent by patients on prescription drugs must be weighed against the more costly and inherently risky alternatives of surgery and hospitalization, which can often be obviated by drug therapy. Consideration must also be given to the enormous expense associated with the development of new drugs. It typically takes 10 or more years from the inception of a drug in the laboratory to registrational approval and marketing at an overall cost which is now estimated to be in excess of $800 million and increasing, a figure that includes the opportunity costs of failed development campaigns. Only 1 out of 10,000–20,000 compounds prepared as potential drug candidates ever reach clinical testing in humans and the attrition rate of those that do is >80%, a success rate that has been stubbornly difficult to change despite advances in improving candidate quality and significant increases in investment in research and development. The expense of developing a promising drug grows steadily the further through the pipeline it progresses; clinical trials can be several orders of magnitude more costly than the preclinical development of a compound. While the sales of drugs that complete clinical trials and reach the shelves of pharmacies can eventually recoup their developmental expenses many times over if successful, many fail to do so and the cost of the drugs that fail is never recovered.

To a large extent, the difficulties associated with bringing a drug to market have arisen from the increasingly stringent but appropriate criteria that have been imposed by the Food and Drug Administration (FDA) in the United States and analogous regulatory agencies in other countries. In the current regulatory environment, it is unlikely that an occurrence similar to that associated with thalidomide, a disaster that resulted in terrible birth defects several decades ago, would recur. Moreover, the era of pursuing "me-too" drugs that offer little advantage over a marketed agent is largely obsolete. It is now necessary for a sponsor to demonstrate that a drug for which a New Drug Application (NDA) is submitted to the FDA offers significant benefits in terms of efficacy and/or safety relative to the existing drug therapy. The approvability bar may be lowered for agents designed to treat life-threatening maladies such as cancer and AIDS or for diseases of the nature of Alzheimer's disease for which no effective therapy currently exists. However, even in these cases it is incumbent upon the sponsoring company to provide compelling evidence that their drug is safe and effective. The restrictions imposed by Health Maintenance

---

J.F. Kadow (✉) • N.A. Meanwell • K.J. Eastman • K.-S. Yeung
Bristol-Myers Squibb Co., Research and Development, Virology
Chemistry, Wallingford, CT 06492, USA
e-mail: John.kadow@bms.com

J. Payack
Bristol-Myers Squibb Co., Research and Development, Process R&D
Chemistry Technologies, New Brunswick, NJ 08903, USA

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_11,

Organizations (HMOs) and other payers may also have a significant impact on the sales of any given drug. Most HMOs list only a select few drugs for which they will reimburse costs within any given category, particularly antidepressants and anti-hypertensive or cholesterol-lowering agents.

A major consequence of the financial and logistical impediments to the successful introduction of new drugs has been the high incidence of mergers and acquisitions among US-based pharmaceutical companies in the recent years. These events have not occurred because bigger is necessarily better but because the critical mass of internal resources required to bring a drug from the test tube to the pharmacy continues to grow. In contrast to this trend among the major drug companies (often dubbed "big pharma"), there has been a proliferation of start-up companies often founded by entrepreneurial scientists with "big pharma" or academic experience and financed by venture capital investment. While many such start-ups are tightly focused on the application of biotechnology to develop protein- and nucleic acid-based drugs, others function as mini-drug companies, staffed by both medicinal chemists and biologists. Unlike their much larger brethren, small companies cannot address a wide range of drug research opportunity, leading to a focus upon a particular therapeutic area and, perhaps, even a specific disease. Their mission is to discover drug candidates which a large company may be interested in licensing and developing. The "big pharma" companies do not rely exclusively upon filling their developmental pipelines with drug candidates that have been discovered in-house but often enter into collaborations and licensing agreements to acquire the rights to promising agents from the laboratories of smaller companies or academic researchers in addition to other members of "big pharma."

Medicinal Chemistry

Chemistry has long been an integral part of the pharmaceutical industry and remains an important capability, relying on both purely synthetic molecules and natural products as a source of lead compounds. Many currently marketed drugs, such as the antineoplastic agent paclitaxel and the antibiotic vancomycin, are natural products. The extracts of plants and marine organisms and the products of soil bacteria fermentation continue to be investigated as potential sources of powerful new drug substances. In this arena of drug discovery, chemists conduct the painstaking isolation, purification, and structural characterization of pharmacologically active components which most often are present in minute amounts in the natural source and which typically have extremely complex chemical structures. More recently, the enormous advances in molecular biology have resulted in the successful development of bio-engineered injectable, protein-based therapeutic agents that are encompassed under the general term biologics and which are described in greater detail later in this chapter. It is anticipated that by 2014, biologic drugs will represent the majority of the top ten selling prescription medicines; however, the impact of this development on the role of medicinal chemists is unclear.

The majority of existing drugs are small organic molecules of molecular weight approximately 200–600 that have been designed and synthesized by medicinal chemists. It is thus important to define what is meant by "medicinal chemist" and what role is played by the practitioners of this sub-discipline in the pharmaceutical industry [4]. A traditional and, perhaps, somewhat narrow definition of a medicinal chemist is that of a researcher engaged in the design and synthesis of bioactive molecules. As part of their academic training, many medicinal chemists performed doctoral and postdoctoral work that involved the total synthesis of natural products and/or the development of synthetic methodology. They are hired by pharmaceutical companies based on their ability and skill in planning and conducting the synthesis of organic compounds. While such skills can remain important throughout the chemists' careers, they alone are insufficient for the challenging task of drug discovery. In contrast to an academic environment where the synthesis of target molecules is frequently the objective, the enterprising young chemical researcher who enters the pharmaceutical industry must be able and willing to undergo an evolution from that of a pure synthetic chemist to that of the medicinal chemist who also has insights into drug design principles and an excellent understanding of the biochemistry underlying therapeutic targets.

Several aspects of drug design and function are important components of the medicinal chemists' knowledge. The essential structural elements that comprise a drug molecule and enable it to interact on the molecular level with a biological macromolecule such as a receptor or enzyme and thus impart a pharmacological effect are defined as pharmacophores, which can be of varying complexity. The medicinal chemist must become skilled at analyzing structure activity relationships (SAR) that pertain to a series of compounds under study. These relate the activity of analogs within an active series tested in a biological assay with changes in structure that depend on the introduction of substituents of various size, polarity, and lipophilicity at various domains of the parent drug molecule. Elucidation of the SAR within a series of active compounds is the key to optimizing the potency and other desirable biological properties in order to identify a new chemical entity (NCE) as a *bona fide* drug candidate. Quantitative structure activity relationships (QSAR) are often employed in this effort; analyses that employ linear free-energy relationships, linear regression, and other techniques that can be utilized to correlate biological activity with the electronic, steric, polarizability, and other physical/chemical parameters of the substituent groups of structurally related compounds. There is also a rising appreciation of the benefits of optimizing compounds based on ligand efficiency, which relates potency to the number of heavy atoms in a molecule, and lipophilic ligand efficiency which assesses the relationship between biological activity and compound lipophilicity. The objective of these analyses is to minimize molecular size and hydrophobicity while maximizing potency and selectivity.

The synthesis and isolation of pure enantiomers has become increasingly important. In the past, chiral drugs were frequently marketed as racemic mixtures since it was not deemed cost-effective to provide them in enantiomerically pure form. However, in many cases one or the other enantiomers of an optically active drug may have a significantly greater level of the desired biological activity and/or lower side effect liability than its antipode. Regulatory agencies such as the FDA now routinely require that each enantiomer of a chiral drug be isolated and evaluated in tests of efficacy, side effects, and toxicity. If one of the enantiomers is shown to be clearly superior, then it is likely that it is the form that will be developed as the drug candidate. Thus, enantioselective chemical reactions which can afford a high enantiomeric excess (ee) of one or the other of a pair of enantiomers are valuable components of the medicinal chemist's synthetic toolkit. Enzymes also play a prominent role in drug development since isolated enzymes or microorganisms can often achieve an enantiospecific chemical transformation much more efficiently and economically than conventional synthetic methods. Many "big pharma" companies now have dedicated groups that exclusively study and develop enzymatic reactions.

Research Strategies

The discovery of a new drug may occur in a serendipitous fashion or as a result of a unique insight. However, pharmaceutical companies cannot depend on chance occurrences as a research strategy. The aforementioned "me-too" approach has not been completely abandoned and it is likely that the clinical evaluation of a novel drug will be followed by a number of competitors' agents but with the caveat that the latter offer some therapeutic advantage over the prototype.

The most scientifically sound approach is that of rational drug design, which is based on an understanding of the biochemical mechanisms underlying a particular disease. If, for example, over activity or under activity of a certain neurotransmitter system is believed to be responsible for a central nervous system (CNS) disorder such as depression, then medicinal chemists can endeavor to design agents capable of normalizing neurotransmission by their action upon the receptor proteins through which interneuronal communication is mediated. The cloning and expression of human genes to afford functional receptors and enzymes that can be studied in cell culture has been a tremendous advance in the ability to evaluate drug action at the molecular level. Similarly, advances in molecular biology have afforded access to quantities of pure macromolecules that are essential to the life cycle of pathogens such as bacteria and viruses, thus enabling novel mechanistic strategies for the treatment of infectious disease. In many cases, X-ray crystallography has provided a detailed 3D structure of a macromolecule such as an enzyme with or without a bound substrate. Researchers with expertise in computer-assisted drug design (CADD) can depict the determined structure on graphics terminals and, in collaboration with medicinal chemists, propose drug molecules designed to effectively complement an active site. This kind of detailed analysis of protein structure was instrumental in the design of a number of drugs that inhibit human immunodeficiency virus (HIV) protease, an enzyme essential to the replication of the virus that causes AIDS.

There have been approximately 1,000 human proteins identified as potential targets for drug intervention in various diseases and it has been estimated that the determination of the human genome would increase that number by at least tenfold. Therefore, it seems safe to predict that the rational approach to drug discovery will grow accordingly and with it the role of synthetic/medicinal chemistry. There will be intense competition within the pharmaceutical industry to determine the functional relevance of this multitude of new targets in the absence and presence of disease, a critical validation exercise imperative if compound attrition in the clinic is to be reduced. A close nexus to this quest will be the search for compounds that can impart selective pharmacological effects upon target proteins. It is unlikely that these goals will be met by employing only the classical iterative approach which entails the synthesis of one compound at a time and low volume testing. Instead, the challenges of this new era of research must be met by methodologies that can synthesize and test large numbers of compounds in a short period of time—that is, combinatorial chemistry and high-throughput screening (HTS). In the context of its application within pharmaceutical research, combinatorial chemistry should not be regarded as a separate discipline but as a technologically specialized part of medicinal chemistry. This topic will be discussed in detail in a later section of the chapter.

Another important interface occurs between discovery medicinal chemists and chemists in process research and development. Medicinal chemists are not overly concerned with the cost, toxicity, or environmental impact of the starting materials, reagents, and solvents they employ to synthesize target compounds since they are dealing with relatively small quantities of compound. Similarly, reaction conditions that employ very low or elevated temperatures and pressures are not problematic on the discovery scale and synthetic sequences are typically designed to maximize access to diverse target structures. However, these and other pragmatic considerations must be taken into account for the bulk-scale preparation of experimental drugs. Process chemists invariably modify the synthetic procedures implemented by medicinal chemists and, in many cases, devise an entirely new synthetic pathway. Process chemistry will also be discussed in an ensuing section.

Pharmacodynamics

Medicinal chemists must be knowledgeable about pharmacodynamics, that is, the effect of drugs upon biological systems. In addition to being aware of the state-of-the-art understanding of the biological mechanisms that underlie the particular diseases for which they are endeavoring to discover drug therapy, they need to understand the basis of the various in vitro and in vivo tests that biologists employ to evaluate both the potential efficacy and side-effect liability of synthesized compounds. Because drug research covers a plethora of human diseases, each with its own unique combination of etiology and biochemical mechanisms, the number and diversity of biological tests are far too great to discuss in this chapter. Suffice it to say that in a general sense, the primary and often even the secondary biological tests of drugs for a particular disease target are in vitro tests that can be run quickly, inexpensively, and using small amounts of compound. For example, these can be receptor-binding assays for CNS drugs, enzyme assays for antihypertensive agents, inhibition of bacterial colony growth by antibiotics, and the killing of cultured cancer cells by oncolytic drugs. Encouraging results from in vitro evaluation lead to in vivo testing in an appropriate animal model. In vivo tests are more laborious and costly, but are necessary to establish that a drug is effective in an intact living organism; they can range from complex behavioral paradigms for CNS drugs to enhancement of survival time of tumor-implanted mice by experimental cancer drugs. An evaluation of the propensity of a drug candidate to cause side effects is as important as efficacy testing. Even if a compound shows an encouraging level of the desired activity, a lack of selectivity can cause it to induce a number of undesirable pharmacological effects thus precluding its further development. The medicinal chemist must be able to interpret the results of tests run on lead compounds and use this information as a guide to further synthetic work.

Table 11.1 Lipinski's rules for drug absorption

Absorption of a drug following oral administration is favored by:
Molecular weight of <500
The drug molecule has <5 hydrogen bond donors
The drug molecule has <10 hydrogen bond acceptors
The octanol/water distribution coefficient, log P, is <5

Source: Lipinski et al. [5]

Pharmacokinetics and Toxicity

It is also necessary that chemists understand the principles of pharmacokinetics (PK), that is, the effects of biological systems upon drugs. The key PK parameters, absorption, distribution, metabolism, and excretion (ADME), are as critical as biological activity in determining whether an NCE is a viable drug candidate. A compound may exhibit high affinity for a biological receptor or potent inhibition of an enzyme in an in vitro assay but if it is poorly absorbed or rapidly metabolized to an inactive species then it will be ineffective as a drug. For example, the empirically based Lipinski's rule of five (Table 11.1) broadly defines the optimal physical/chemical properties, including molecular weight, lipophilicity, and hydrogen-bond forming moieties, that promote oral absorption of drug candidates [5]. A compound with potent intrinsic activity can be rendered ineffective in vivo by its rapid conversion to inactive metabolites. The susceptibility of compounds to metabolic conversion can be assessed by incubating them with liver microsomal homogenates from various species including rodent, dog, monkey, and human or with cloned, expressed human hepatic metabolizing enzymes. Analysis of the incubated drugs by liquid chromatography/mass spectometry (LC/MS) can quantify the extent of metabolism and even identify some specific metabolites. In vivo administration of a drug candidate to one or several animal species is required to determine its oral bioavailability, half-life, and other PK properties such as distribution and elimination. If an unsatisfactory PK profile threatens the demise of an otherwise promising drug candidate, it falls upon the medicinal chemist to make structural permutations designed to correct the problem. If poor absorption is the problem, this may entail modifying the lipophilicity of the drug molecule to render it more membrane permeable. A metabolic liability might be rectified by blocking the site of biotransformation with a metabolically inert atom or group.

Toxic effects of a compound on blood or organs or the potential to cause gene aberrations will eliminate a compound from further consideration even if biological activity and PK properties are excellent. Promising lead compounds are screened using in vitro tests in bacteria and mammalian cells to determine whether they cause damage to DNA or induce genetic mutations. If they pass this hurdle, compounds are dosed on a daily basis for periods ranging from several weeks to several months to both a rodent and non-rodent (usually dog or monkey) species with the animals observed closely for any adverse effects. At the completion of the study period, test animals are necropsied to ascertain whether any organ or tissue damage occurred at the microscopic level. Unacceptable toxicological findings will invariably lead to the termination of a drug candidate, with responsibility once again falling to the medicinal chemist to devise and implement structural modifications designed to eliminate the toxicity. This may be a more daunting task than overcoming a side effect or metabolic issue, since the underlying cause of toxicity is frequently not defined or the toxicity may be related to the mechanism of action of the compound.

Drug Delivery

Drugs can be administered to patients in many ways. The most common and preferred route is oral administration and oral drugs are generally formulated as tablets or capsules in which a specific dose of the drug substance is homogeneously mixed with an inert filler or excipient designed to modulate drug dissolution. Some oral medications, such as antibiotics for use in the pediatric population, are formulated as solutions, as are injectable drugs. Obviously this requires satisfactory solubilization of the drug, preferably in aqueous medium, and compounds bearing an ionizable group such as a basic amine or an acidic functionality can usually be converted into water-soluble salts. However, neutral molecules present greater difficulties. In some cases, the results of clinical trials have indicated that an experimental injectable drug showed promising efficacy but failed to elicit a robust response because poor intrinsic solubility limited the amount that can be administered, resulting in low plasma levels of the drug. Inadequate membrane permeability can restrict the absorption and, hence, the bioavailability of an orally administered drug. Medicinal chemists can respond to such findings by investigating the feasibility of preparing a suitable prodrug [6]. A prodrug is a derivative in which a cleavable solubilizing or permeabilizing group is covalently appended to the parent drug molecule, most often via a hetero atom such as oxygen or nitrogen. An effective prodrug is the one which has much higher solubility or permeability than the parent drug but which is rapidly cleaved in vivo following administration to achieve a therapeutically beneficial plasma concentration of the biologically active parent drug.

Patents

The granting of patent protection on both approved and experimental drugs is of critical importance to the pharmaceutical industry. Issued patents provide a company with exclusivity for the manufacture, use, and sale of drug products and it is highly unlikely that a company would undertake the risks and costs of developing an agent for which it had no patent protection. There are several types of patents of which the “composition of matter” (COM) or “product” patent may be deemed to have the greatest value. An approved COM patent covers specifically claimed compounds of a defined structural chemotype and provides tangible evidence that the claimed compounds have been prepared, characterized, and found to have utility, the latter a critical aspect of an invention. In order to be patentable, the compounds must have structural novelty and cannot have been publicly disclosed either in the scientific or patent literature or by oral presentation. However, structural novelty alone is an insufficient basis for a patent to be granted since it must be demonstrated that the compounds have utility in the treatment of a disease. The basis of such utility is activity in appropriate and relevant biological tests. Clinical data may also be used in support of a patent application although in the great majority of cases the applications are filed well before any compound within the application reaches clinical trials. Medicinal chemists are closely affiliated with the patent process and are most commonly named as inventors listed on COM patents covering drug substances. The chemists and other researchers with whom they collaborate must provide the chemical and biological data that is included in a patent application and the chemists will also provide input that defines the scope and claims of the invention. Since patents are legal documents that provide the assignee exclusive proprietary rights to the claimed subject matter for 20 years from the date issuance, it is essential that all supportive data be accurate and instructive. If a patent is challenged by another party and is found to contain erroneous information, it may be invalidated. Moreover, in the United States, patents are granted on a “first-to-invent” basis. Thus, if two or more parties submit applications to the US Patent Office claiming identical subject matter, then the patent will be awarded to the party that can prove that it had the earliest conception and reduction to practice of the subject matter. Therefore, it is imperative that chemists maintain accurate records of all experimental work in approved notebooks and that such records are dated, signed, and witnessed in a timely fashion. In March of 2013, the United States will change from "First to Invent" to "First to File" in order to harmonize with the rest of the world.

Other types of drug-related patents include process, use, and formulation patents. Chemists are responsible for process patents, which describe an improved method of preparation of some drug substance, but are typically minimally involved with the others. Use patents are based on the discovery of some non-obvious utility of a compound that is either part of the public domain or covered by an existing patent; such discoveries are most likely to be made by biologists. Formulation patents disclose a preferred means of drug delivery of a known drug substance.

Clinical Trials

Even though there is no involvement of chemistry in the clinical evaluation of drugs, any discussion of the pharmaceutical industry must include clinical trials since the results determine whether an experimental drug has the combination of efficacy, safety, and tolerability that will allow it to achieve registrational approval and reach the market. If a drug candidate survives the hurdles of preclinical pharmacological, pharmacokinetic, and toxicological testing, the next customary step in the United States is the sponsoring firm's filing of an Investigational New Drug (IND) application with the FDA. This is a formal request to initiate clinical investigation in humans that describes the preclinical profile of a molecule and is accompanied by a detailed description of the planned studies and clinical protocols. Upon approval of the IND, Phase I clinical studies are initiated.

Phase I studies are conducted in healthy volunteers in order to establish the drug's safety and to determine appropriate dosage levels. If the drug is found to have an acceptable human pharmacokinetic profile and to be free of untoward side effect liabilities, it is advanced into Phase II trials, which are typically carried out in as many as several hundred patients and may last from 6 months to 2 years. Phase II trials are designed to ascertain the appropriate dosing regimen for a drug and whether it is effective in treating the targeted disease. Only about one-third of drugs pass Phase II trials, most failing because of a lack of efficacy or unacceptable side effects. Those that pass are advanced into Phase III trials which may involve from several hundred to several thousand patients and which can last from 1 to more than 3 years, depending on the type of drug and disease under study and the complexities of the study design. Phase III trials provide the ultimate test of an experimental drug since they are designed to verify the drug's effectiveness against the target disease and define its long-term safety profile. For agents that are intended for chronic use, studies also monitor adverse reactions that may develop only after long-term use and the development of tolerance. Clinical studies of many drug classes will commonly employ several patient groups of approximately equal size with one group receiving the experimental drug, another placebo (nondrug), and/or another a positive control; that is, a marketed drug used to treat the same disease for which the experimental agent is being evaluated. In order to minimize the possibility of bias in favor of the test drug, such studies are most often run in a double-blinded manner with neither patients nor clinical investigators knowing which group is receiving treatment until the conclusion of the trial.

If a drug candidate is among the 20–25% that successfully negotiates Phases I–III and if statistical analysis of the clinical data supports its efficacy, then the sponsor will assemble the voluminous data generated into a NDA which is submitted to the FDA. The standard period for review of the NDA is 10 months but a priority review can be requested based on unmet medical need which reduces the time to 6 months. Actual approval can take longer, as long as 1–2 years if the FDA requests that additional information be provided or even that some additional studies be conducted. When approval is granted, the company is free to market the drug. The most typical current practice is for NDAs to be filed globally with regulatory agencies around the world essentially simultaneously with the submission in the United States. This is a resource-intensive practice which reflects the importance of non-US markets. The results of clinical evaluation of an experimental drug can feed back into medicinal chemistry. For example, if a drug is found to fail because of poor bioavailability in humans, medicinal chemists will endeavor to design and prepare an analog with improved pharmacokinetic properties.

Summary

The preceding sections, by necessity, present a succinct and somewhat superficial overview of the very extensive and complex role of a medicinal chemist in the pharmaceutical industry and the interface with other disciplines. An acquired understanding of relevant biology, pharmacology, toxicity, and its application to drug design is not simply of heuristic value but is essential for the chemist to engage in meaningful dialogue with colleagues who work in these specialties. Successful drug discovery and development cannot be performed by individuals working in isolation but requires the interactive collaboration of a multitude of researchers representing multiple scientific disciplines, as depicted in Fig. 11.1. It can be argued that medicinal chemists are the most versatile generalists among these researchers since they must command a primary expertise in chemistry along with an extensive knowledge of the other disciplines that contribute to drug discovery and development. The following section presents examples of marketed drugs in a number of different therapeutic categories.

Cardiovascular Agents

Hypertension

A variety of agents of several mechanistic types are currently available for the treatment of hypertension (elevated blood pressure). The dihydropyridine derivative amlodipine (Norvasc®) is a receptor-operated, calcium channel blocker that prevents Ca^{2+} entry into vascular smooth muscle cells that is also useful for the treatment of

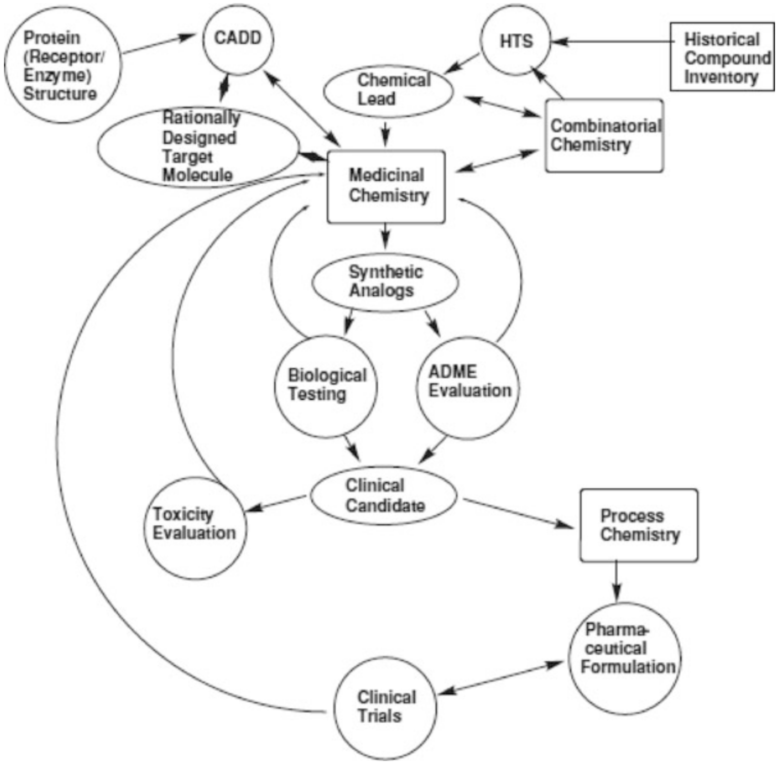
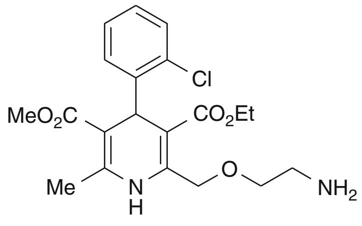
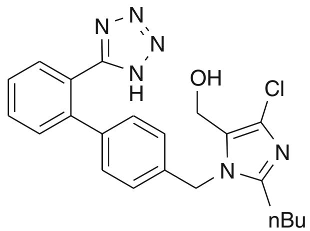


Fig. 11.1 Drug discovery and development is a complicated process that involves the interaction of researchers in a range of disciplines. Medicinal chemists may synthesize analogs based on chemical leads arising from the high-throughput screening of combinatorial libraries or historical compound inventories. Alternatively, analog synthesis can be based on a collaboration between medicinal chemistry and computer-assisted drug design scientists to rationally design small molecules capable of interacting with a macromolecular biological target (receptor or enzyme). Subsequent biological, pharmacokinetic, and toxicological evaluations lead to the identification of a drug candidate that, following development of a suitable bulk scale synthesis by process chemistry and pharmaceutical formulation, is advanced into clinical trials. Feedback to medicinal chemistry from any of these developmental steps can give rise to further optimization of structure by synthetic modifications and refinement

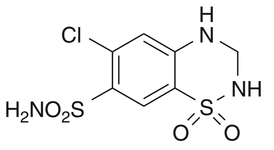
angina. Losartan (Cozaar®) is an angiotensin receptor antagonist that inhibits the action of the peptide angiotensin II on the AT1 receptor and is often prescribed in combination with the diuretic hydrochlorothiazide (Hyzaar®). Metoprolol (Toprol®) is a cardioselective, β1-adrenergic receptor blocking agent that is also useful in the treatment of angina.



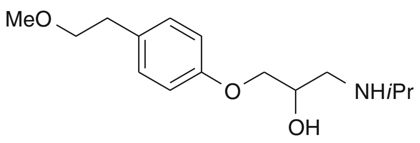
amlodipine



losartan



hydrochlorothiazide

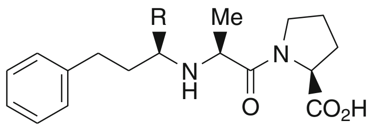


metoprolol

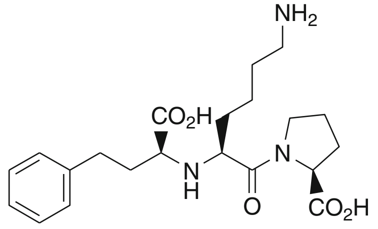
Congestive Heart Failure, Migraine, and Antithrombotic Agents

Enalapril (Vasotek®) and lisinopril (Zestril® and Prinvil®) are angiotensin-converting enzyme (ACE) inhibitors, useful in the treatment of congestive heart failure and hypertension by suppression of the renin–angiotensin–aldosterone system. Enalapril is an ethyl ester prodrug that is hydrolyzed in the liver to the active carboxylic acid enalaprilat. Aliskiren (Tekturna®) is the first direct-acting renin inhibitor to be marketed for the treatment of hypertension.

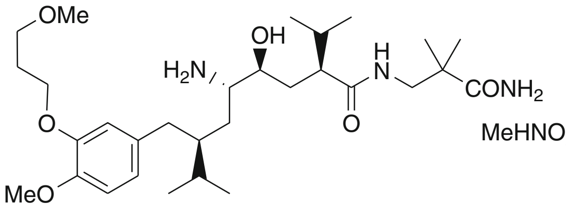
Sumatriptan is a selective agonist at serotonin (5-hydroxytryptamine) type-1 receptors (most likely the 5-HT1$_B$ and 5-HT1$_D$ subtypes) in the vasculature. It is thought to exert its beneficial effects on migraine headaches by selectively constricting certain large cranial blood vessels and/or possibly through suppression of neurogenic inflammatory processes in the CNS. Clopidogrel (Plavix®) is the prodrug of an inhibitor of ADP-induced platelet aggregation that is useful in the prevention of thrombotic events such as stroke and myocardial infarction in patients with symptomatic atherosclerosis and acute coronary syndrome without ST-segment elevation. Rivaroxaban (Xarelto®) is the first direct-acting inhibitor of the coagulation enzyme factor Xa to be approved in Europe for the prevention of venous thromboembolism in adults undergoing total hip replacement or total knee replacement surgery. Dabigatran etexilate (Pradaxa®) is a double prodrug of the first direct-acting thrombin inhibitor to be approved for the prevention of stroke and systemic embolism in patients with atrial fibrillation. Rivaroxaban and dabigatran are the first members of a new class of anti-coagulants designed to replace warfarin.



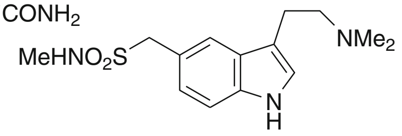
enalapril: R = CO_2Et
enalaprilat: R = CO_2H



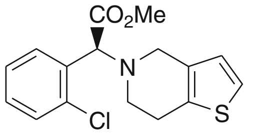
lisinopril



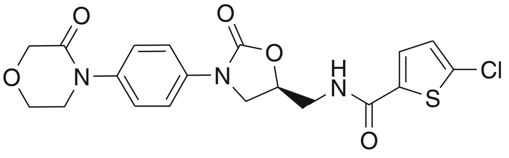
aliskiren



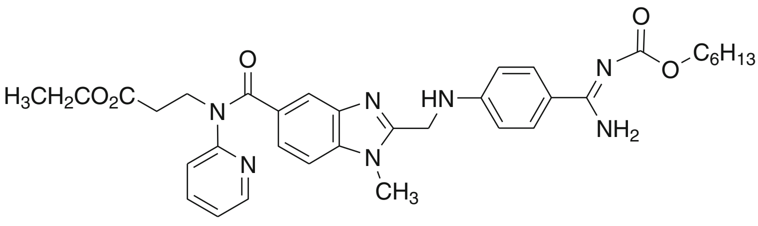
sumatriptan



clopidogrel



rivaroxaban

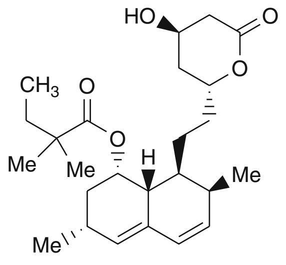


dabigatran

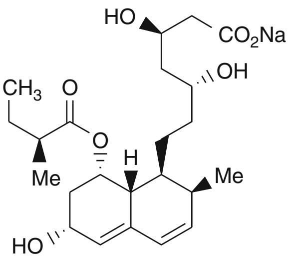
Metabolic Agents

Hyperlipidemia

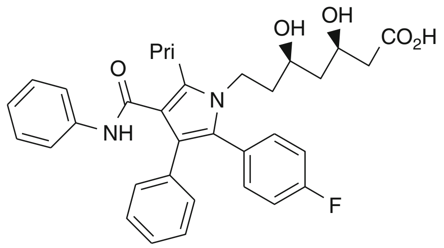
Simvastatin (Zocor®), pravastatin sodium (Pravachol®), atorvastatin (Lipitor®/Pfizer), and rosuvastatin (Crestor®) are hydroxymethylglutaryl-CoA (HMG-CoA) reductase inhibitors (statins) that lower serum lipid levels by inhibiting cholesterol biosynthesis. Simvastatin and pravastatin sodium are semi-synthetic, mevalonic acid-derived antilhyperipidemic agents whereas atorvastatin is a wholly synthetic, pentasubstituted pyrrolo heptanoic acid. Unlike pravastatin, simvastatin is a lactone prodrug which must be converted to the corresponding, ring-opened δ-hydroxy acid in vivo. Ezetimibe (Zetia®) is a newer antihyperipidemic agent with a novel mechanism of action. Ezetimibe does not inhibit cholesterol biosynthesis in the liver as do the statins but rather inhibits cholesterol absorption in the intestine. This novel action is complementary to the HMG-CoA reductase mechanism displayed by the statins. An innovative new product for the treatment of hyperlipidemia is Vytorin® which consists of a mixture of simvastatin and ezetimibe in one pill.



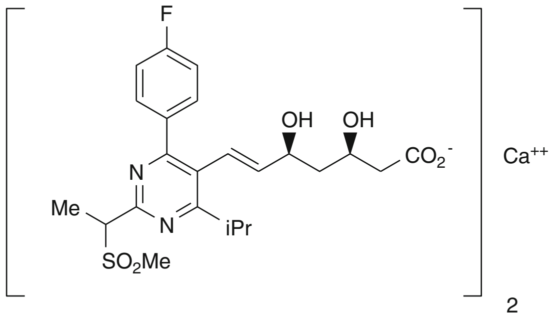
simvastatin



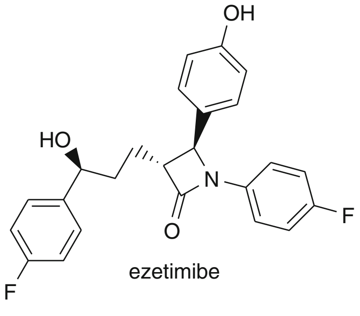
pravastatin sodium



atorvastatin



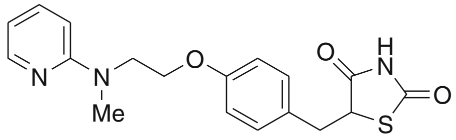
rosuvastatin



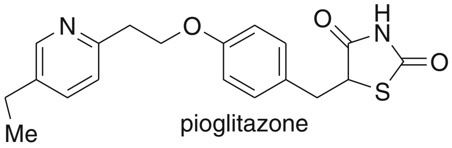
ezetimibe

Diabetes

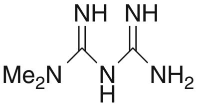
A variety of mechanistic agents are currently available for the treatment of type 2 (non-insulin dependent) diabetes mellitus (NIDDM). Rosiglitazone (Avandia®) and pioglitazone (Actos®) are thiazolidinedione antidiabetic agents and agonists at the peroxisome proliferator-activated receptor-gamma (PPARgamma). Activation of this receptor enhances insulin sensitivity in target tissues by increasing insulin-responsive gene transcription. Concern about adverse effects has reduced the use of Avandia® and made its manufacturer the target of many lawsuits despite its sustained effects on glycemic control. Some reviewers of available data have concluded that Avandia® caused more deaths than Actos® but an FDA panel disagreed and it remains on the market in the US, subject to significant restrictions. Metformin (Glucophage®) is an antihyperglycemic agent that improves glucose tolerance in patients with type 2 diabetes. The compound acts by decreasing both hepatic glucose production and intestinal absorption of glucose, and improves insulin sensitivity by increasing peripheral glucose uptake and utilization. Glimepiride (Amaryl®) is a member of the sulfonylurea class of antidiabetic agents that is thought to lower blood glucose concentration by stimulating insulin secretion in pancreatic beta cells. Saxagliptin (Onglyza®), sitagliptin (Januvia®), and vildagliptin (Galvus®) are newer drugs for the treatment of type 2 diabetes mellitus. These agents inhibit the proteolytic activity of the serine protease dipeptidylpeptidase-4, thereby potentiating the glucose regulatory action of the endogenous incretin peptides. Therapy that combines these DPP4 inhibitors with metformin has demonstrated beneficial effects in improving glycemic control compared to monotherapy with either type of agent alone.



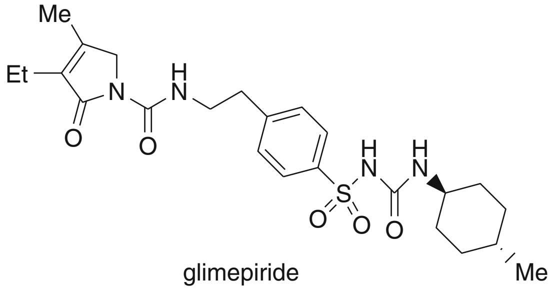
rosiglitazone



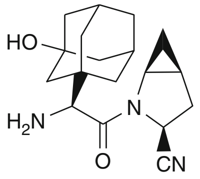
pioglitazone



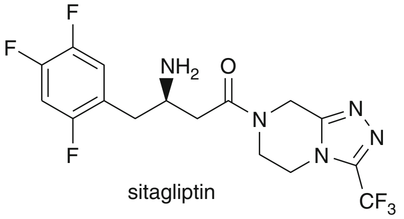
metformin



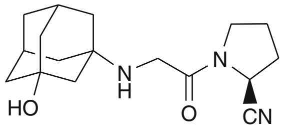
glimepiride



saxagliptin



sitagliptin



vildagliptin

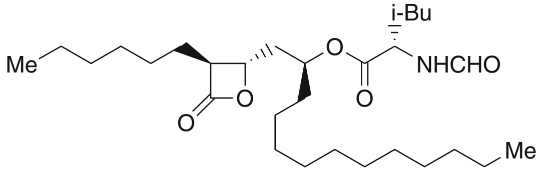
Obesity

Orlistat (Xenical®) is a reversible gastric and pancreatic lipase inhibitor. The compound has no effect on appetite suppression but rather acts by inhibiting dietary fat absorption from the GI tract. Obesity is an area that remains one of compelling and urgent significant medical need but it has been very difficult to identify effective and safe anti-obesity agents with alternative mechanisms of action. Sibutramine (Meridia®) and its major active metabolites are re-uptake inhibitors of norepinephrine, serotonin, and dopamine and exert their beneficial effect through appetite suppression but marketing of the compound in the United States was discontinued in October of 2010 due to elevated risk of cardiovascular side effects identified in post-marketing studies. The cannabinoid-1 receptor blocker, rimonabant (Acomplia®), decreases the activity of the endocannabinoid system, which is essential in regulating body weight and energy balance in addition to glucose and lipid metabolism. However, rimonabant has recently been suspended from the European market due to concerns over psychiatric adverse events and has not been approved for use in the United States.

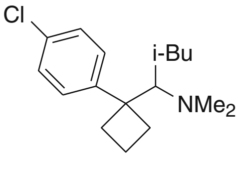
Gastrointestinal and Genitourinary Agents

Antisecretory

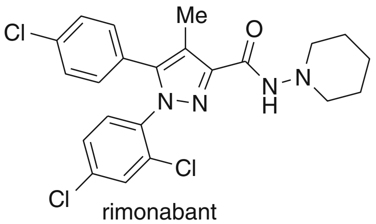
Ranitidine (Zantac®) is a histamine H_2-receptor antagonist that inhibits the release of gastric acid and is useful in the treatment of a variety of hypersecretory conditions (dyspepsia, heartburn, duodenal and gastric ulcers, and gastroesophageal reflux (GERD)). Lansoprazole (Prevacid®), omeprazole (Prilosec®), and esomeprazole (Nexium®) are benzimidazole gastric antisecretory agents and are unrelated both chemically and pharmacologically to the H_2-receptor antagonists. These agents are known as proton pump inhibitors due to their ability to inhibit the H^+K^+-ATPase (the proton pump) in gastric parietal cells thereby blocking the secretion of hydrochloric acid. Esomeprazole is the (*S*)-enantiomer of omeprazole which is racemic and is thus a mixture of both its (*R*)- and (*S*)-enantiomers. Lansoprazole and omeprazole are also useful in the management of duodenal and gastric ulcers, and GERD.



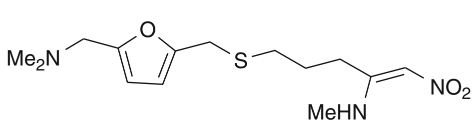
orlistat



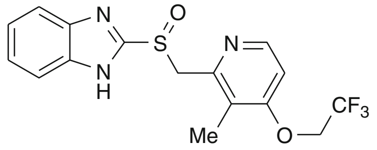
sibutramine



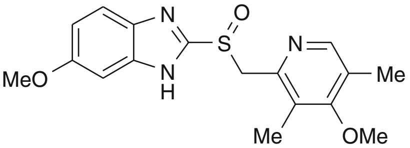
rimonabant



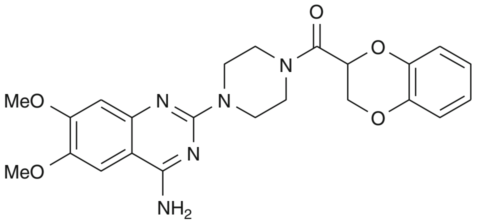
ranitidine



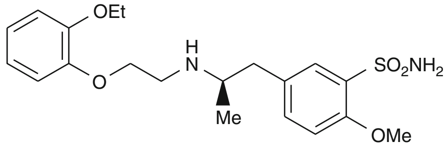
lansoprazole



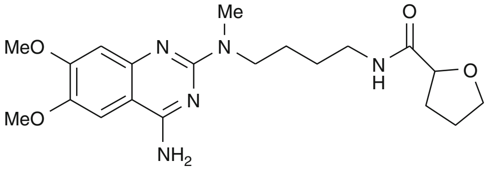
omeprazole



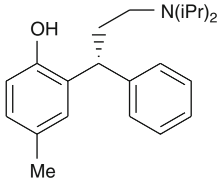
doxazosin



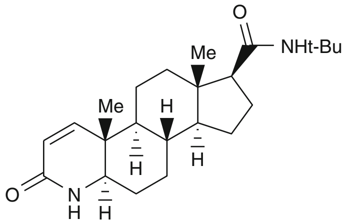
tamsulosin



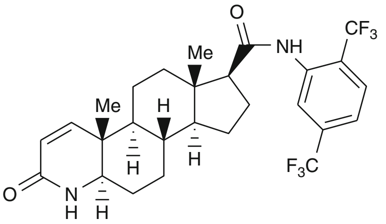
alfuzosin



tolterodine



finasteride



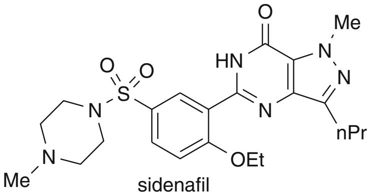
dutasteride

Benign Prostatic Hyperplasia and Urinary Urge Incontinence

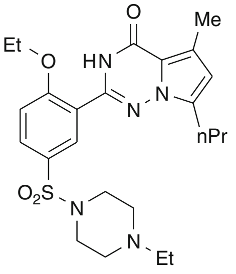
Doxazosin (Cardura®), tamsulosin (Flomax®), and alfuzosin (Uroxatral®) are used in the management of benign prostatic hyperplasia (BPH). These compounds are postsynaptic, α_1-adrenergic blocking agents that relax prostatic tissue and increase urinary outflow in men. Because tamsulosin demonstrates selectivity for the α_{1a}-adrenergic receptor subtype located in prostate over that of α_{1b}-subtype located in vascular tissue, there is a reduced incidence of cardiovascular side effects (hypotension, dizziness, and syncope). Because doxazosin is not selective for the α_{1a}-subtype, it is also useful in the treatment of hypertension. Finasteride (Proscar®) and dutasteride (Avodart®) are 5α-reductase inhibitors that block the conversion of testosterone to 5α-dihydrotestosterone (DHT). Because DHT is an androgen responsible for prostatic growth, inhibition of the 5α-reductase enzyme is beneficial in reducing prostatic enlargement.

Erectile Dysfunction

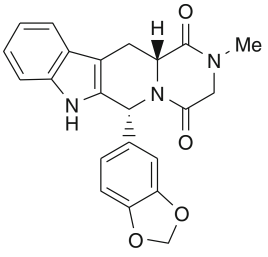
The pyrazolopyrimidinone derivative sildenafil (Viagra®), the indolopyrazinone derivative tadalafil (Cialis®), and the imidazotrizinone derivative vardenafil (Levitra®) are selective inhibitors of the phosphodiesterase (PDE) type 5 enzyme marketed for the treatment of erectile dysfunction. Selective blockade of the PDE type 5 isoenzyme enhances cGMP levels, causing vasodilation in corpus cavernosal tissue which increases penile blood flow and, in turn, penile tumescence and rigidity.



sidenafil



vardenafil



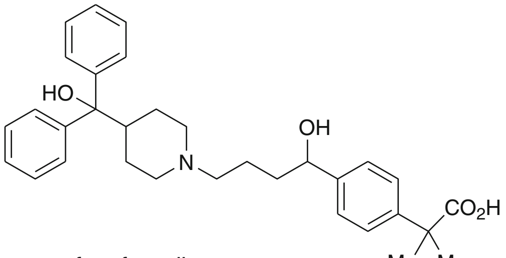
tadalafil

Pulmonary Agents

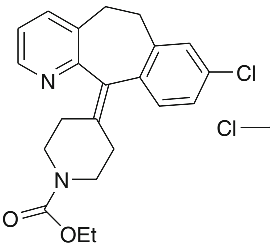
Asthma and Allergic Rhinitis

The butyrophenone derivative fexofenadine (Allegra®) is an antihistamine and used in the treatment of seasonal allergic rhinitis. Because fexofenadine does not readily cross the blood–brain barrier, it is considered to be a "non-sedating" antihistamine. Loratadine (commonly known as Claritin®) and cetirizine (Zyrtec®) are other non-sedating antihistamines that have been converted from prescription-only to over-the-counter status.

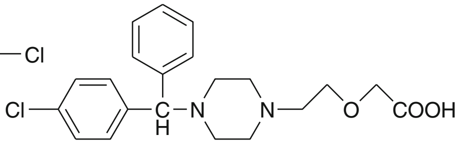
Fluticasone propionate (Flovent®) is a synthetic corticosteroid derivative that is a selective agonist at the human glucocorticoid receptor and useful in the treatment of asthma. Although the precise mechanism of fluticasone action in asthma is unknown, it is believed that its anti-inflammatory property contributes to the observed beneficial effects. Combination of an inhaled corticosteroid with a long-acting β_2-adrenergic agonist is a major therapeutic regimen for treating asthma and chronic obstructive pulmonary disorder (COPD). Fluticasone/salmeterol (Advair®) and budesonide/formoterol (Symbicort®), taken by oral inhalation, are examples of this class. More recently, the quaternary ammonium tricyle tiotropium bromide (Spiriva®) has been introduced. It is a long-acting bronchodilator useful in the treatment of asthma and COPD, and exerts its pharmacological effect through inhibition of the muscarinic M3 receptor. Montelukast (Singulair®) is a leukotriene receptor antagonist indicated for the treatment of asthma, as well as seasonal and perennial allergic rhinitis, that acts through the inhibition of the leukotriene-mediated inflammatory processes.



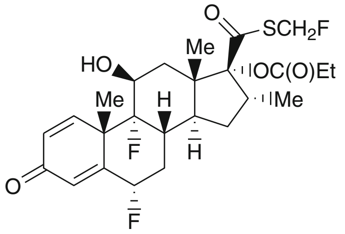
fexofenadine



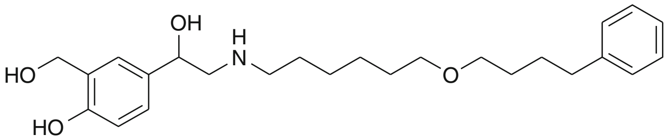
loratidine



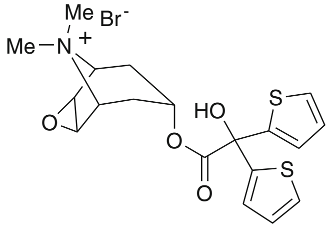
cetirizine



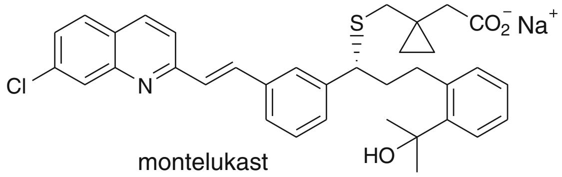
fluticasone propionate



salmeterol (β_2-adrenergic agonist)



tiotropium bromide

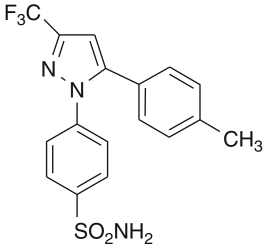


montelukast

Inflammation and Osteoporosis

Arthritis

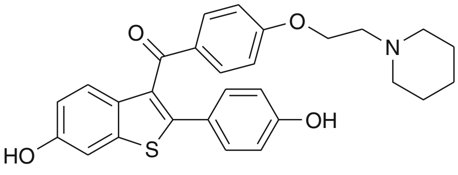
The diaryl pyrazole derivative celecoxib (Celebrex®) is a selective cyclooxygenase type-2 (COX-2) inhibitor useful in the treatment of arthritis. The compound exerts its pharmacological effect by selectively blocking the COX-2 enzyme to produce an anti-inflammatory effect without the adverse gastrointestinal side effects that accompany COX-1 inhibition. In addition, celecoxib also displays analgesic and antipyretic activities in animal models.



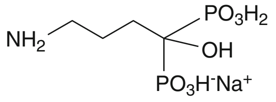
celecoxib

Osteoporosis

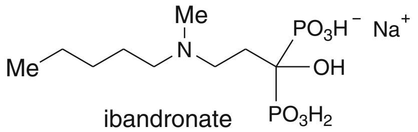
The benzothiophene derivative raloxifene (Evista®) is a selective estrogen receptor modulator (SERM). Raloxifene produces its biological actions via modulation (both activation and blockade) of estrogen receptors that ultimately results in decreased resorption of bone. The bisphosphonate derivative alendronate (Fosamax®), an inhibitor of osteoclast-mediated bone resorption, is also useful in the treatment of osteoporosis. Ibandronate sodium (Boniva®) that can be taken once monthly is another bisphosphonate derivative which, like alendronate and raloxifene, is useful in the treatment of osteoporosis in postmenopausal women.



raloxifene



alendronate

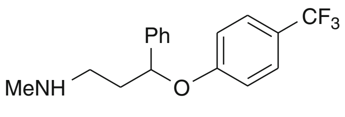


ibandronate

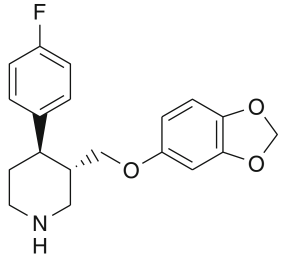
Central Nervous System Agents

Antidepressants

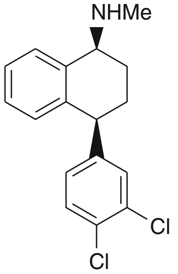
Fluoxetine (Prozac®), paroxetine (Paxil®), and sertraline (Zoloft®) are selective serotonin reuptake inhibitors (SSRIs) that are useful for the treatment of depression. These agents potentiate the pharmacological actions of the neurotransmitter serotonin by preventing its reuptake at presynaptic neuronal membranes. In addition to its SSRI properties, venlafaxine (Effexor®) also appears to be a potent inhibitor of neuronal norepinephrine reuptake and a weak inhibitor of dopamine reuptake, thereby concomitantly enhancing the actions of these neurotransmitters. Venlafaxine is indicated for use in the treatment of anxiety and depression.



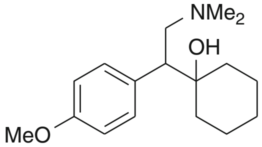
fluoxetine



paroxetine



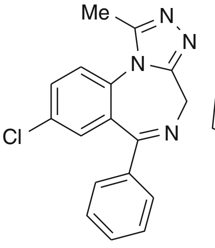
sertraline



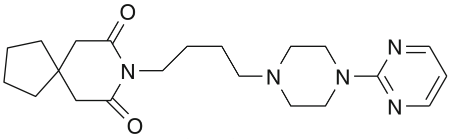
venlafaxine

Anxiolytics

Alprazolam (Xanax®), a benzodiazepine derivative, is used for the treatment of both anxiety and panic disorder while buspirone (Buspar®) is indicated for the treatment of anxiety. The mechanism of action of buspirone is distinct from that of the benzodiazepines and is believed to be mediated mainly through modulation of serotonergic neurotransmission via its interaction with the 5-HT$_{1A}$ serotonin receptor subtype.



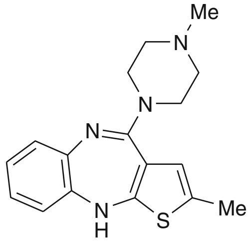
alprazolam

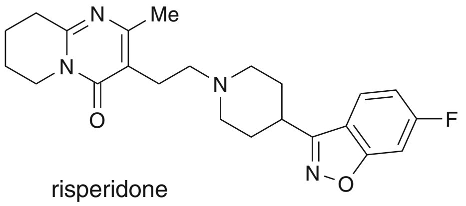


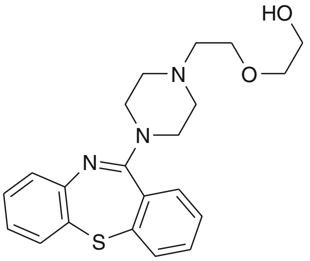
buspirone

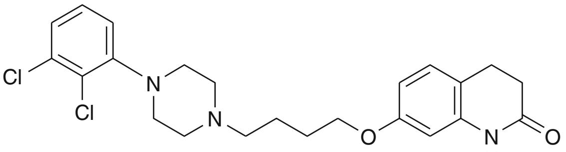
Bipolar Disorders, Schizophrenia, and Epilepsy

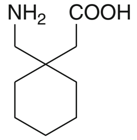
The thienobenzodiazepine derivative olanzapine (Zyprexa®), the benzisoxazole risperidone (Risperidal®), and the thiazepine quetiapine (Seroquel®) are atypical antipsychotic agents. Olanzapine is used in the treatment of bipolar disorder and risperidone is useful in the management of schizophrenia. It is believed that both compounds exert their beneficial effects through antagonism of serotonergic and dopaminergic receptors. Quetiapine is approved in the US for use as monotherapy in both bipolar mania and depression, offering potential compliance advantages. The precise mechanism of action is unknown, although its antidepressant activity may be linked to its higher affinity for serotonergic than dopaminergic receptors. A newer agent for the treatment of schizophrenia is aripiprazole (Abilify®). It is believed the pharmacological effects of aripiprazole are mediated through a combination of partial agonist activity at the dopamine D_2 and serotonin 5-HT$_{1a}$ receptors and antagonism at the serotinergic 5-HT$_2$ receptor. The γ-aminobutyric acid derivative (GABA) gabapentin (Neurontin®) is useful in the treatment of epilepsy. Although gabapentin is structurally related to GABA, it has no GABA-ergic activity and the mechanism of its anticonvulsive actions is currently unknown.


olanzapine


risperidone

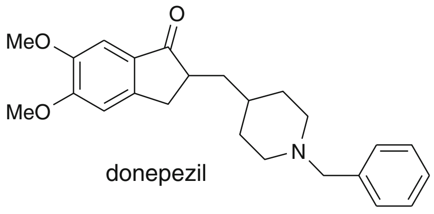

quetiapine

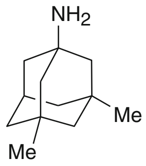

aripiprazole


gabapentin

Alzheimer's Disease

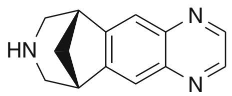
The indanone derivative donepezil (Aricept®) is an acetycholinesterase inhibitor and is structurally unrelated to other cholinesterase inhibitors. Because it increases the concentration of the neurotransmitter acetycholine at cholinergic sites, it is used in the treatment of Alzheimer's disease (dementia) and appears to provide modest benefits in cognition and behavior. Another agent useful in the treatment of Alzheimer's disease is the adamantylamine derivative, memantine (Namenda®). Memantine is an *N*-methyl-D-aspartate (NMDA) receptor antagonist that is thought to exert its pharmacological effect by blocking the excitatory action of the amino acid glutamate at this receptor. However, memantine has shown no evidence of preventing or slowing the neurodegeneration associated with Alzheimer's disease.


donepezil


memantine

Smoking Cessation

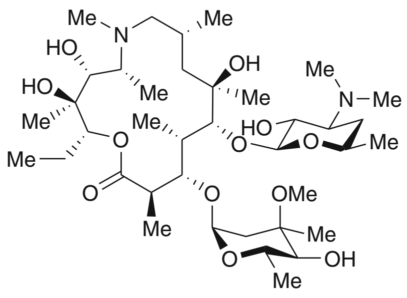
The pyrazino-benzaepine derivative varenicline (Chantix®) is a partial agonist of the $\alpha_4\beta_2$ subtype of the nicotinic acetylcholine receptor that activates the receptor without creating the full effect on the release of dopamine that is associated with nicotine. Varenicline is indicated as a treatment for smoking cessation.


varenicline

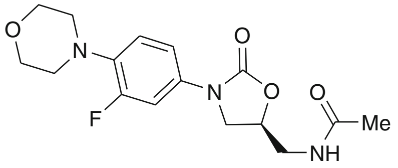
Infectious Diseases

Antibacterial Agents

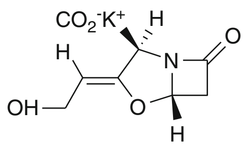
The primary driver for research in the antibacterial area over the past decade has been the emergence of organisms resistant to the existing antibiotics. Important members of the armamentarium of antibacterials include azithromycin (Zithromax®), linezolid (Zyvox®), amoxicillin/potassium clavulanate (Augmentin®), levofloxacin (Levaquin®), ciprofloxacin (Cipro®), and daptomycin (Cubicin®). Azithromycin is a semisynthetic 9α-azalide analog of erythromycin possessing improved resistance to acid-mediated degradation, increased activity against gram-negative organisms, and improved pharmacokinetics. Its indications include the treatment of mild to moderate upper and lower respiratory tract infections and otitis media in pediatric patients. Interestingly, azithromycin tends to concentrate in lung tissue which is the site of many bacterial infections. Linezolid is a completely synthetic oxazolidinone derivative which has a unique mechanism of action, resulting in a low potential for cross resistance to other antibacterial agents. Linezolid is indicated for the treatment of community-acquired pneumonia, MRSA, and VRE infections and has the distinctive characteristic of being nearly 100% orally bioavailable. Amoxicillin/clavulanate is one of the few approved drug mixtures and is a drug of choice for the treatment of acute otitis media. Amoxicilin is a beta lactam antibiotic that works by inhibiting bacterial cell wall synthesis. However, bacteria have developed enzymes called beta lactamases that inactivate amoxicillin. Clavulanate has no antibacterial activity of its own but inactivates beta lactamases thus augmenting the efficacy of the amoxicillin. Amoxicillin/clavulanate is also an alternative treatment for anthrax exposure in pediatrics. Levofloxacin is a totally synthetic fluoroquinolone antibiotic indicated for the treatment of pneumonia; chronic bronchitis; and sinus, urinary tract, kidney, prostate, and skin infections. It is administered by injection and the balance of superior efficacy coupled with reduced side effects positions levofloxacin for administration after other antibiotics have failed. Ciprofloxacin is another synthetic fluoroquinolone antibacterial that acts as a DNA gyrase inhibitor. It is active against a broad range of pathogens, including both gram-positive and gram-negative aerobic bacteria and is effective against urinary tract and lower respiratory tract infections. Daptomycin, a naturally occurring cyclic lipopeptide of molecular formula $C_{72}H_{101}N_{17}O_{26}$ is a cidal antibacterial agent used for the treatment of infections caused by gram-positive bacteria, including those that are resistant to standard antibacterial regimens. It is a large molecule with a complex structure, so it is not depicted in the scheme.



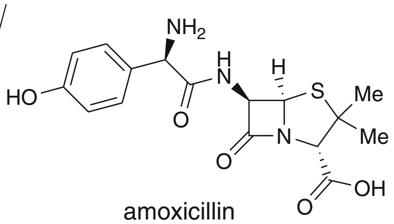
azithromycin



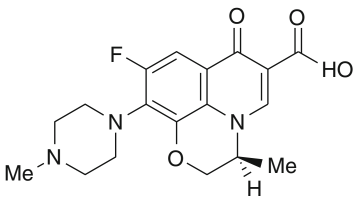
linezolid



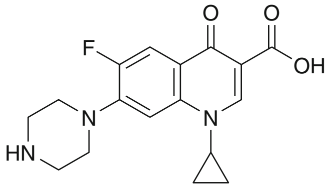
potassium clavulanate



amoxicillin



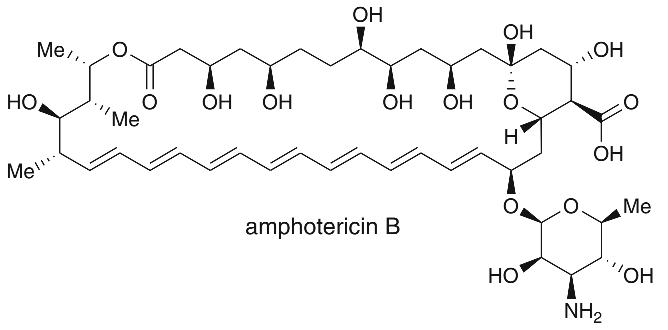
levofloxacin

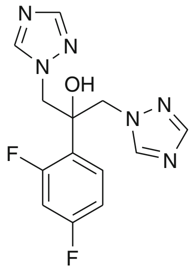


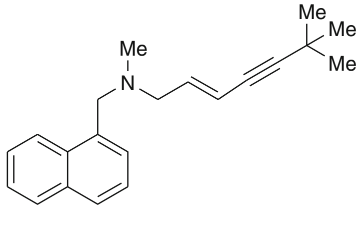
ciprofloxacin

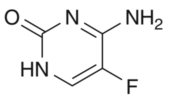
Antifungal Agents

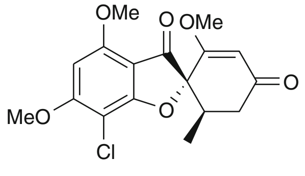
The increasing immunocompromised patient population has exacerbated the need for effective antifungal agents to combat the opportunistic fungal infections that arise in these patients. Current therapy for localized or systemic fungal infections comprises several groups including azoles, polyenes, allylamines, flucytosine, and griseofulvin. Many of these antimycotic agents target the production and integrity of the fungal cell wall, ergosterol biosynthesis and membrane function, nucleic acid biosynthesis, and mitosis [7]. In particular, fluconazole (Trican®), an achiral triazole derivative, inhibits the fungal cytochrome P450 enzyme 14α-demethylase, effectively preventing the conversion of lanosterol to ergosterol, an essential component of the fungal cytoplasmic membrane. Fluconazole is indicated for the treatment of systemic candidiasis as well as meningitis caused by *Cryptococcus neoformans*. Amphotericin B is a naturally occurring polyene antimycotic used topically for localized fungal infections or intravenously for systemic fungal infections. Amphotericin B irreversibly binds to ergosterol forming a transmembrane channel that results in monovalent cation leakage, which is thought to contribute to cell death. Similar to other allylamines, terbinafine (Lamisil®) inhibits ergosterol synthesis by binding to squalene epoxidase, an enzyme that is part of the fungal cell membrane biosynthesis pathway. This is thought to change cell membrane permeability, causing fungal cell lysis. Terbinafine's highly lipophilic nature results in its accumulation in skin, nails, and fatty tissues. It can be administered either systemically or topically and is a treatment of choice for dermatophytosis. Flucytosine (Ancobon®) effectively inhibits fungal protein synthesis by the action of multiple metabolites on fungal RNA and DNA. However, due to documented primary and secondary resistance, flucytosine is not an optimal monotherapy. Instead, flucytosine is often co-administered with amphotericin B. Indications include treatment of serious infections caused by susceptible strains of *Candida* or *Cryptococcus neoformans*. It can also be used for the treatment of chronomycosis (chromoblastomycosis), if susceptible strains cause the infection. Finally, griseofulvin exerts its activity against dermatophytes via binding to tubulin, thus interfering with microtubule function and, consequently, inhibiting mitosis. Its primary indication is for the treatment of mycoses of the hair, nail, and skin. In addition to continued research efforts directed at targeting historical mycotic vulnerabilities, there are signs of an emerging paradigm targeting virulence as a potential treatment for fungal diseases.


amphotericin B


fluconazole

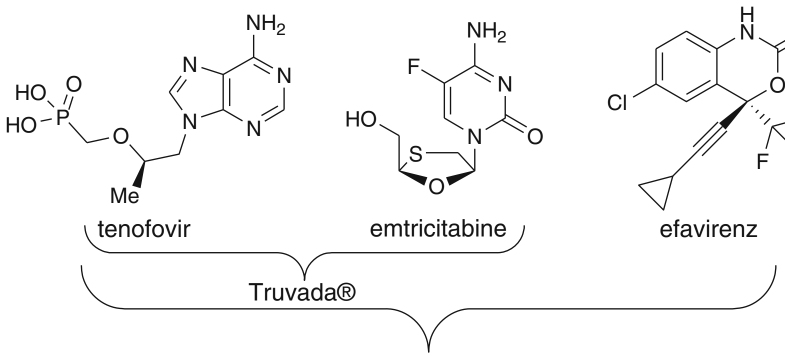

terbinafine


flucytosine

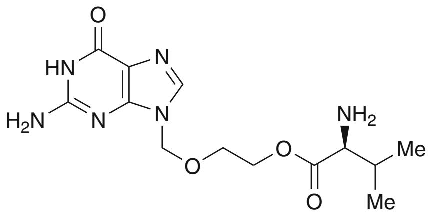

griseofulvin

Antiviral Agents

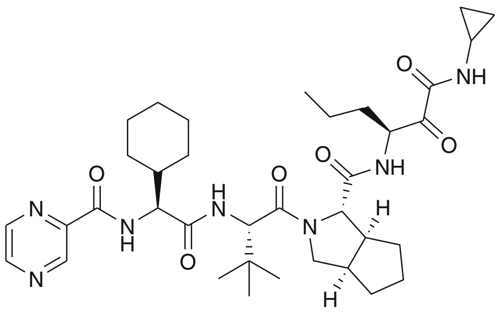
Antiviral research has become a major focus in the pharmaceutical industry over the past decade as evidenced by the marketing of a plethora of antiviral agents active against HIV/AIDS, hepatitis B and C, influenza, and respiratory syncytial virus (RSV). Tenofovir combined with emtricitabine (Truvada®), efavirenz combined with tenofovir and emtricitabine (Atripla®), valaciclovir (Valtrex®), and pegylated interferon alfa-2a (Pegasys®/ pegylated interferon alfa-2b (Peg-Intron®) represent the top four selling antiviral agents worldwide. The combination of nucleoside analog transcriptase inhibitors tenofovir and emtricitabine into a single pill (truvada) provides a well-tolerated and effective dosage form that increases compliance by the patient due to the reduced pill burden. In addition to its efficacy as a treatment for HIV, Truvada® has also been shown to have potential as a preventative measure if administered regularly prior to exposure. The most convenient triple therapy, Atripla®, provides the combination of efavirenz, a non-nucleoside reverse transcriptase inhibitor, with tenofovir and emtricitabine in the first effective single pill cocktail for once a day HIV treatment. Atazanavir sulfate (Reyataz®) or Lopinavir (sold as Kaletra® which is a coformulation with ritonavir) are members of a group of HIV protease inhibitors that are used in conjunction with other anti-retroviral chemotherapeutic agents for the treatment of HIV in adults and adolescents. Atazanivir is a Phe-Pro scissile bond peptidomimetic that clinical data suggests may have a more favorable hypertriglyceridemia profile as compared to other protease inhibitors. Darunavir (Prezista®) is the most recently marketed HIV protease inhibitor. Valaciclovir is an antiviral compound used in the treatment of herpes simplex and herpes zoster that has been marketed generically since late 2009. The valine moiety of this prodrug, which facilitates oral absorption by active transport, is cleaved in vivo converting it to acyclovir. Pegylated interferon α-2a or α-2b bolsters the immune system and weekly injections combined with twice daily dosing of ribavirin, a synthetic nucleoside analog, was the current optimal therapy for the treatment of hepatitis C virus. The serious side effects associated with this treatment and the modest efficacy (40–50%) have stimulated a substantial effort toward developing improved HCV treatments and several compounds are in late stage clinical trials, the NS3 protease inhibitors telaprevir or Boceprevir have now been approved for use in combination with pegylated interferon and ribavirin. Oseltamivir phosphate (Tamiflu®) is an orally active anti-influenza drug that inhibits neuraminidase and is being stockpiled for use in future outbreaks of influenza.



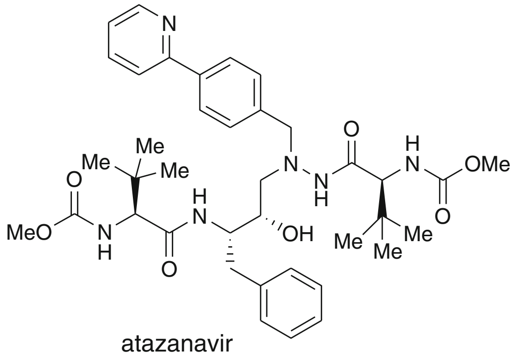
Atripla®



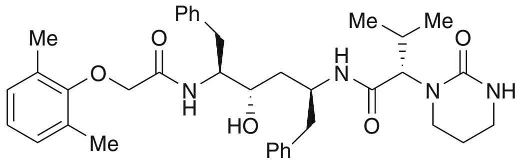
valaciclovir



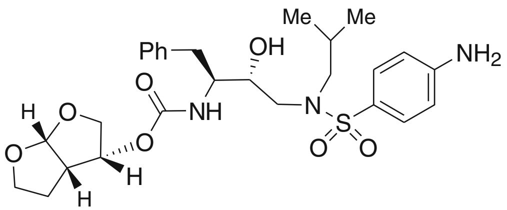
telaprevir



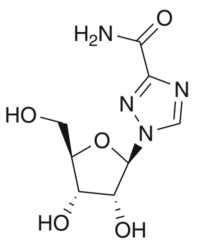
atazanavir



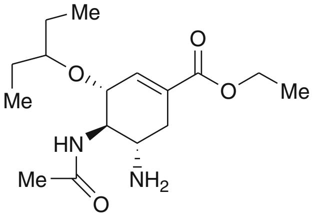
lopinavir



darunavir



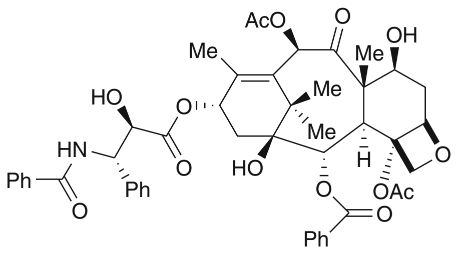
ribavirin



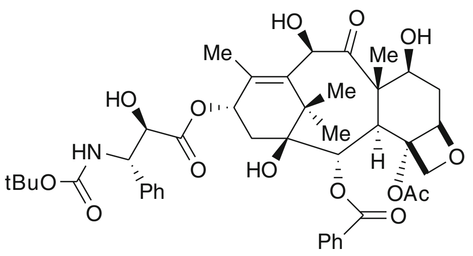
oseltamivir

Antineoplastic Agents

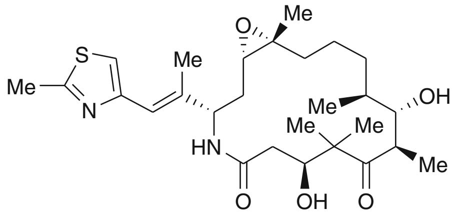
Paclitaxel (Taxol®) and irinotecan (Camptosar®) were discovered as a result of natural product extract screening. Paclitaxel is a naturally occurring diterpene that exerts its antineoplastic effect via stabilization of the mitotic spindle during cell replication. It is used for the treatment of non-small cell lung, breast, ovarian, and esophageal carcinomas as well as Kaposi's sarcoma. The semi-synthetic taxane analog docetaxel (Taxotere®), the epothilone analog ixabepilone (Ixempra®), and eribulin mesylate (Halaven®) have a similar mechanism of action in that they also target the microtubules used to form the mitotic spindle. Halaven® received FDA approval in November 2010 and is a synthetic analog of a marine natural product isolated from the sea sponge *Halichondria okadai*. Irinotecan is a prodrug that upon release of the piperidinylpiperidine carbamate moiety reveals the pharmacologically active parent SN-38 which is a derivative of the naturally occurring camptothecin. Irinotecan exerts its antineoplastic activity via the inhibition of Type I DNA topoisomerase and stabilization of the transiently formed topoisomerase I/DNA-cleavable complex. Tamoxifen (Nolvadex®), a nonsteroidal antiestrogen chemotherapeutic possessing both agonistic and antagonistic properties, is used for the treatment and prevention of breast cancer. Imatinib (Gleevec®), an inhibitor of the Bcr-Abl tyrosine kinase, is used for the treatment of chronic myelogenous leukemia. Dasatinib (Sprycel®) inhibits Bcr-abl, Src family kinases and others and has recently been approved for first-line treatment of the same indication. Bortezomib (Velcade®), a ubiquitin proteosome inhibitor administered parenterally, is used for the treatment of multiple myeloma in patients who have been refractory to other chemotherapeutic regimens. Erbitux (Cetuximab®), a human-murine chimeric monoclonal antibody that blocks the epidermal growth factor receptor (EGFR), was developed for the treatment of irinotecan-refractory colorectal cancer. This agent is also used in patients who are intolerant of irinotecan-based therapy. A number of small molecule kinase inhibitors have been approved for oncology applications including the small molecule EGFR tyrosine kinase inhibitor gefitinib (Iressa®) used to treat non-small cell lung cancer. Several additional important oncology drugs are described in the section on biologics.



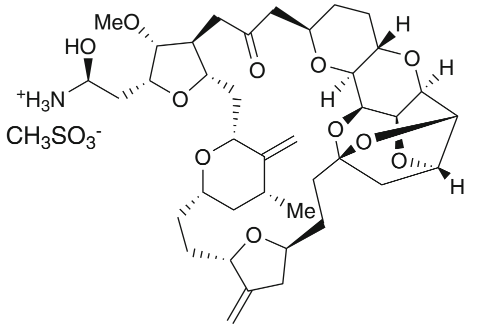
paclitaxel



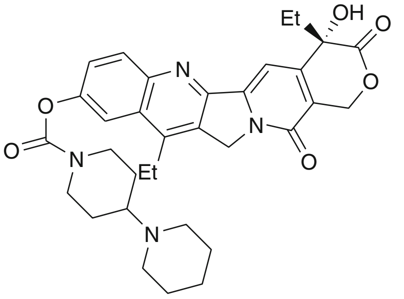
docetaxel



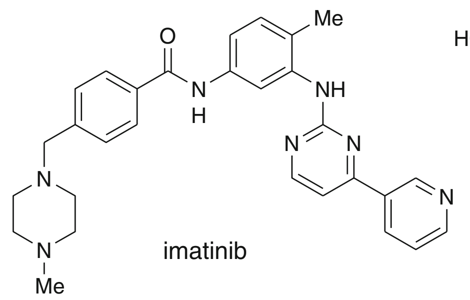
ixabepilone



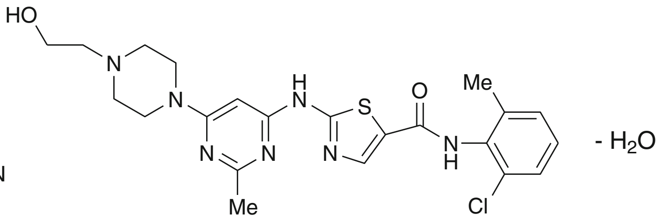
eribulin mesylate



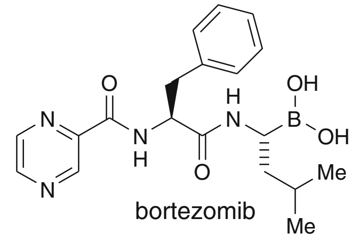
irinotecan



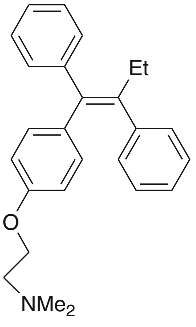
imatinib



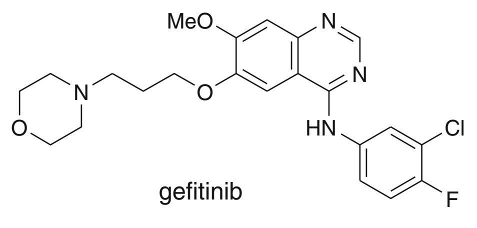
dasatinib



bortezomib



tamoxifen

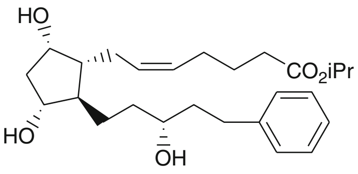


gefitinib

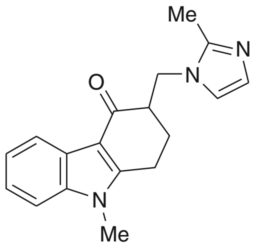
Miscellaneous Agents

Glaucoma and Nausea

Latanoprost (Xalatan®) is a topical, ocular hypotensive agent used to treat glaucoma. The compound is a synthetic analog of the naturally occurring prostaglandin PGF2α and is thought to reduce intraocular pressure by increasing outflow of the aqueous humor. Odansetron (Zofran®) is a selective serotonergic 5-HT_3 receptor antagonist that is used to ameliorate the nausea and vomiting frequently associated with chemotherapy.



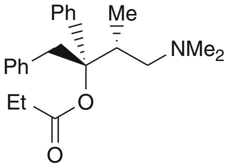
latanoprost



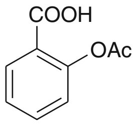
odansetron

Analgesics

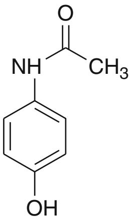
Propoxyphene, along with aspirin, acetominophen, and ibuprofen are among the most widely used agents for the treatment of mild to moderate pain.



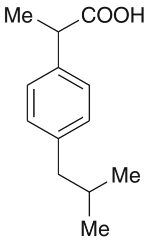
propoxyphene



aspirin



acetaminophen



ibuprofen

Biologics

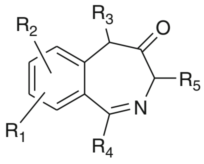
A key question that will affect the future of medicinal chemists in the pharmaceutical industry is what will be the balance of research funding and efforts between discovering and developing traditional small molecule drugs and biologics. Biologics are defined as a preparation, such as a drug, a vaccine, or an antitoxin, that is synthesized from living organisms or their products and used as a diagnostic, preventive, or therapeutic agent, Biologics are typically large macromolecules that, to date, have been limited to injectable administration methods and not been suitable for oral dosing. In 2006, none of the top 20 brand name drugs by retail sale dollars was a biologic but in 2009 that number rose to 8 out of 20. These include the tumor necrosis factor inhibitor etanercept, marketed as Enbrel® for the treatment of rheumatoid arthritis; the immunosuppresive agent infixamade (Remikade®); the specific antirheumatic agent adalimumab (Humira®); the antitneoplastics bevacizumab (Avastin®), a VEGF kianase inhibitor; rituximab (Mabthera®); the monoclonal antibody trastuzumab, marketed as Herceptin® for the treatment of breast cancer; the heparin enoxeparin (Lovenox®) for anticoagulative indications; and human insulin analogs of which Lantus® is representative. The role that medicinal chemists may play in the future design and development of biologics continues to be defined.

Small Molecule High Throughput Synthesis

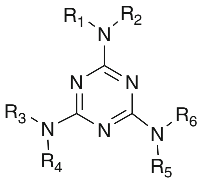
The field of non-oligomeric, small molecule, high throughput synthesis came into existence in 1992 and has developed into an important strategy for both lead identification and lead optimization [8, 9]. Synthetic libraries have been categorized as follows: (1) discovery libraries: opportunistic libraries synthesized with no preconceived notion about which molecular target it may be active against. These libraries tend to be large in size, typically >5,000 compounds; (2) targeted libraries: these libraries are biased in their design and contain a pharmacophore known to interact with a specific target, or a family of targets; (3) optimization libraries: libraries are constructed around an existing lead with the intent to improve potency, selectivity, and pharmacokinetic profile. These libraries tend to be smaller in size, usually ranging from tens up to a few hundred compounds.

Discovery Libraries

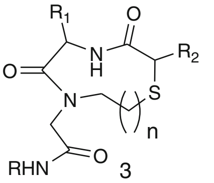
Researchers have employed several different strategies to create populations of molecules that are used for broad-based screening. One strategy is to synthesize libraries based on "privileged pharmacophores," heterocyclic scaffolds such as benzodiazepines (**1**) or triazines (**2**) that have been shown to possess biological activity in multiple settings. A second strategy is to design scaffolds or templates for library synthesis that are based on important molecular recognition recognitions elements including, for example, peptide β-turn mimetics (**3**). The discovery of chemical encoding technologies and radio-frequency (Rf)-encoded synthesis has had a major impact on the synthesis of lead discovery libraries. Both technologies utilize resin-based split-pool synthesis to prepare large arrays of compounds with libraries of >50,000 members typically prepared using chemically encoded beads. Chemical encryption, in the form of unique chemical markers (tags), is associated with synthetic identity of the library member tethered to the resin bead. However, the restricted amounts typically prepared (200–300 μg), the lack of analytical characterization of library members, and the requirement of a specialized screening format for chemically encoded libraries have limited the utility of this technology. Radio frequency encoded synthesis overcomes some of these limitations while retaining the efficiency of split-pool synthesis. Libraries of 10,000–15,000 members can be prepared, with individual members quantified and characterized by LC/MS. Most pharmaceutical companies have utilized Rf-encoded synthesis in their lead identification efforts.



1



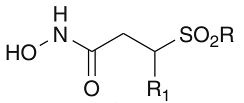
2



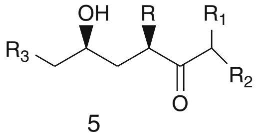
3

Targeted Libraries

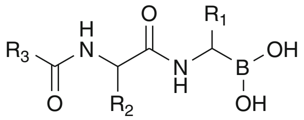
Libraries targeted towards proteolytic enzymes, non-proteolytic enzymes, G-protein coupled receptors (GPCRs), and ion channels have been very successful in lead identification. Libraries of hydroxamates (**4**), hydroxy ethylenes (**5**), boronic acids (**6**), and β-keto sulfonamides (**7**) have been prepared as inhibitors of metallo-, aspartyl, serine, and cysteine proteases, respectively, using either solid phase or solution phase synthesis.



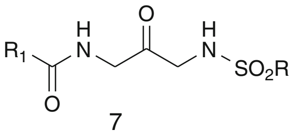
4



5

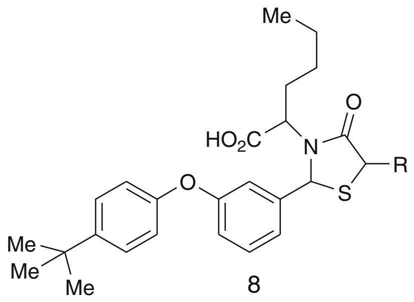


6



7

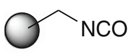
Structure-based design has been effectively utilized in the synthesis of inhibitors of nonproteolytic enzymes. Inhibitors of MurB, an essential bacterial enzyme required for the biosynthesis of peptidoglycan, were identified using the X-ray structure of the enzyme as a tool for library design. Thiazolidinone inhibitors (**8**) thus identified are the first examples of small molecule inhibitors of MurB.



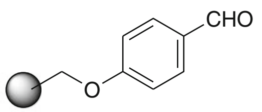
8

Substituted indoles ($5HT_{2a}$; D_4 and α_{2a} receptor antagonists) and piperazines (opioid antagonists) are representative chemotypes targeted towards GPCRs.

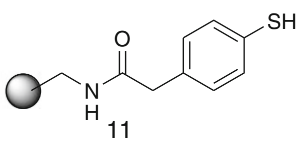
Advances made in solid-phase extraction (SPE) and in the development of resin-based scavengers have increased the versatility of chemistries implemented for the synthesis of targeted libraries. A combination of cation exchange (SCX) and anion exchange (SAX) resin was effectively utilized to prepare libraries of highly substituted amides. Scavenger resins and polymer-bound reagents are routinely used to prepare medium-sized (500–1,000 members) libraries. Polymer-bound isocyanates (**9**) and aldehydes (**10**) are used to remove amines from reaction mixtures, while polymer-bound thiols (**11**) are used to scavenge halides.



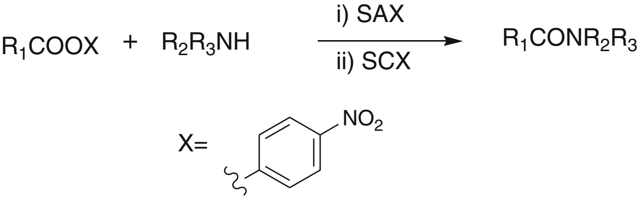
9



10

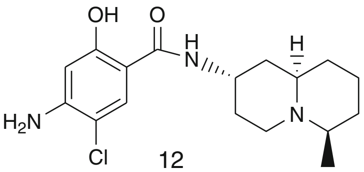


11

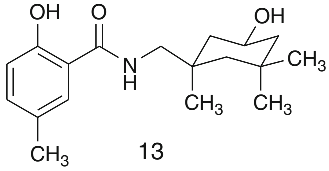


Optimization Libraries

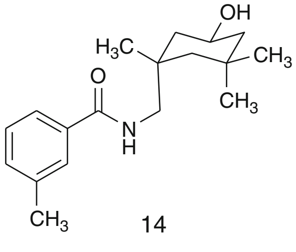
Starting with a lead structure, researchers have demonstrated that parallel synthesis can be effectively utilized to optimize biological activity while reducing timelines for the iterative process. A parallel synthesis strategy was implemented to identify structurally simpler and more potent analogs of the influenza hemagglutinin inhibitor **12**, $EC_{50} = 4$ μg/mL. SPE was used to automate the preparation of >400 analogs, resulting in the identification of **13**, $EC_{50} = 20$ ng/mL, and **14**, $EC_{50} = 20$ ng/mL as simpler and more potent influenza inhibitors.



12

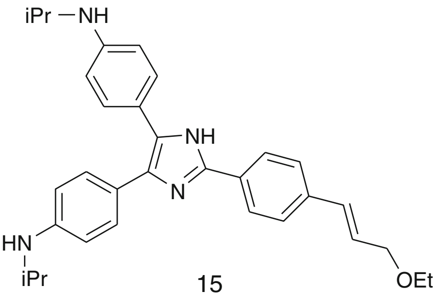


13

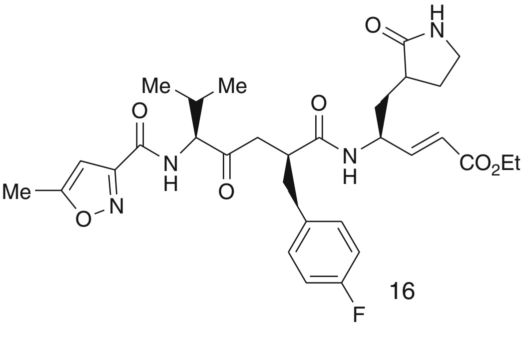


14

During the past 10 years, the pharmaceutical industry has expended significant resources in developing and assimilating technologies to increase synthesis throughput and decrease preclinical time lines. There are numerous examples in the literature demonstrating the effective use of high-throughput synthesis for lead discovery and optimization. There are two publicly known examples of clinical candidates that have emerged directly from optimization libraries. OC144-093 (**15**) is a P-glycoprotein modulator, $IC_{50} = 50$ nM, while AG-7088, (**16**) a potent rhinovirus 3C protease inhibitor, k_{obs}/1,470,000 m/s, advanced into clinical trials for the treatment of the common cold.



15



16

Chemical Process R&D in the Pharmaceutical Industry

Most of the active pharmaceutical ingredients (APIs) of commercially available pharmaceuticals are manufactured either by chemical syntheses or by microbial fermentations. However, some of the active ingredients are directly obtained from natural sources. This section addresses the development and manufacture of APIs [10]. Drug candidates in development are becoming increasing complicated, with multiple chiral centers and/or difficult to assemble heterocyclic systems. Incessant demand to shorten the timelines for the discovery, development, and launch of NCEs coupled with environmental concerns has necessitated the development of higher yielding, more robust, and environmentally friendly processes in shorter times. The success of a pharmaceutical company greatly depends not only on discovering important NCEs that fulfill unmet medical need but also on its ability to design, optimize, and scale up a chemical process to commercial manufacturing with increasing rapidity. The chemical manufacturing process must be a robust procedure capable of operating routinely in a manufacturing environment. The recent trend towards outsourcing the preparation of key intermediates (or even the entire API) during the early phases of a program has added additional time constraints on the development of robust processes. Often the early preclinical and phase 1 supplies are made via "fit for purpose" strategies, and the process chemist is handed the program well into development, with shortened timelines before process lock.

Considerable attention has to be given to various parameters in developing a manufacturing process for an API including, for example, efficiency of the synthesis, availability and cost of starting materials, toxicity of the reagents, stability and toxicity profiles of intermediates, formation of by-products, and safe disposal of waste materials. Data from various aspects of chemical process development, including process structure and flowsheet, operational guidelines, optimization, process management, process control, fault diagnostics, and equipment management need to be in place in order to support a smooth transition from the laboratory to the manufacturing plant. Safety is another critical factor requiring consideration for large-scale manufacture. All reactions should undergo a process hazard analysis for incident-free and successful plant implementation before scale up. The use of automation in accelerating the design of cost-effective and well-understood synthetic processes is growing very rapidly [12]. Automation concepts and tools such as statistical design of experiments (DOE) and parallel experimentation using in-house built reactor blocks or commercially available robots play an ever expanding role in increasing the productivity of process R&D with respect to speed and economics, as well as obtaining process knowledge. Production and logistical processes are becoming more complex due to an increasing number of products and smaller batch sizes. To manage this, supply chain optimization and production planning activities need to be addressed. Production simulation can be used for performance measurements and capacity assessments of manufacturing as well as material and information flow processes. Some applications of production simulation include bottleneck analysis, examination of process alternatives, and assessment of investment decisions.

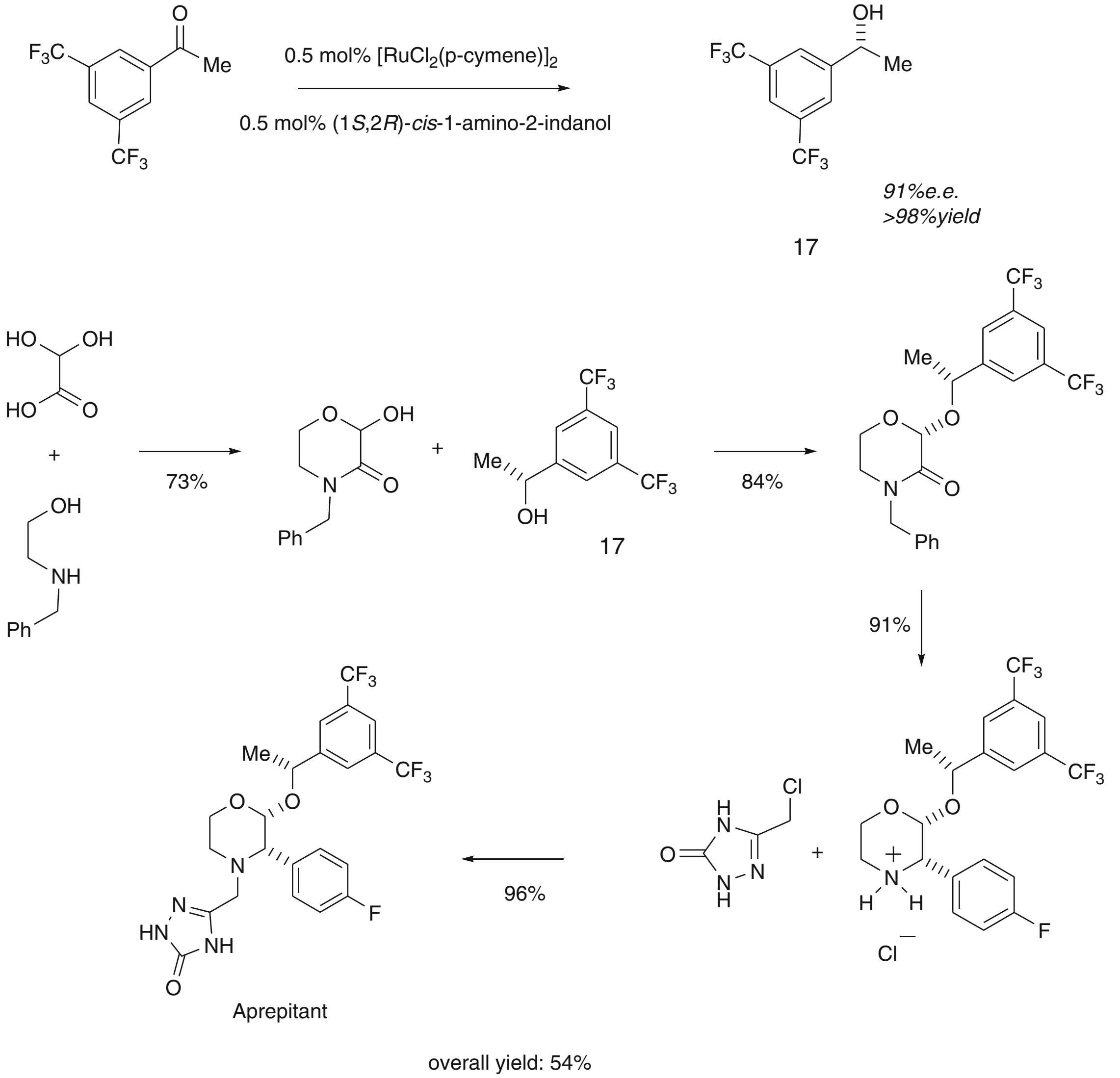
Batch process development is a fairly complex series of engineering tasks. In the pharmaceutical industry, the production of a majority of APIs is based on a batch concept. This concept offers many advantages with respect to quality assurance as an individual batch can be accepted or rejected. However, the scale up of the batch size without proper controls may lead to problems. As a new product approaches commercialization, a database of appropriate controls and processing parameters is developed for each late-stage step. This leads to a 'design space' of acceptable processing parameters, ensuring that the desired yield and quality is obtained. The FDA has encouraged manufacturers to adopt quality by design (QbD) principles which are described in a document *ICH Q8R1* that was authored by pharmaceutical industry experts from the United States, Europe, and Japan with the goal of standardizing development parameters so that they will be applicable worldwide. Examples of topics covered include quality management, personnel qualifications and hygiene, process equipment cleaning and calibration, documentation and records, material management, production and in-process controls, packaging and labeling, and many others. In QbD, a design space is developed, either by traditional one-variable experimentation or, more commonly, by multivariate DOE. By thoroughly defining the design space, product quality can be assured by maintaining processing parameters within the acceptable ranges. In-line testing (also known as process analytical techniques) has become a mainstay of modern manufacturing, lessening the need for traditional off-line testing.

Quality standards for pharmaceutical products have always been (understandably) more rigorous than for other commercial chemicals. The QbD paradigm is a step towards limiting batch release testing; however, the goal of eliminating analytical release may prove elusive. Even a well-defined process operating within its quality design space will require that the raw materials used meet established specifications. This may push analytical testing earlier in the synthesis, perhaps even onto the suppliers, but will not eliminate it.

Potential genotoxic impurities in both marketed and experimental medicines have become a large regulatory and quality issue. Indeed, nearly an entire issue of the respected journal *Organic Process Research & Development* was devoted to the subject [11]. Potential genotoxic impurities are required to be below the default threshold of toxicological concern limit (TTC) of 1.5 μg/day. The genotoxic potential of intermediates, impurities, and reagents in a process to make API must all be assessed. An initial in silico evaluation using programs such as DEREK, TOPKAT, and MCASE provide "alerting structures" that then require further investigation, usually via Ames testing. Compounds testing Ames positive are considered putative mutagens and must be controlled at the TTC limit of 1.5 μg/day. The level of detection required for high-dose compounds can be <1 ppm, demanding ever more sensitive analytical techniques. Indeed, satisfactory routes to a given API may be discarded because of potential genotoxic impurities if those impurities cannot be removed or adequately controlled.

Crystallization, filtration, drying, and milling (if required) are other important factors that need to be defined well before a process to manufacture solid APIs is finalized. Physicochemical properties of APIs play a vital role in providing the pharmaceutical drug products with desired bioavailability, manufacturing properties, and good final product quality. Particle size, density, flowability, polymorphism, hygroscopicity, and stability are critical properties for solid APIs in the development of formulations. Polymorphism is very important in determining the physical properties of various crystal forms of a drug for optimal chemical and formulation processing, as well as for satisfying regulatory and patent issues for producing consistent solid forms of a drug.

Asymmetric enantioselective syntheses have begun to supplant traditional resolution methods for chiral intermediates leading to APIs [13]. The most elegant of these asymmetric approaches involves the use of chiral catalysts. The following flow diagrams highlight the use of enantioselective chiral catalysts in preparation of APIs of some widely used pharmaceutical drugs in today's market. Scheme 11.1 shows the preparation of aprepitant, the active ingredient of the anti-

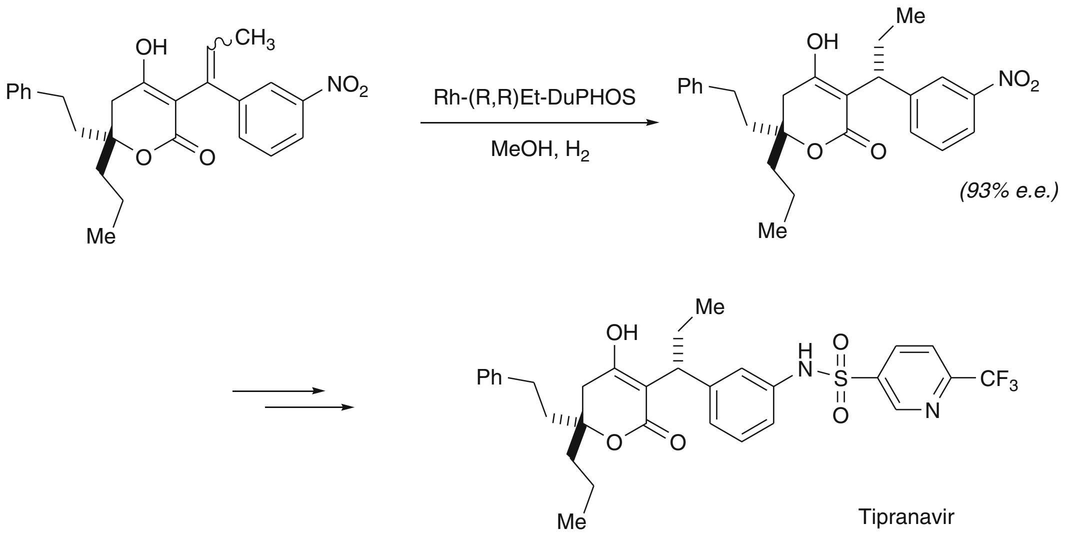


overall yield: 54%

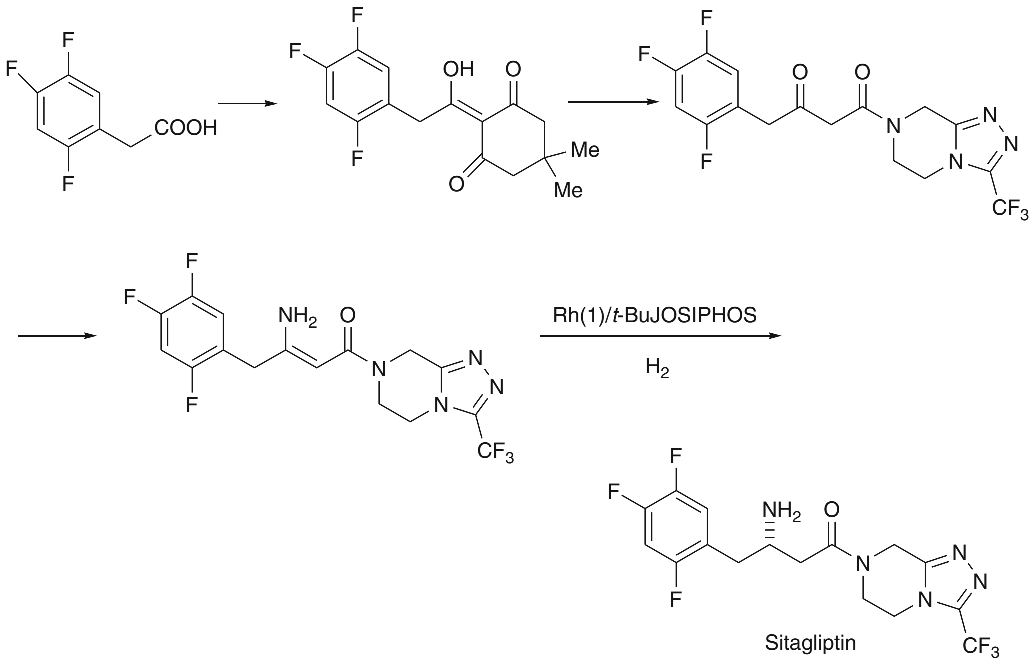
Scheme 11.1 Preparation of aprepitant emetic Emend®. A chiral reduction to form alcohol **17** proceeds in high yield and good ee. Recrystallization increased chiral purity to over 99%. All three chiral centers in the API are derived from the initial chiral reduction [14].

The key step toward the synthesis of the HIV protease inhibitor tipranavir is shown in Scheme 11.2. A rather unique alkene hydrogenation is used for the late-stage installation of a chiral center [15].

An elegant and efficient route to sitagliptin was developed that utilizes a previously undocumented asymmetric hydrogenation of an enamine in the final step [16] (Scheme 11.3).



Scheme 11.2 Key step in the preparation of tipranavir



Scheme 11.3 Preparation of sitagliptin

Conclusion

The discovery and development of novel therapeutic agents by the pharmaceutical industry has afforded physicians an extensive armamentarium to fight a wide range of human diseases. Undoubtedly, there remains the opportunity for even more effective drugs with greater benefit-to-risk ratios than those currently available. The elucidation of the human genome may eventually lead to the identification of many new macromolecular targets for drug intervention and the application of new technologies has, to date, been modestly beneficial to the overall success of drug discovery. The considerable challenges faced by pharmaceutical research will undoubtedly best be addressed by creative and observant scientists operating cohesively and openly in multidisciplinary teams. Medicinal chemists have been and will likely continue to play integral roles in these groups, operating at the forefront of pharmaceutical research to identify innovative drugs that the future demands.

References

1. Krogsgaard-Larsen P, Liljefors T, Madsen U (eds) (1996) A textbook of drug design and development, 2nd edn. Harwood Academic, Amsterdam
2. Spilker B (1989) Multinational drug companies; issues in drug discovery and development. Raven, New York
3. Wermuth CG (ed) (1996) The practice of medicinal chemistry. Academic, San Diego
4. Wess G, Urmann M, Sickenberger B (2001) Medicinal chemistry: challenges and opportunities. Angew Chem Int Ed 40:3341–3350
5. Lipinski CA, Lombardo F, Dominy DW, Feeny PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 23:3–25
6. Stella VJ, Borchardt RT, Hageman MJ, Oliyai R, Maag H, Tilley JW (eds) (2007) Prodrugs: challenges and rewards, vols. 1–2. Springer, New York
7. Gauwerky K, Borelli C, Korting HC (2009) Drug Discov Today 14:214–222
8. Miertus S, Fassina G (eds) (1999) Combinatorial chemistry and technology; principles, methods and applications. Marcel Dekker, New York
9. Dolle R (2001) Comprehensive survey of combinatorial library synthesis: 2000. J Comb Chem 3:477–518
10. Anderson NG (2000) Practical process research and development. Academic, New York
11. Laird T (2010) Organic process research and development 14(4): 942–1045
12. Rubin AE, Tummala S, Both DA, Wang C, Delaney E (2006) Emerging technologies supporting chemical process R&D and their increasing impact on productivity in the pharmaceutical industry. Chem Rev 106:2794–2810
13. Farina V, Reeves JT, Senanayake CH, Song JJ (2006) Asymmetric synthesis of active pharmaceutical ingredients. Chem Rev 106:2734–2793
14. Brands KM, Payack JF, Rosen JD, Nelson TD, Candelario A, Huffman MA, Zhao MM, Li J, Craig B, Song ZJ, Tschaen DM, Hansen K, Devine PN, Pye PJ, Rossen K, Dormer PG, Reamer RA, Welch CJ, Mathre DJ, Tsou NN, McNamara JM, Reider PJ (2003) Efficient synthesis of NK(1) receptor antagonist aprepitant using a crystallization-induced diastereoselective transformation. J Am Chem Soc 125:2129–2135
15. Hewitt Bradley D, Burk Mark J, Johnson Nicholas B (2000) Improved process for asymmetric hydrogenation. Pharmacia & Upjohn. WO 00/055150, Sep, 21 2000.
16. Hansen KB, Hsiao Y, Xu F, Rivera N, Clausen A, Kubryk M, Krska S, Rosner T, Simmons B, Balsells J, Ikemoto N, Sun Y, Spindler F, Malan C, Grabowski EJ, Armstrong JD III (2009) Highly efficient asymmetric synthesis of sitagliptin. J Am Chem Soc 131:8798

Manufactured Textile Fibers*

12

Bhupender S. Gupta

Textile Background

The first conversion of naturally occurring fibers into threads strong enough to be looped into snares, knit to form nets, or woven into fabrics is lost in prehistory. Unlike stone weapons, such threads, cords, and fabrics—being organic in nature—have in most part disappeared, although in some dry caves traces remain. There is ample evidence to indicate that spindles used to assist in the twisting of fibers together had been developed long before the dawn of recorded history. In that spinning process, fibers such as wool were drawn out of a loose mass, perhaps held in a distaff, and made parallel by human fingers. (A maidservant so spins in Giotto's *The Annunciation to Anne*, ca. AD 1306, Arena Chapel, Padua, Italy [1].) A rod (spindle), hooked to the lengthening thread, was rotated so that the fibers while so held were twisted together to form additional thread. The finished length then was wound by hand around the spindle, which, in becoming the core on which the finished product was accumulated, served the dual role of twisting and storing and, in so doing, established a principle still in use today. (Even now, a "spindle" is 14,400 yards of coarse linen thread). Thus, the formation of any threadlike structure became known as spinning, and it followed that a spider spins a web, a silkworm spins a cocoon, and manufactured fibers are spun by extrusion, although no rotation is involved.

It is not surprising that words from this ancient craft still carry specialized meanings within the textile industry and have entered everyday parlance, quite often with very different meanings. Explanations are in order for some of the words used in the following pages. For example, as already indicated, "spinning" describes either the twisting of a bundle of essentially parallel short pieces of wool, cotton, or precut manufactured fibers into thread or the extrusion of continuous long lengths of manufactured fibers. In the former case, the short lengths are known as "staple" fibers, and the resulting product is a "spun yarn," whereas the long lengths are called "continuous-filament yarn" or merely "filament yarn." Neither is called a "thread," for in the textile industry that term is reserved for sewing thread and rubber or metallic threads. Although to the layperson "yarn" connotes a material used in hand knitting, the term will be used in the textile sense hereinafter.

Before manufactured fibers are discussed, it is necessary to define some terms.[1] The "denier" of a fiber or a yarn defines its linear density, that is, the mass in grams of a 9,000 m length of the material at standard conditions of 70°F and 65% relative humidity. Although denier is actually a measure of linear density, in the textile industry, the word connotes the size of the filament or yarn. Fibers usually range from 1 to 15 denier and yarns from 15 to 1,650. Single fibers, usually 15 denier or larger, used singly, are termed "monofils." The cross-sectional area of fibers of identical deniers will be inversely related to their densities, which range from 0.92 g/cc for polypropylene to 2.54 g/cc for glass. The approximate densities of some of the other commonly used fibers are given in Table 12.1. Because by definition denier is measured at standard conditions, it describes the amount of "bone-dry" material plus the moisture regain, which ranges from zero for glass and polypropylene to 14% for rayon. It should be mentioned that some years ago, scientific organizations throughout the world accepted the word "tex," this being the mass (g) of 1 km of

---

*The author dedicates this chapter to the memory of late Dr. Robert W. Work, Professor Emeritus, a longtime friend and mentor.

B.S. Gupta (✉)
College of Textiles, North Carolina State University,
PO Box 8301, 2401 Research Drive, Raleigh, NC 27695, USA
e-mail: bs_gupta@ncsu.edu

[1] Each year, the ASTM publishes in its *Book of Standards*, the most recent and accepted definition and test methods used in the textile and fiber industries.

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_12, © Springer Science+Business Media New York 2012

Table 12.1 Typical (average) values of tensile and physical properties of some commonly used textile fibers

Fiber	Breaking stress (cN/tex)	Strain to fail (%)	Moisture regain (%)	Density (g/cc)
Natural				
Cotton	40	7	8	1.52
Flax	54	3	12	1.52
Silk	38	23	11	1.33
Wool	14	40	14	1.30
Regenerated				
Acetate	13	25	6	1.32
Rayon	25	20	14	1.51
Synthetic				
Acrylic	26	25	1.5	1.18
Modacrylic	28	32	1.5	1.32
Nylon	50	25	4	1.14
Polyester	50	15	0.4	1.38
Polypropylene	53	17	0	0.92

the material, as a more useful term than denier. “Tex” is an accepted adjunct to the SI or International System of Units, but it has received only limited acceptance in commerce, whereas the SI units are being employed increasingly in scientific organizations. Furthermore, the sizes of cotton, wool, and worsted yarns, and yarns containing manufactured fibers but produced by the traditional cotton, wool, or worsted systems, still are expressed in the inverse-count system that has been used for centuries.

The “breaking tenacity” or more commonly, “tenacity,” is the breaking strength of a fiber or a yarn expressed in force per unit denier, that is, in grams per denier, calculated from the denier of the original unstretched specimen. “Breaking length” expresses the theoretical length of yarn that would break under its own weight and is used mostly in Europe. “Elongation” means “breaking elongation” and is expressed in units of increase in length to break calculated as a percentage of the original specimen length.

Typical force–elongation curves of some manufactured and natural staple fibers and textile-type manufactured filaments are shown in Figs. 12.1 and 12.2. Table 12.1 gives the values of some of the physical and tensile properties of textile fibers.

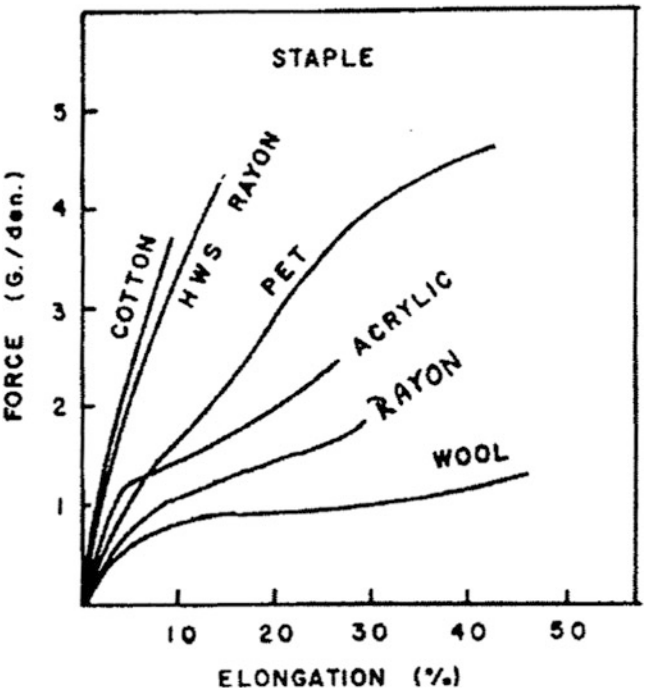


Fig. 12.1 Force–elongation curves of natural and manufactured staple fibers at standard conditions of 70°F and 65% humidity

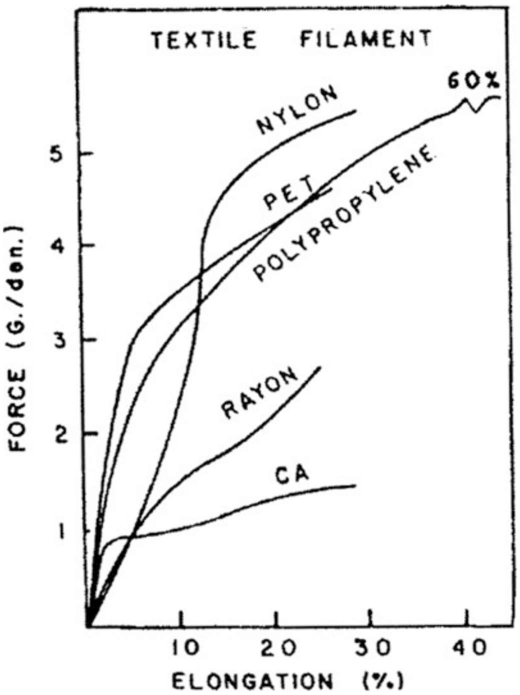


Fig. 12.2 Force–elongation curves of manufactured textile continuous-filament yarns at standard conditions of 70°F and 65% relative humidity

History

Early humans, over time, became aware of the presence and usefulness of fibrous materials available from such sources as seeds, leaves, stems, animal coverings, and cocoons. They learned to spin, weave, knit, felt, or braid these fibers to protect their bodies and improve their lifestyle. A few hundred years ago, it has been suggested, someone took a clue from a busy worm and thought that it should be possible for humans to make a silklike fiber that would be of commercial value. Curiosity combined with simple experiments strengthened that premise and much later led to the development of viscose rayon, only partially a manufactured fiber, in the nineteenth century. Further curiosity, wartime need, and superior commitment by modern researchers led to the synthesis of a totally synthetic fiber in the 1930s. The knowledge gained and the spark thus ignited resulted in the development of other fibers to such an extent that manufactured fibers now dominate the market in the industrialized nations in all major categories of apparel, home furnishing, and industrial end uses.

The story of the development of manufactured fibers is of great historical interest, beginning in 1664 when Robert Hooke, an Englishman, suggested that it should be possible to make a fiber much like silk that could be of value in the marketplace. Audemars, a Swiss chemist, received the first patent for making silklike fiber in 1855. He drew fibers by dipping a needle and pulling it out from a solution of cellulose nitrate containing some rubber.

The credit for using a spinnerette and forcing a solution through it for producing a fiber, however, goes to the English scientist Sir Joseph W. Swan, in the early 1880s. The first person to put the idea into commercial practice was the French chemist Count Hilaire de Chardonnet, who built the first plant to commercially produce a fiber based on regenerated cellulose, called “artificial silk,” at Besançon, France, in 1891. During the last years of the nineteenth century and the beginning of the twentieth century, progress was so rapid that the production of this fiber increased from several thousand pounds in 1891 to over two million in 1910. Commercial production of the fiber in the United States began in 1910 with the opening of the first plant. In 1924, the industry gave the “artificial silk” fiber a new name, rayon.

By the year, 1910, the brothers Camille and Henry Dreyfus had discovered a practical method for producing cellulose acetate polymer and were making plastic film and toilet articles in Basel, Switzerland. During World War I, they built a plant in England to produce acetate dope for painting airplane wings to render them air impervious. The success of the product led the U.S. government to invite the Dreyfus brothers to build a plant in the United States, which started commercial production in 1924.

The successful manufacture of these two fibers, although based on fibrous materials available in nature, marked the beginning of the development of manufactured fibers in the 1930s. This effort, initiated by a technological breakthrough, was marked by the work of W. H. Carothers, aimed at learning how and why certain molecules joined to form large molecules or polymers [2]. Fibers were described as being composed of high-molecular-weight linear polymers, and the first one to be manufactured, nylon 66, was synthesized and produced on a commercial scale in 1939. It was quickly followed by nylon 6, the second most widely used nylon, and modacrylic (1949), olefin (1949), acrylic (1950), polyester (1953), and triacetate (1954). Glass had joined this group of large-production items earlier in 1936.

Several other fibrous materials have been produced, but they are regarded as fibers with special performance characteristics, used either in limited textile or in specialized industrial applications. Some of these, worth noting, are spandex (1959), aramid (1961), polybenzimidazole (PBI) (1983), and sulfar (1983).

Thus, the period from the 1930s to the 1960s can be considered as a time of discoveries and innovations for manufactured fibers, when the majority of the basic fibers were developed. The years since then may be thought of as a period of modification of performance characteristics. The basic generic materials have been manipulated both chemically and physically to produce a wide variety of different fibers, tailored to secure the desired characteristics for specific end-use products. Thus, fibers can be extruded in different shapes and sizes for special purposes. They can be modified to offer greater comfort, flame resistance, or static-free behavior in apparel; they can offer soil release and other desirable characteristics for carpets; they can be developed with unique surface characteristics, easier dyeability, or better blending qualities. The industry has begun to discover many possibilities for modifying the behavior of a given fibrous material. It has learned how to produce new fibers with greater strength, greater thermal resistance, or other special qualities. In particular worth mentioning are the biodegradable polymers, and the fibers based on them, that had controlled absorbability and entered the market later in the twentieth century. These materials, limited to specialized applications and therefore produced in low volume, were developed primarily for use in medical textile products.

Essentially, then, no new, large-volume, highly profitable fibers have been developed since the mid-1950s. Instead, the existing ones have become commodities with all the economic impact thereby implied. No major chemical engineering processes have been added, although the previously described ones have been modified to allow for spinning of liquid crystalline polymers or the formation of gel-spun fibers. Research activity has been reduced and centered essentially on modifications of fiber size, shape, and properties, and many variants now are successfully marketed. Production volumes have increased enormously for nylon, polyester, and polyolefin.

Fiber Consumption

Figure 12.3 compares population growth with the production of manufactured fibers and the mill consumption of natural fibers in the United States. Per capita consumption of all fibers, starting at a level in the 1920s of about 30 lb, rose to approximately 40 lb following World War II and reached a level at or about 50 lb in the 1970s. It topped 60 lb in 1973, dropped below 45 lb in 1982, and in 1991 was at 55 lb. But clearly overshadowing the increases resulting from population growth and a higher standard of living are the volumes produced of, first, the cellulosic manufactured fibers and, second, the

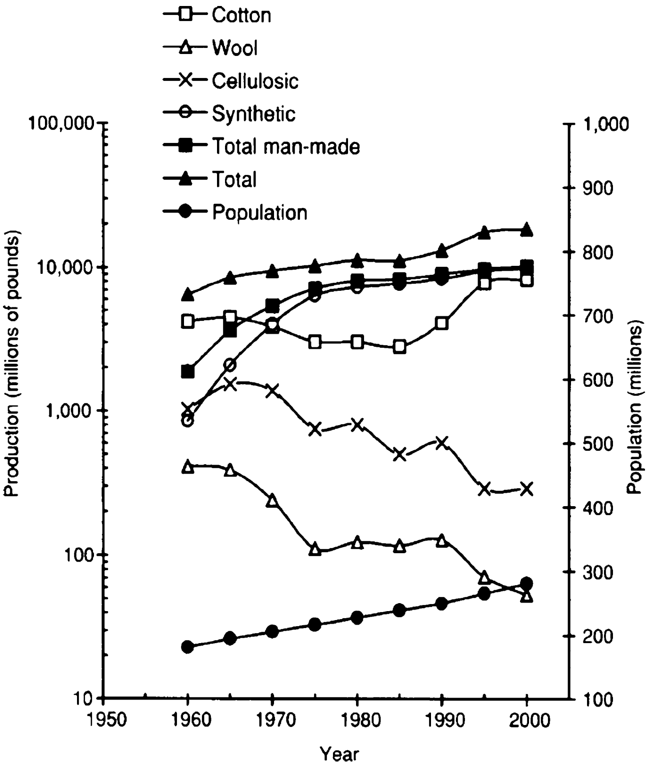


Fig. 12.3 Comparative growth of population and fiber consumption in the United States, 1960–2000

noncellulosic or completely synthetic fibers. The consumption of manufactured fibers increased from about two billion lb in 1960 to nearly nine billion lb in 1990. During the 1980s, the increase was about 11%, made up of a decrease in cellulosics of about 200 million lb (25%) and an increase in synthetics of about 1.2 billion lb (17%). Of the 13.2 billion lb of fibers used in the United States in 1990, 4.6% was the manufactured cellulosics, 63.3% was the synthetics (67.9% manufactured), 31.1% was cotton, and less than 1% was wool. A significant increase was noted in the use of cotton during the 1980s, from 3.0 billion lb in 1980 to 4.1 billion lb in 1990, believed to be in large measure due to advances in chemical finishes that made caring for the natural fibers easier than it was with the older technology. To some extent, this increase also may have occurred because people were willing to accept a more wrinkled look in order to gain the comfort of hydrophilic fibers. Still, the general figures represent quite a reversal in the consumption of fiber types seen three or more decades ago. For example, in 1960, of the 6.5 billion lb of fibers used, manufactured fibers accounted for only 29% of the total, with cotton 65% and wool 6%.

Major applications of fibers lie in apparel, home furnishing, and industrial products. In each of these, manufactured fibers have made large inroads, and currently, their usage dominates. As an illustration, consider the changes that have taken place in the use of the materials required in the manufacture of tire cords. Originally made from cotton, rayon took a commanding position during World War II. But as late as 1951, cotton comprised about 40% of the total output of tire cords of approximately half a billion lb, and nylon was at a negligible level of four million lb. By 1960, however, cotton had all but disappeared; nylon represented about 37% of the total (on a weight basis), even though only about 0.8 lb of nylon is needed to replace 1.0 lb of rayon. Whereas rayon for several years had dominated the so-called original-equipment tire market and nylon had held a corresponding position for replacement tires, more recently, glass and polyester have made heavy inroads into both—especially in belted constructions. The situation continued to change in favor of noncellulosic manufactured fiber usage in tires so that by 1972, rayon was down to 14%, nylon up to 42%, polyester up to 32%, glass up to 7%, and steel at 5%, all on a weight basis. By the late 1970s, tire markets were dominated wholly by manufactured fibers with polyester holding over 90% of the passenger car original-equipment market and nylon commanding over 90% of the truck original-equipment market. This division of markets is a direct result of the performance characteristics of the two fibers. Polyester-containing tires are free of "flat-spotting" or cold-morning thump and so are preferred in passenger cars for their smooth ride. On the other hand, nylon-containing tires are tougher and more durable and so are the choice for trucks and off-road vehicles.

The production of manufactured fibers throughout the world has developed in a manner that rather parallels the situation in the United States, as may be seen in Fig. 12.4. There are some expected differences, and obviously, the data for world usage are strongly influenced by the large components attributable to the United States, which currently account for about 23% of the manufactured fiber and about 15% of the total fiber consumption. The output of the world cellulosics has leveled off, but expansion of the noncellulosics has continued unabated. The use, or at least the recorded use, of the natural fibers, cotton and wool, rose rapidly in the 1950s, as the world economy rebounded at the conclusion of World War II. Since then (1960–1970), a modest increase has continued, essentially parallel to the growth of world population. But, in comparison with population trends, it appears that the great demand has been for manufactured textile fibers. Much of this increase has resulted from an improved standard of living and the absence of major wars.

A detailed economic examination of the processing of fibers and the changes that have taken place during the last half century would show two rather vivid occurrences. The first of these is a rapid decrease in the prices of the newer fibers as they became established, followed by a leveling out and stabilization. The second is the relative stability of prices of the manufactured fibers on short-term and even

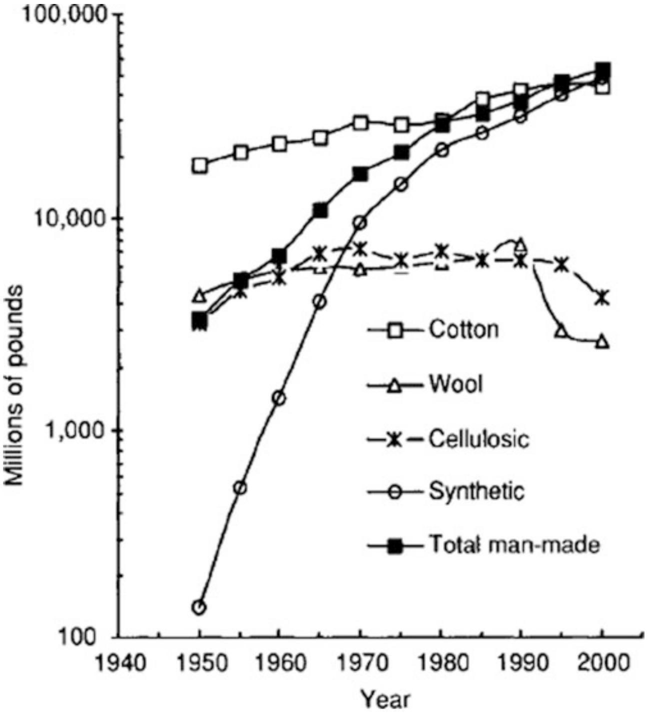


Fig. 12.4 Production of fibers in the world, 1950–2000

long-term bases, as compared with fluctuations in the prices for the natural fibers where governmentally imposed stability has not been in effect. Data are not presented about it in this text, but in the first half of the twentieth century, there was a saying in the textile industry that the person who made or lost money for the company was the one who was responsible for buying cotton and wool "futures." However, it should also be emphasized that list prices of manufactured fibers are ceiling prices and do not reflect the short-term discounts, allowances, and special arrangements that are given in a free marketplace when the demand for any manufactured fiber softens.

A presentation of complete information about the consumption of raw materials, chemical reactions, reagents and catalysts used, and efficiencies of operation in the production of manufactured fibers undoubtedly would contribute to a better understanding of the industrial chemistry involved. Several factors have prevented this, however. In the first half of the twentieth century, a historical belief in the efficacy of trade secrets still permeated the chemical industry. Even with the increased mobility of technical and scientific personnel during and following World War II, the idea still prevailed that if nothing other than patents was allowed to become public knowledge, so much the better. The situation has changed considerably since about 1960, as can be noted from the availability of information contained in the list of suggested readings that follows this chapter, yet, secrecy tends to be maintained despite the fact that key employees move from company to company, and the chemical engineering knowledge available in chemical companies that produce large volumes of fibers permits an almost complete appraisal of a competitor's activities.

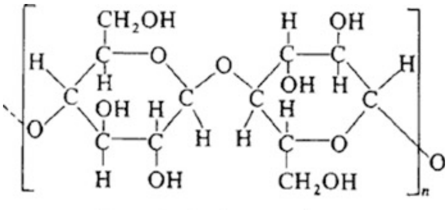
In general, in the early period of production of a fiber, the cost of the original raw material may have had very little bearing on the selling price of the final fiber. The most important factor is the action of the producer's competitors and the conditions of the market and the demand that can be developed. But the complexity of the processes involved in conversion determines the base cost of the fiber at the point of manufacture. As the process becomes older, research reduces this complexity; with simplification, there may be rapid drops in plant cost. If demand remains high, such reductions will not be expected to be reflected in selling prices; rather, profits are high. As more producers enter the field in order to share in those profits, output capacity surpasses demand, and in accordance with classical economic theory, major selling price reductions result. This was happening, in general, in the 1960s and for cellulose-based manufactured fibers in the 1970s. But beginning in 1973, the cost of petroleum-based products started to rise steeply and erratically. This rise was based not on economic considerations alone but on political considerations among the oil-producing and exporting countries (OPEC) as well. Further upward pressure on manufactured fiber prices has resulted from governmental limitations placed on chemical usage and exposure and on amounts of chemicals that can be discharged into the air and water. To meet these limitations, the manufactured fiber industry has had to supply large infusion of capital. In some instances, such expenditures could not be justified, and plant capacity was shut down permanently. This was particularly true in the case of filament rayon. In recent years, the factors of rapidly rising raw material/energy prices and the costs of meeting environmental regulations have not allowed the prices of manufactured fibers to fall as production experience has been gained and technological advances have been introduced. Instead, selling prices have been continually adjusted upward in an effort to pass along unavoidable cost increases so as to maintain profitability. In areas of application in which a manufactured fiber replaced a natural one because of lower prices and stable availability, swings in fashion and increases in imports sometimes have caused a reduction in fiber utilization. This scenario, combined with environmental concerns, is believed to have particularly applied to acrylic fiber, whose production has decreased recently in both Europe and the United States.

The great importance of manufactured fibers in the chemical industry and in the overall economy of the United States (and, in general, the developed countries) becomes apparent when the volume of production of these materials is considered and compared with the market value of even the least expensive of the raw materials used by them. The amounts of oil and natural gas consumed by the manufactured fiber industry represent around 1% of national annual usage. Of this amount, about one half is used to produce raw materials from petrochemicals, with the other half used for energy to convert trees to wood pulp for cellulose-based fibers and to convert the wood pulp and petrochemical-derived raw materials to fibers.

Rayon

Chemical Manufacture

Rayon, the first of the manufactured fibers produced in large volume, is based on the natural polymer cellulose, a repeat unit of which is shown below:



Two anhydroglucose units

Although in the early days the main source of this raw material was cotton linters, a combination of improved technologies for obtaining alpha cellulose from wood and the shortage of cotton linters used for the manufacture of cellulose nitrate during World War II resulted in a shift in raw material to wood pulp. Only certain trees constitute the most economical supply of dissolving pulp, as the final product is called; the process economics depend upon the cost of logs delivered at the pulping mill and the relative yield of alpha cellulose after the unusable lignin, and the other components of the wood are discarded.

A general flow diagram for the manufacture of rayon is given in Fig. 12.5. The dissolving pulp is received by the rayon manufacturer in sheet or roll form. In the manufacturing process, impurities are removed, with special attention being given to removal of traces of such metallic elements as manganese and iron, the former having an effect on the manufacturing process (as will be noted later) and the latter an effect on the color of the final product. The production of dissolving pulp involves drastic chemical action at elevated temperatures, which substantially reduces the originally very high molecular weight of the cellulose. The portion not soluble in 17–18% aqueous caustic, known as alpha cellulose, remains, and the lower-molecular-weight beta and gamma fractions are largely soluble and lost. The composition of the pulp is aimed at high alpha content. A typical economic trade-off is involved. The pulp producers can secure an alpha content of up to 98% by means of a cold caustic extraction, or, on the other hand, the rayon manufacturer can use a less expensive, lower alpha content pulp (90–96%) and expect to secure a lower yield. The sellers have numerous grades available to meet the specific process needs and end-product requirements of each of the buyers.

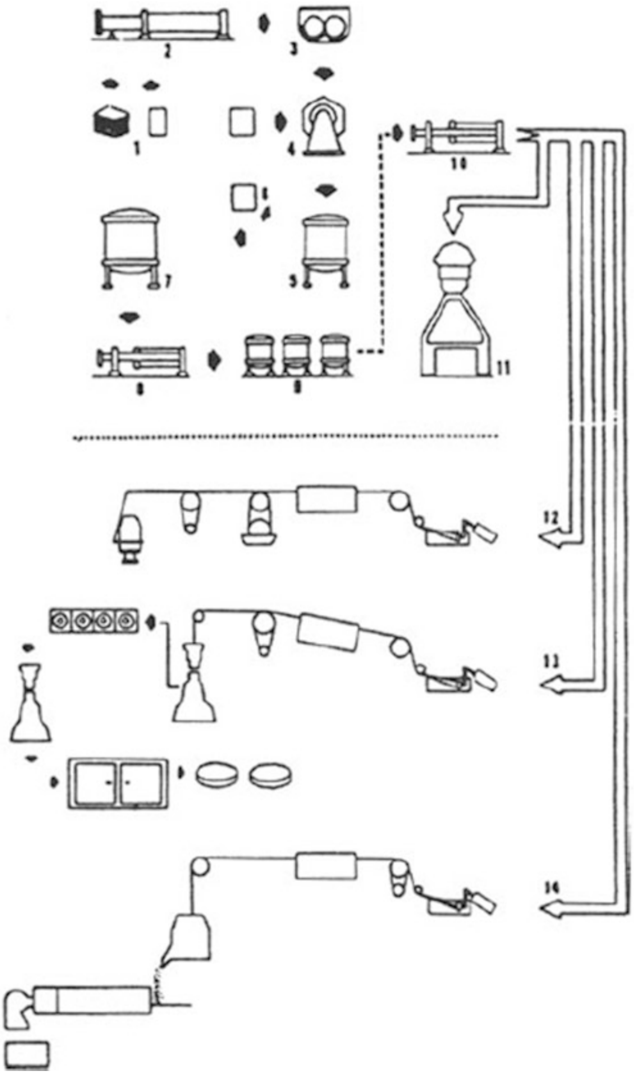


Fig. 12.5 Flow diagram for manufacture of viscose yarn: (1) cellulose sheets and caustic soda, (2) steeping press, (3) shredder, (4) xanthating churn, (5) dissolver, (6) caustic supply, (7) ripener, (8) filtration, (9) deaeration, (10) filtration, (11) continuous process, (12) tire cord, (13) pot spinning, (14) staple spinning

In the manufacture of rayon, it is the usual practice to begin "blending" at the first step, which involves steeping the pulp. Further blending proceeds throughout successive steps. The warehouse supply of pulp consists of numerous shipments, and in making up the batches for the conventional process, a few sheets are taken from each of several shipments. This serves two purposes. It prevents a slight variation in a single pulp lot from unduly affecting any

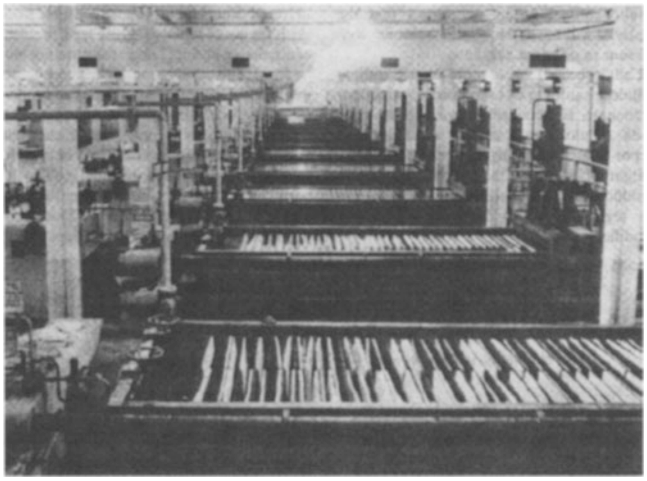


Fig. 12.6 Steeping of cellulose in the manufacture of viscose rayon (Courtesy of Avtex Fibers, Inc.)

given volume of production, and it provides a moving average so that changes with time are reduced to a minimum.

The cellulose sheets are loaded vertically, but loosely, into a combination steeping bath and press (Fig. 12.6), which is slowly filled with a solution of 17–19% caustic, where they remain for about 1 h. In the steeping, the alpha cellulose is converted into alkali or "soda" cellulose; at the same time, as already mentioned, the caustic solution removes most of the beta and gamma celluloses (also called hemicelluloses). The exact chemical composition of the soda cellulose is not known, but there is evidence that one molecule of NaOH is associated with two anhydroglucose units in the polymer chain:

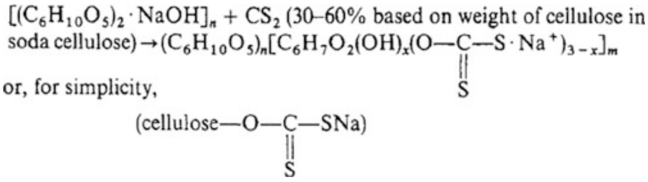
$$\begin{aligned}&(C_6H_{10}O_5)_n\ (\text{cellulose}) + 18\%\ \text{aqueous NaOH}\\&\quad\rightarrow [(C_6H_{10}O_5)_2 \bullet \text{NaOH}]_n\\&\quad\times (\text{swollen, insoluble, soda cellulose I})\\&\quad+ \text{soluble soda cellulose from } \beta \text{ and } \gamma \text{ celluloses}\end{aligned}$$

The excess caustic solution is drained off for reuse. Additional amounts are removed by forcing the sheets through a press. The sheets are still in a swollen state and retain from 2.7 to 3.0 parts of the alkali solution. The spent steeping solution squeezed out of the pulp is processed for recovering the caustic from the organic materials.

The sheets of soda cellulose are discharged into a shredder. If blending is desired, the charges from two or more steeping presses are mixed in a single shredder, where the already soft sheets are torn into crumbs; cooling is provided to prevent thermal degradation. Shredding is controlled to produce crumbs that are open and fluffy, and that will allow air to penetrate the mass readily; this is essential in aging.

Soda cellulose is aged by holding it at a constant temperature in perforated containers. The oxygen in the air produces uniform aging accompanied by a reduction in molecular weight and an increase in the number of carboxyl groups present. The target of aging is an average molecular weight high enough to produce satisfactory strength in the final fiber but low enough so that the viscosity of the solution will not be excessively high at the desired concentration for spinning. Each of the various rayon end products has its optimum degree of polymerization or chain length, ranging from about one fourth the original length for regular rayon to one half for certain high-performance fibers. As noted earlier, this optimum size is generally established by effecting a compromise between process economics and desired end-product properties. The aging proceeds for periods of up to 2 or 3 days, although the tendency is to speed up the operation by using higher temperatures and traces of metal ions, such as manganese or cobalt, to catalyze the reaction. A combination of experience and constant quality-control testing guarantees that the material will reach the correct point for conversion to cellulose xanthate.

Cellulose xanthate, or more exactly, sodium cellulose xanthate or sodium cellulose dithiocarbonate, is obtained by mixing the aged soda cellulose with carbon disulfide in a vapor-tight xanthating churn. Based upon weight of cellulose, the amount of carbon disulfide used will be in the range of 30% for regular rayon to 50–60% for modified varieties.



$$[(C_6H_{10}O_5)_2 \cdot \text{NaOH}]_n + CS_2\ (30\text{–}60\%\ \text{based on weight of cellulose in soda cellulose}) \rightarrow (C_6H_{10}O_5)_n[C_6H_7O_2(OH)_x(O\text{—}\underset{\underset{S}{\|}}{C}\text{—}S \cdot Na^+)_{3-x}]_m$$

or, for simplicity,

$$(\text{cellulose—O—}\underset{\underset{S}{\|}}{C}\text{—SNa})$$

The xanthate is soluble in a dilute solution of sodium hydroxide—a characteristic discovered by Cross and Bevan in 1892—and this property makes the spinning of rayon possible. It is a yellow solid; when dissolved in a dilute solution of alkali, it becomes a viscous, honey-colored liquid—hence the word "viscose." At this stage, the viscous solution may contain about 7.25% cellulose as xanthate in about a 6.5% solution of sodium hydroxide, although concentrations of both vary, depending on what end products are desired. The solution is ready for mixing with other batches to promote uniformity, to be followed by filtration, ripening, deaeration, and spinning. The filtration process usually involves several stages so that filters of decreasing pore size may be used to secure a balance of throughput and stepwise particle and gel removal.

Such an operation is a straightforward one for "bright" rayon, but only in the days of "artificial silk" did the shiny fiber alone satisfy the market. After a few years, a dull-appearing fiber also was demanded. At first, fine droplets of oil in the filaments were used to produce dullness until it was discovered that titanium dioxide pigment having a particle size smaller than 1 μm in diameter was even more satisfactory. The latter has since become the universal delustrant for all manufactured fibers. With the use of pigments of any type, problems of dispersion and agglomerate formation must be faced. The usual practice has been to add this pigment when mixing the cellulose xanthate into the dilute solution of caustic.

However, there are many other chemicals and additives that a producer may be required to add to the solution, including (1) a few parts per million of a tracer element for later identification of the product, (2) coloring pigments for “dope-dyed” rayon (“dope dyeing” will be discussed in greater detail under another heading); (3) chemicals for controlling the rate of precipitation and regeneration for obtaining rayon with so-called high performance), and (4) polymers and chemicals to impart specific properties to the fiber. From the standpoint of chemical processing, it is obvious that these additives may also be added when the sodium cellulose xanthate is dissolved in dilute caustic solution or may be injected into the solution before it enters the spinnerette prior to being extruded. To keep the operations as flexible as possible, the additives should be injected at the last possible moment so that when a changeover is desired, there will be a minimum amount of equipment to be cleaned. On the other hand, the farther along in the operation that additives are placed into the stream, the greater the problem of obtaining uniformity in an extremely viscous medium and the greater the difficulty in maintaining exact control of proportions before the viscous solution is passed forward and spun. Furthermore, all insoluble additives must be of extremely small particle size, and all injected slurries must be freed of agglomerates by prefiltration; if not, the viscous solution containing the additives must be filtered. Each manufacturer of viscose rayon develops the particular conditions for making additions, depending on a multitude of factors, not the least of which is the existing investment in equipment. All manufacturers must face the universal necessity of filtering the solution with or without pigments or other additives, so that all impurities and agglomerates that might block the tiny holes in the spinnerette are removed.

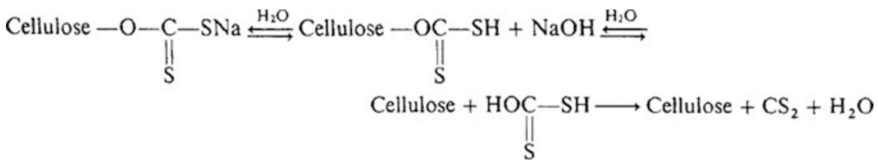
Although it was known in the years following the discovery by Cross and Bevan that a viscose type of solution could be used in the preparation of regenerated cellulose, the conversion of this solution into useful fibers was not possible until the discovery that the solution required aging until “ripe.” Ripening is the first part of the actual chemical decomposition of cellulose xanthate, which, if allowed to proceed unhampered, would result in gelation of the viscose solution.

Experience has taught the manufacturer the correct time and conditions for the aging operation, but the requirement of aging itself demands that the entire process be so planned that the viscose solution will arrive at the spinnerette possessing, as nearly as possible, the optimum degree of ripeness, to produce fibers having the desired characteristics. This degree of ripeness is determined by an empirical test made periodically, which is a measurement of the resistance of the solution to precipitation of the soda cellulose when a salt solution is titrated into it. Thus, it is known as the “salt index” or “Hottenroth number” after its originator. An additional step in the overall ripening operation involves the removal of dissolved and mechanically held air by the use of a vacuum on a moving thin film of the viscose solution.

It should be mentioned that so inevitable is decomposition of cellulose xanthate and consequent gelation of the contents of pipes and tanks that all viscose rayon plants must be prepared to pump in-process viscose solutions to other spinning machines or to a waste receiver, purging the entire system with dilute caustic solution, in the event of a long delay in spinning.

Wet Spinning

Spinning a viscose solution into rayon fibers (wet spinning) is the oldest of the three common ways of making manufactured fibers. In this method, the polymer is dissolved in an appropriate solvent, and this solution is forced through fine holes in the face of the spinnerette, which is submerged in a bath of such composition that the polymer precipitates. The pressure necessary for this extrusion is supplied by a gear pump, which also acts as a metering device; the solution is moved through a final or “candle” filter before it emerges from the holes of the spinnerette. There is immediate contact between these tiny streams and the liquid or “wet” bath. As the bath solution makes contact with the material extruded from the holes, chemical or physical changes take place. These changes, whether of lesser or greater complexity, convert the solution of high-molecular-weight linear polymer first to a gel structure and then to a fiber. As will be observed in what follows, it is an interesting fact that the spinning of viscose rayon, with all of the ramifications made possible by variations in the composition of the solution and the precipitating bath, as well as in the operating conditions, presents the chemist and the chemical engineer with both the oldest and the most complex wet-spinning process.

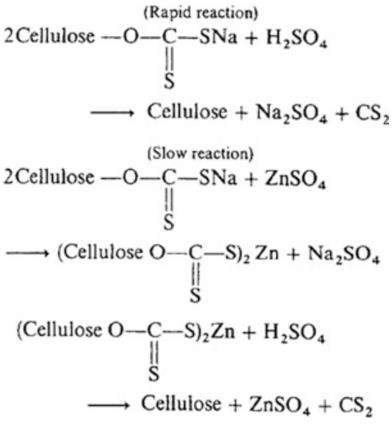


$$\text{Cellulose—O—}\underset{\underset{\displaystyle S}{\|}}{\text{C}}\text{—SNa} \overset{H_2O}{\rightleftharpoons} \text{Cellulose—O}\underset{\underset{\displaystyle S}{\|}}{\text{C}}\text{—SH} + \text{NaOH} \overset{H_2O}{\rightleftharpoons}$$

$$\text{Cellulose} + \text{HO}\underset{\underset{\displaystyle S}{\|}}{\text{C}}\text{—SH} \longrightarrow \text{Cellulose} + CS_2 + H_2O$$

The formation of rayon fibers from viscose solution is far from being simple, from either a physical or a chemical standpoint. The spinning bath usually contains 1–5% zinc sulfate and 7–10% sulfuric acid, as well as a surface-active agent, without which minute deposits will form around the holes in the spinnerette. Sodium sulfate (15–22%) is present, formed by the reactions, and as sulfuric acid is depleted and sodium sulfate concentration builds up, an appropriate replenishment of the acid is required. There is a coagulation of the organic material as the sulfuric acid in the spinning bath neutralizes the sodium hydroxide in the viscose solution; at the same time, chemical decomposition of the sodium cellulose xanthate takes place to regenerate the cellulose. If zinc ions are present, which is the usual situation in the production of the improved types of rayon, an interchange takes place so that the zinc cellulose xanthate becomes an intermediate. It reacts at a slower rate, causing slower decomposition to cellulose. This provides conditions for more effective stretching or drawing of the fiber. Chemical additives usually are present to repress hydrogen ion action. The gel-like structure, the first state through which the material passes, is not capable of supporting itself outside the spinning bath. As it travels through the bath, however, it quickly becomes transformed into a fiber that can be drawn from the spinning bath and that can support itself in subsequent operation (Fig. 12.7).

The reaction between the bath and the fiber that is forming is paramount in determining the characteristics of the final product; it is for this reason that additives (previously mentioned), as well as zinc ions, may be used to control both the rate of coagulation and regeneration. In this manner, the arrangement of the cellulose molecules may be controlled to produce the conformational structure desired. A practical application of this will be discussed later.



$$\text{(Rapid reaction)}$$
$$2\text{Cellulose}-\text{O}-\underset{\underset{\displaystyle \text{S}}{\|}}{\text{C}}-\text{SNa} + \text{H}_2\text{SO}_4 \longrightarrow \text{Cellulose} + \text{Na}_2\text{SO}_4 + \text{CS}_2$$
$$\text{(Slow reaction)}$$
$$2\text{Cellulose}-\text{O}-\underset{\underset{\displaystyle \text{S}}{\|}}{\text{C}}-\text{SNa} + \text{ZnSO}_4 \longrightarrow (\text{Cellulose O}-\underset{\underset{\displaystyle \text{S}}{\|}}{\text{C}}-\text{S})_2\,\text{Zn} + \text{Na}_2\text{SO}_4$$
$$(\text{Cellulose O}-\underset{\underset{\displaystyle \text{S}}{\|}}{\text{C}}-\text{S})_2\text{Zn} + \text{H}_2\text{SO}_4 \longrightarrow \text{Cellulose} + \text{ZnSO}_4 + \text{CS}_2$$

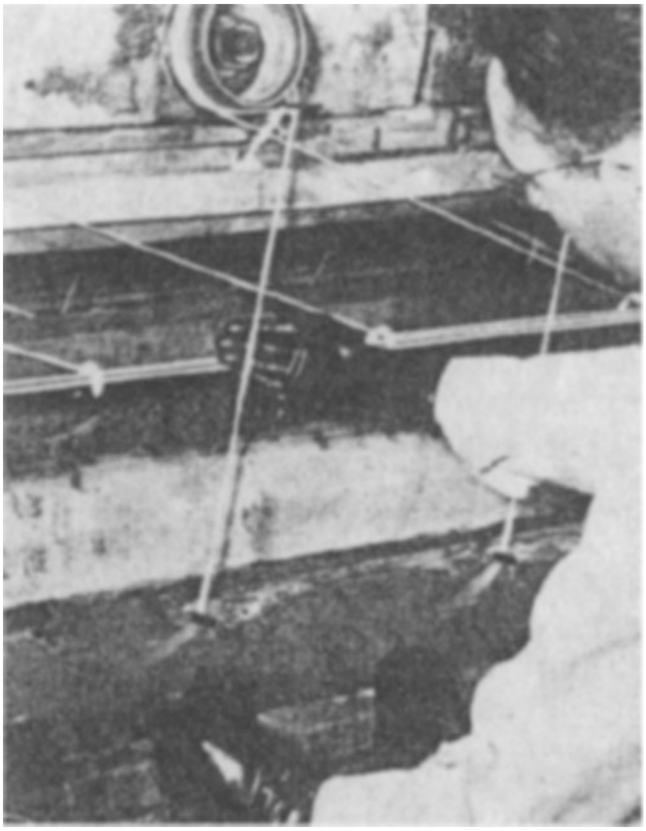


Fig. 12.7 Spinning of viscose rayon (Courtesy of Avtex Fibers, Inc.)

Because of hydraulic drag, stretching occurs in the bath and also in a separate step after the yarn leaves the bath. In both cases, the linear molecules of cellulose are oriented from random positions to positions more parallel to the fiber axis. If a rayon tire cord is to be the final product, the fibers must be severely stretched to produce a very high orientation of the molecule. This is the basis of the tire cord's high strength and ability to resist stretching, without which growth of the tire body would occur. For regular textile and nonwoven uses, such high strengths are not desired, and the spinning and stretching conditions are controlled to produce rayon of lower strength and greater stretchability under stress.

In order to stretch the yarn uniformly during the manufacturing process, two sets of paired rollers or "godets" are employed, each of the two sets operating at different rotational speeds. The yarn is passed around the first set of godets several times to prevent slippage and is supplied to the stretching area at a constant speed. A second set of godets moves it forward at a more rapid rate, also without slippage. Stretching may range from a few to 100 or more percent. Spinning speeds are of the order of 100 m/min but may vary with both the size of the yarn and the process used.

Spinning conditions, composition of the spinning bath, and additives to the viscose solution determine the physical characteristics of the rayon—its breaking strength and elongation, modulus, ability to resist swelling in water, and characteristics in the wet state as compared with those of the dry material. Not only must the chemical composition of the spinning bath be carefully controlled but the temperature must be regulated at a selected point, somewhere in the range of 35–60°C, to ensure those precipitation and regeneration conditions that are essential to the manufacture of any particular viscose rayon having the properties needed for a selected end use.

After precipitation and regeneration of cellulose have been completed and raw rayon fiber has been formed, the subsequent steps must be controlled so that differences in treatment are minimized; otherwise, such sensitive properties as “dye acceptance” will be affected, and the appearance of the final product will vary.

Minute traces of suspended sulfur resulting from the chemical decomposition of cellulose xanthate must be removed by washing with a solution of sodium sulfide. It is expedient to bleach the newly formed fibers with hypochlorite to improve their whiteness; an “antichlor” follows. The chemicals originally present and those used to purify the fibers must be removed by washing. As a final step, a small amount of lubricant is placed on the filaments to reduce friction and improve processibility in subsequent operations.

Several different processes are used for the steps involved in spinning and purifying continuous-filament rayon. One of the most common involves the formation of packages of yarn, each weighing several pounds, for separate treatment. After it has been passed upward out of the spinning bath and stretched to the desired degree, the yarn is fed downward vertically into a rapidly rotating can-like container called a spinning pot or “Topham” box (after the man who invented it in 1900). It is thrown outward to the wall of the pot by centrifugal force and gradually builds up like a cake, with excess water being removed by the same centrifugal force. This cake is firm, although it must be handled with care, and is sufficiently permeable to aqueous solution to permit purification.

In another method of package spinning, the yarn is wound onto a mandrel from the side at a uniform peripheral speed. With this process, the yarn may be purified and dried in the package thus formed. In any of these systems, the spinning and stretching, as well as subsequent steps, may involve separate baths.

The continuous process for spinning and purifying textile-grade rayon yarn merits particular mention from the standpoint of industrial chemistry, as it is rather an axiom that a continuous process is to be preferred over a batch or discontinuous operation. This method employs “advancing rolls” or godets that make it possible for the yarn to dwell for a sufficient length of time on each pair, thus allowing the several chemical operations to take place in a relatively small area. Their operation depends on the geometry existing when the shafts of a pair of adjacent cylindrical rolls are oriented slightly askew. Yarn led onto the end of one of these and then around the pair will progress toward the other end of the set with every pass, the rate of traversing, and therefore the number of wraps, being determined by the degree of skewness.

The production of rayon to be converted to staple fiber also is amenable to line operation. Here, the spinnerette has many thousands of holes, and a correspondingly large number of filaments are formed in the precipitation bath. The resulting tow then is stretched to the desired degree and immediately cut in the wet and unpurified condition. The mass of short lengths can be conveyed through the usual chemical treatments, after which it is washed and dried. It is fluffed to prevent matting and is packaged for shipment in large cases.

Cuprammonium, Nitrocellulose, and Cellulose Acetate Processes for Rayon

Cuprammonium Cellulose. Cellulose forms a soluble complex with copper salts and ammonia. Thus, when cellulose is added to an ammoniacal solution of copper sulfate that also contains sodium hydroxide, it dissolves to form a viscous blue solution, and in this form, it is known as cuprammonium cellulose. The principles on which the chemical and spinning steps of this process are based are the same as those for the viscose process. Cellulose is dissolved, in this case, in a solution containing ammonia, copper sulfate, and sodium hydroxide. Unlike the viscose solution, the cuprammonium solution need not be aged and will not precipitate spontaneously on standing except after long periods. It is, however, sensitive to light and oxygen. It is spun into water and given an acid wash to remove the last traces of ammonia and copper ions. Although this rayon was never manufactured in a volume even approaching that achieved by the viscose process, the smaller individual filaments inherent in it made it useful in certain specialty markets. It no longer is manufactured in the United States but continues to be made abroad.

Nitrocellulose and Cellulose Acetate. Although nitrocellulose and cellulose acetate intermediates have been made and regenerated to form cellulose fibers, neither of these historical processes associated with rayon manufacture are still in operation.

Textile Operations

After the filament rayon fiber has been spun and chemically purified, much of it passes through what are known as “textile operations” before it is ready to be knitted or woven. Because these steps of twisting and packaging or beaming are common to the manufacture of all manufactured fibers, it is advisable to review briefly the background and the processes.

Rayon, the first manufactured fiber, not only had to compete in an established field but also had to break into a conservative industry. Silk was the only continuous-filament yarn, and products made from it were expensive and possessed of high prestige so that they offered a tempting market for rayon. Thus, the new product entered as a competitor to silk and, as already noted, became known as "artificial silk." Under the circumstances, it was necessary for rayon to adapt itself to the then-existing silk-processing operations and technologies. It was customary to twist several silk filaments together to secure a yarn of the desired size, strength, and abrasion resistance. Because rayon was weaker than silk and its individual filaments were smaller, it required as much twisting as silk or even more.

This twisting could have been carried out in the same plant where the yarn was spun, but the existence of silk "throwsters" (from the Anglo-Saxon thrawan, to twist or revolve) made that unnecessary. However, as the rayon industry developed, the amount of yarn twisted in the producing plant or sent forward to throwsters decreased. Over the years, the trend has been to use less twist and to place, instead, several thousand parallel ends directly on a "beam" to form packages weighing as much as 300–400 lb, which are shipped directly to a weaving or knitting mill. The advent of stronger rayons, as well as other strong fibers, and a diminishing market for crepe fabrics which required highly twisted yarns accelerated the trend away from twisting.

In all twisting and packaging operations, the yarn makes contact with guide surfaces and tensioning devices, often at very high speeds. To reduce friction, it is necessary to add a lubricant as a protective coating for the filaments. This is generally true of all manufactured fibers, and it is customary to apply the lubricant or "spinning finish" or "spinning lubricant" as early in the manufacturing process as possible. For those materials that develop static charges in passing over surfaces, this lubricant also must provide antistatic characteristics.

It is difficult to overstate the importance of fiber lubricants to the successful utilization of manufactured fibers. Few problems can be more damaging to a fiber-handling operation than a lubricant upset. A separate chapter could be written on lubricant usage, but some of its more important aspects will be mentioned here. Obviously, lubricants must reduce friction between the fiber and various surfaces to allow movement without excessive damage to the fiber or the surface contacted, the latter being any one of a variety of metals or ceramics. Similarly, a fiber comes in contact with the surface of other fibers in staple fiber processing and in packaging. The lubricant composition must be stable under a variety of storage conditions, without decomposing or migrating within the package or being lost from the fiber surface by adsorption into the fiber, and must be nontoxic and nondermatitic. It also must be compatible with other materials added during textile processes, such as the protective size coat applied to warp yarns before weaving or the wax coat often applied to yarns before circular knitting. Possible metal corrosion must be evaluated for each lubricant composition. Finally, after having performed its function, the fiber finish or lubricant must be completely scoured from the fabric to permit uniform adsorption of dyes and fabric finishes. The application of sewing lubricants to fabric to be cut and sewn is yet another area requiring attention.

Spun Yarn. After rayon became established in the textile industry, where it could be used as a silklike fiber, and its selling price was greatly reduced, other markets for it were developed. The cotton, wool, worsted, and linen systems of converting short discontinuous fibers to yarns were well established, and their products were universally accepted. Here again, it was necessary to make rayon fit the requirements of existing equipment and historically acceptable operations. The first of these was that it be cut into the same lengths as those found in cotton and wool. Fortunately, the viscose rayon process was and is eminently suited to the production of tows containing thousands of filaments. The pressure required to force the solution through the holes is so low that neither thick metal sections nor reinforcement of the surface is necessary to prevent bulging, and large spinnerettes containing several thousand holes can be used. Furthermore, the spinning bath succeeds in making contact with all the filaments uniformly. As a result, the spinning of viscose rayon tow is very similar in principle to the production of the smaller continuous-filament yarns.

Because both cotton and wool possess distortions from a straight rodlike structure, machinery for their processing was designed to operate best with such crimped fibers. Thus, it was necessary for rayon staple to possess similar lengths and crimpiness in order to be adapted to existing equipment. The crimp, that is, several distortions from a straight path per inch, is produced in rayon "chemically" by modification of the structure. The precipitation-stretching step in spinning is carried out so that the skin and the core of the individual filaments are radially nonuniform and constantly changing over very short lengths along the filaments. Because the skin and the core differ in sensitivity to moisture, the two components shrink differentially, leading to the development of permanent distortions of the filaments. The latent chemical crimp may be enhanced by a thermomechanical step. In this, the tow is fed between two wheels, which in turn force it into a chamber, called a "stuffer box," heated with steam, in which it is forced against the compacted material ahead of it, causing the straight filaments to collapse immediately. As the mass of material is pressed forward, it becomes tightly compacted, and it tends to remain axially distorted after it escapes through a pressure-loaded door at the opposite end of the chamber.

Modified Viscose Rayon Fibers

The variations of chemical ingredients, their concentrations, and the temperature of the spinning bath determine the rates of coagulation and regeneration and thus the relative amounts of "skin" and "core" in the cross section of the fiber. The skin is known to possess a higher degree of order and better mechanical properties than the core, so an increase in its proportion is desired in higher-performance fibers. The degree of orientation is determined by the stretch imparted. Inherent in-process variations such as these are what may be called the "art" of viscose rayon manufacturing, whereby great diversity in rayon properties can be obtained. The cross-sectional morphologies of some rayons are illustrated in Fig. 12.8.

High-Wet-Modulus Rayon. One of the important innovations in the rayon industry has been the development of high-wet-modulus (HWM) rayon. For its manufacture, the cellulose molecules require a higher degree of polymerization (DP) than regular rayon, so aging and ripening times are decreased, and the processing temperature is lowered. As the viscose solution flows through the spinnerette into the bath, coagulation takes place to form the needed skin. However, in order for an increase in tenacity to occur, regeneration should proceed slowly; this is aided by using a lower concentration of acid in the spin bath. Also, zinc is added because zinc ions in the bath slow down regeneration by forming zinc cellulose xanthate, which is more resistant to acid decomposition than is sodium xanthate. Because of the slow generation process, what is actually formed is nearly an "all-skin" rayon with a round cross section. Because the structure also is stretched before it crystallizes, a higher stretch is possible; this gives a higher orientation. The fiber thus has significantly higher strength than regular rayon. This higher tenacity exists not only when the material is dry but also when it is wet, hence the name HWM rayon. Because of its high structural orientation and greater order, the fibers have fewer physically accessible sites for water molecules; thus, it is less susceptible to swelling and to the adverse effects of basic cleaning solutions so that the fiber's launderability is improved. The HWM rayons are used extensively for blending with cotton, wool, silk, and all other manufactured fibers.

High Absorbency Rayons. Over the past years, disposable products have become commonplace, especially in the United States and Europe. Cellulosic fibers, particularly rayons, have served the needs of the disposables industry because of their absorbent qualities. The most useful fibers for disposable/absorbent applications are the rayons with crenulations, crimp, and hollow regions, all of which add to the absorbency of the fiber. These characteristics are achieved in varying degrees by physical and chemical alterations in the spinning process. Crenulations or random irregularities in the shape of the cross section, typical for most rayon fibers, are caused by the rapid formation of skin before the dehydration is complete. As the fiber interior loses solvent, it collapses in certain areas and produces the crenulated shape. Furthermore, fabricators have learned how to control the cross-sectional shape of filaments by using spinnerettes containing other than round holes. One example is fibers having Y or trilobal-shaped cross sections, which have been found to be capable of picking up more water and at a faster rate than possible with fibers of round cross sections [4].

Hollow viscose fibers contain gas pockets produced by adding "blowing" agents, such as sodium carbonate, to the viscose. When carbon dioxide is released during regeneration, the fibers inflate, leading to the formation of hollow filaments. The added free volume and decrease in molecular order increase the ability of the fibers to pick up water.

Other New Developments

A number of other developments are taking place in the rayon industry, the target of one being the manufacture of lint—free rayon for use in products such as circuit boards. In another, graphite particles are blended in with the viscose to reduce static buildup. Production of flame-retardant rayon has received increased attention, in one case being achieved by the addition of phosphorus compounds to the spinning dope. The advanced technology now can produce flame-resistant fibers that, when exposed to high temperatures, will not shrink or emit toxic gases. Other developments are in the area of finding better and environmentally safer solvents for the cellulosic raw material. Searches are under way for solvents that may lead to lyotropic liquid crystalline polymer solutions from which ultrahigh-strength and high-modulus fibers can be spun.

Environmentally Friendly High-Wet-Strength Rayon: Lyocell

Low wet strength of rayon in general has restricted the application of the fiber to disposable and semidurable materials. Additionally, the environmental concerns associated with its manufacture have resulted in significant curtailment in production of the fiber during the recent past. Search for ways and means to produce a high-wet-strength fiber using

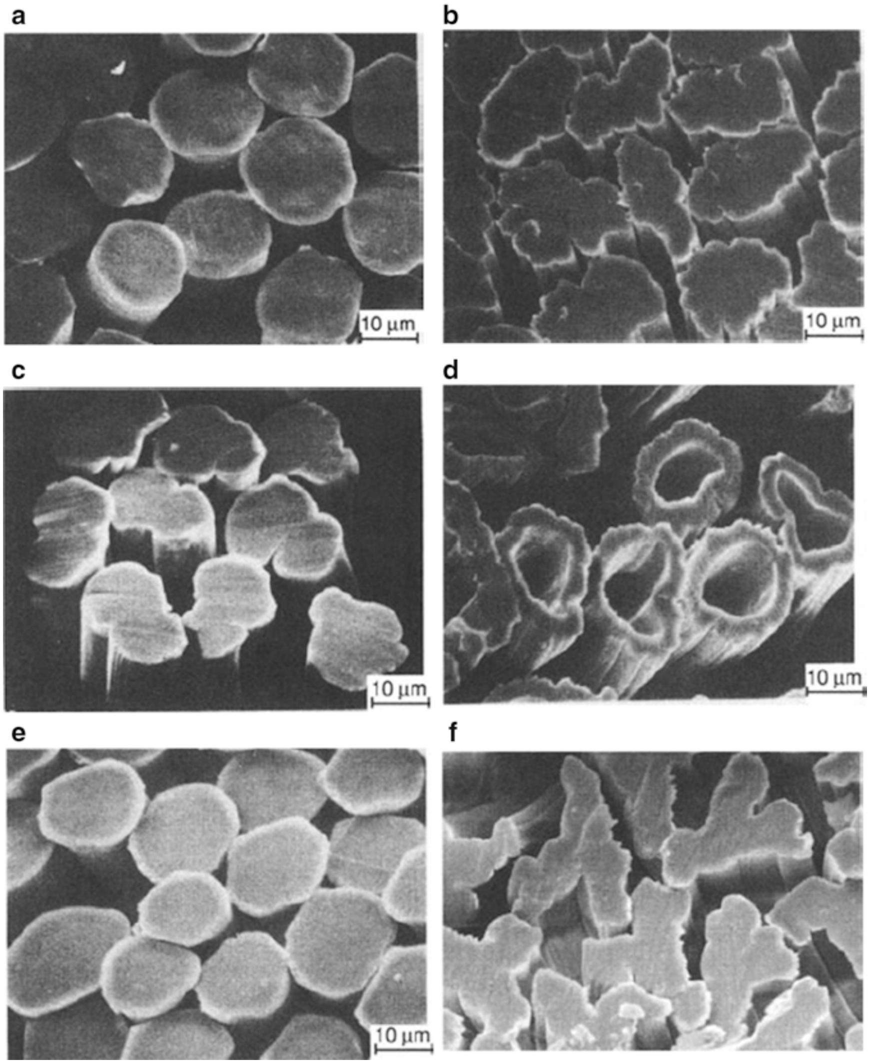


Fig. 12.8 Cross-sectional morphologies of some of the rayon fibers. (**a**) High-wet modulus, (**b**) regular rayon, (**c**) crimped HWM, (**d**) hollow, (**e**) cuprammonium, (**f**) trilobal (*Sources*: All except trilobal: Turbak [3], copyright John Wiley & Sons, Inc., New York, 1985 and used with permission of the copyright owner; trilobal photo: Gupta and Hong [4])

an environmentally acceptable process has occupied much of the research effort during the past two decades. One of the latest additions to the family of rayon fibers is lyocell. The fiber has wet strength comparable to that of the natural cellulosic fibers and is manufactured with a solvent that is essentially totally recovered and recycled [5]. The solvent used is *N*-methyl morphine oxide, $O(C_4H_8)$ $NOCH_3$, popularly known as amine oxide. The manufacturing process involves the dissolving of pulp in hot amine oxide, filtration of the solution, and then spinning into a bath containing a dilute solution of the solvent. The bath removes the amine oxide from the fibers, which are washed and dried, and the removed solvent is almost totally reclaimed for further use. The final fiber is said to have a different molecular structure from that of normal rayon and a smooth surface and a round cross section. The fiber is noted to be stronger than cotton and normal rayon in both the dry and the wet states [6].

Cellulose Acetate

Historical

Cellulose acetate was known as a chemical compound long before its potential use as a plastic or fiber-forming material was recognized. The presence of hydroxyl groups had made it possible to prepare cellulose esters from various organic acids, as cellulose consists of a long molecular chain of beta-anhydroglucose units, each of which carries three hydroxyl groups—one primary, the other two secondary. The formula for cellulose (already noted) is $[C_6H_7O_2(OH)_3]_n$; when this is fully esterified, a triester results. It was learned quite early that although cellulose triacetate is soluble only in chlorinated solvents, a product obtained by partial hydrolysis of the triester to a "secondary" ester (having about 2.35–2.40 acetyl groups per anhydroglucose unit) was easily soluble in acetone obtaining a small amount of water. Many

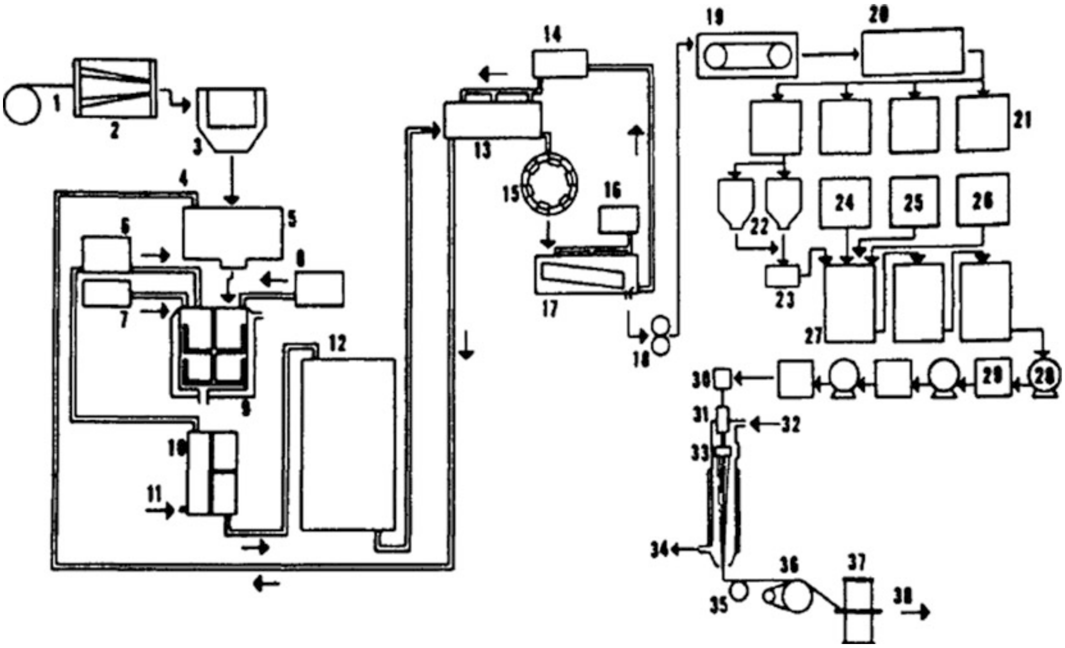


Fig. 12.9 Flow diagram for manufacture of cellulose acetate yarn: (1) wood pulp, (2) attrition mill, (3) cyclone, (4) 35 % acetic acid, (5) pretreater, (6) magnesium acetate solution, (7) precooled acetylation mix, (8) sulfuric acid, (9) acetylator, (10) ripener, (11) steam, (12) blender, (13) precipitator, (14) dilute acetic acid, (15) hammer mill, (16) water, (17) rotary screen washer, (18) squeeze rolls, (19) drying oven, (20) blender, (21) storage bins, (22) silos, (23) weight bins, (24) acetone, (25) wood pulp, (26) pigment, (27) mixers, (28) hold tanks, (29) filter press, (30) pump, (31) filter, (32) air, (33) jet, (34) acetone recovery, (35) oiling wheel, (36) feed roll, (37) bobbin, (38) inspection

other cellulose esters have been prepared, but only the acetate has been commercialized successfully as a manufactured fiber. Propionates and butyrates, and mixed esters of one or both the acetates, have applications as plastics. The first acetate fibers were produced in 1921 in Europe and in 1924 in the United States.

Manufacture of Cellulose Secondary Acetate

Cellulose acetate originally was made from purified cotton linters, but this raw material has been entirely replaced by wood pulp. The other raw materials used are acetic acid and acetic anhydride.

Cellulose acetate is manufactured by a batch process (see Fig. 12.9). There has been mention in the patent literature of a continuous system, but its utilization as a production process has not been announced. The "charge" of cellulose, purified, bleached, and shredded, is of the order of 800–1,500 lb. It is pretreated with about one third its weight of acetic acid and a very necessary amount of water, about 6% of its weight. If it is too dry at the time of use, more H_2O must be added to the acetic acid. A small amount of sulfuric acid may be used to assist in swelling the cellulose and to make it "accessible" to the esterifying mixture.

Although there has been much discussion of the chemistry of cellulose acetylation, it is now generally agreed that the sulfuric acid is not a "catalyst" in the normal sense of the word but rather that it reacts with the cellulose to form a sulfoester. The acetic anhydride is the reactant that provides the acetate groups for esterification. The acetylation mixture consists of the output from the acetic anhydride recovery unit, being about 60% acetic acid and 40% acetic anhydride, in an amount 5–10% above the stoichiometric requirement, to which has been added 10–14% sulfuric acid based on the weight of cellulose used. The reaction is exothermic and requires that the heat be dissipated.

In preparing for acetylation, the liquid reactants are cooled to a point (0°C) where the acetic acid crystallizes, the heat of crystallization being removed by an appropriate cooling system. The slush of acetic acid crystals in the acetic anhydride–sulfuric acid mixture is pumped to the acetylizer, a brine-cooled mixer of heavy construction (see Fig. 12.10).

The pretreated cellulose is dropped in from the pretreating unit located above. The reaction is highly exothermic, and at the start, large amounts of heat are produced. As the temperature of the reaction mixture rises to the melting point of the acetic acid (16.6°C), its large heat of fusion (45.91 cal/g) prevents a dangerous rise in temperature that would degrade the molecular weight of the cellulose chain.

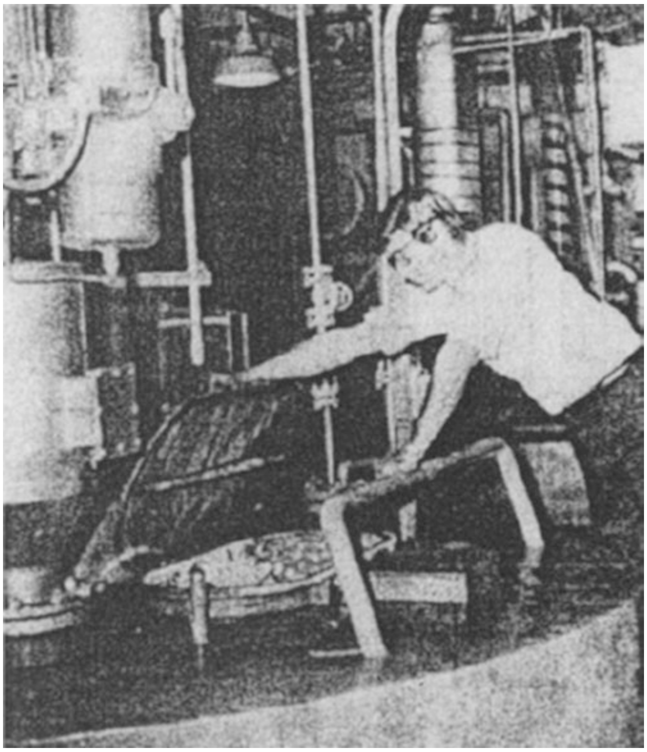


Fig. 12.10 Process vessel for acetylation of cellulose (Courtesy of Celanese Fibers Co.)

As the reaction proceeds, brine in the jacket of the acetylizer provides additional cooling:

$$\begin{aligned}&\text{Cellulose} + (CH_3CO)_2O + H_2SO_4(10-15\%\\ &\text{based on weight of cellulose}) \xrightarrow[\text{[Anhydrous]}]{CH_3COOH}\\ &[C_6H_7O_2(OSO_3H)_{0.2}(CH_3COO)_{2.8}]_n\end{aligned}$$

The reaction product is soluble in the acetylation mixture; as it is formed and dissolved, new surfaces of the cellulose are presented to the reagents. One variation of this procedure uses methylene chloride, rather than an excess of acetic acid in the reaction mixture. This chemical is used both to dissipate the heat by refluxing (boiling point, 41.2°C) and to dissolve the cellulose ester as it is formed. As the reaction proceeds, the temperature is allowed to rise. Because cellulose is a natural product obtained from many sources, it varies slightly in composition, and at the end of the reaction, it's dissolvability cannot be predicted exactly; the disappearance of fibers as determined by microscopic examination thus is the usual means of following its progress.

During the acetylation operation, a certain amount of chain fission is allowed to take place in the cellulose molecule. This is to ensure that the viscosity of the cellulose acetate spinning solution will be low enough for ease of handling but high enough to produce fibers with the required strength. The temperature of the reaction controls the rates of both acetylation and degradation of molecular weight.

The next step in the manufacture is "ripening," whose object is to convert the triester, the "primary" cellulose acetate, to a "secondary" acetate having an average of about 2.35–2.40 acetyl and no sulfo groups (if any sulfuric acid is used in pretreatment) per anhydroglucose unit. While the cellulose sulfoacetate is still in the acetylizer, sufficient water is added to react with the excess anhydride and start the hydrolysis of the ester. Usually, the water is used as a solution of sodium or magnesium acetate, which increases the pH and promotes hydrolysis. The temperature is raised to about 70–80°C by direct injection of steam to speed up the reaction. Hydrolysis is continued until the desired acetyl content is obtained. When this value is reached, an aqueous solution of magnesium or sodium acetate is added to cool the batch and stop the hydrolysis. It is then ready for precipitation. For example,

$$\begin{aligned}&[C_6H_7O_2(OSO_3H)_{0.2}(CH_3COO)_{2.8}]_n\\ &\quad + (CH_3COO)_2Mg \xrightarrow{\text{Aqueous conc. } CH_3COOH}\\ &\quad [C_6H_7O_2(OH)_{0.65}(CH_3COO)_{2.35}]_n + MgSO_4\end{aligned}$$

The solution is carried to the verge of precipitation by adding dilute acetic acid. Then it is flooded with more dilute acetic acid and mixed vigorously, so that the cellulose acetate comes out as a "flake" rather than a gelatinous mass or fine powder. The flake then is washed by standard countercurrent methods to remove the last traces of acid and is dried in a suitable dryer.

Manufacture of Cellulose Triacetate

To obtain completely acetylated cellulose, the reaction requires the use of perchloric acid rather than sulfuric acid as the catalyst. In the presence of 1% perchloric acid, a mixture of acetic acid and acetic anhydride converts a previously "pretreated" cellulose to triacetate without changing the morphology of the fibers. If methylene chloride rather than an excess of acetic acid is present in the acetylation mixture, a solution is obtained. However, usually a degree of substitution between 92% and 100% is acceptable. For obtaining such a triester, it is possible to use about 1% sulfuric acid instead of perchloric acid. When the sulfoacetate obtained from such a reaction is hydrolyzed with the objective of removing only the sulfoester groups, the resulting product has about 2.94 acetyl groups per anhydroglucose unit. The preparation, hydrolysis, precipitation, and washing of "triacetate" are in all other respects similar to the corresponding steps in the manufacture of the more common secondary acetate. Cellulose triacetate,

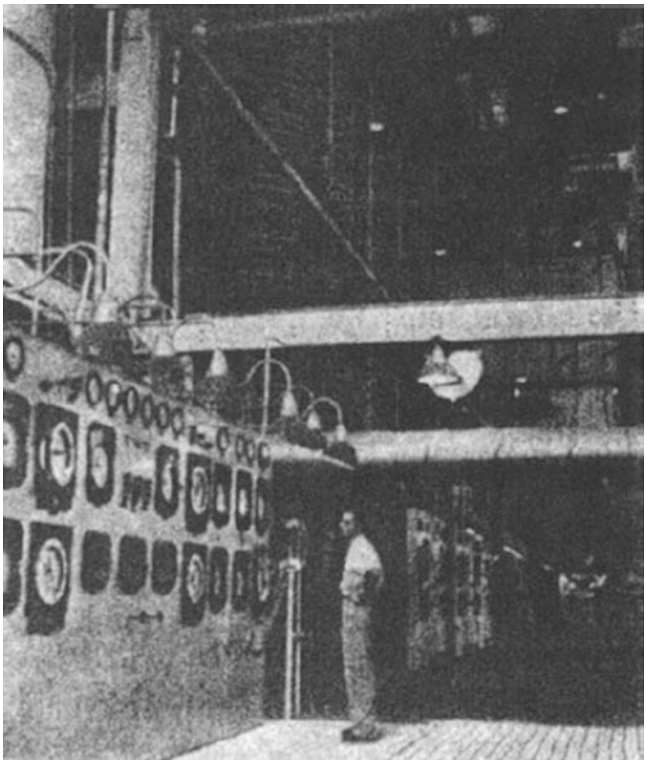


Fig. 12.11 Recovery of acetic acid (Courtesy of Celanese Fibers Co.)

formerly produced under the trade name Arnel® by Celanese Corporation, is no longer in production in the United States.

Acid Recovery. In the manufacture of every pound of cellulose acetate, about 4 lb of acetic acid is produced in 30–35% aqueous solution. The accumulated acid contains a small amount of suspended fibers and some dissolved cellulose esters. To remove the suspended material, the acid is passed slowly through settling tanks. Then it is mixed with organic solvents, so that the acid becomes concentrated in an organic layer, which is decanted. Distillation separates the acid from the organic solvent.

To produce the acetic anhydride, the acid is dehydrated to ketene and reacted with acetic acid using a phosphate catalyst at 500°C or higher in a tubular furnace:

$$CH_3COOH \xrightarrow[\text{Catalyst}]{\text{Heat}} H_2O + CH_2 = C = O$$

$$CH_2 = C = O + CH_3COOH \rightarrow (CH_3CO)_2O$$

The mixture of unreacted acid, water, and anhydride is fed to a still, which yields dilute acetic acid overhead and an anhydride–acetic acid mixture at the bottom (see Fig. 12.11). Conditions are controlled in such a way that the raffinate is about 40% anhydride and 60% acetic acid. As already mentioned, this is the desired ratio for the reaction mixture used for acetylation of cellulose.

Blending of Flake. As in the manufacture of viscose, the products of batch operations are blended to promote uniformity in the manufacture of cellulose acetate. Although a blend of different celluloses is selected in the beginning, the pretreatment, acetylation, and ripening are batch operations with little or no mixing. Before precipitation, a holding tank provides an opportunity for mixing; then precipitation, washing, and drying—all continuous—promote uniformity. The dried cellulose acetate flake moves to holding bins for analysis—the moisture content, acetyl value, and viscosity being especially important. The results of the analyses determine how much further blending is necessary. After blending and mixing of portions of selected batches, the lot is air-conveyed to large storage bins or "silos," which are filled from the center of the top and emptied from the center of the bottom, thus bringing about further mixing.

Spinning Cellulose Acetate

Acetone is metered into a vertical tank equipped with a stirrer, and the cellulose acetate flake and filter aid are weighed in an automatic hopper; all operations are controlled by proportioning methods common to the chemical industry. The ratio of materials is about 25% cellulose acetate, 4% water, less than 1% ground wood pulp as a filter aid, and the remainder acetone. The mixture moves forward through two or three stages at the rate at which it is used, the hold time being determined by experience. After dissolution is completed, filtration is carried out in batteries of plate and frame filter presses in three or even four stages, the passage of the "dope" being through presses of decreasing porosity.

Much of the cellulose acetate is delustered by the addition of titanium dioxide pigment, as with viscose rayon. Between filtrations (and after the last filtration), the dope goes to storage tanks that serve to remove bubbles; in this case, a vacuum is not necessary. From the final storage tank, it is pumped into a header located at the top of each spinning machine; then it is directed to a series of metering gear pumps, one for each spinnerette. Because the holes in the cellulose acetate spinning spinnerette are smaller (0.03–0.05 mm) than those in the corresponding viscose devices, great care must be taken with the final filtration. An additional filter for the removal of any small particles that may have passed through the large filters is placed in the fixture, sometimes called the "candle," to which the spinnerette assembly is fastened. A final filter is placed in the spinnerette-assembly unit over the top of the spinnerette itself.

The method used for spinning cellulose acetate is "dry" spinning. The dope is heated (in some cases above the boiling point of acetone, 56.5°C) to lower its viscosity and thus reduce the pressure required to extrude it and to supply

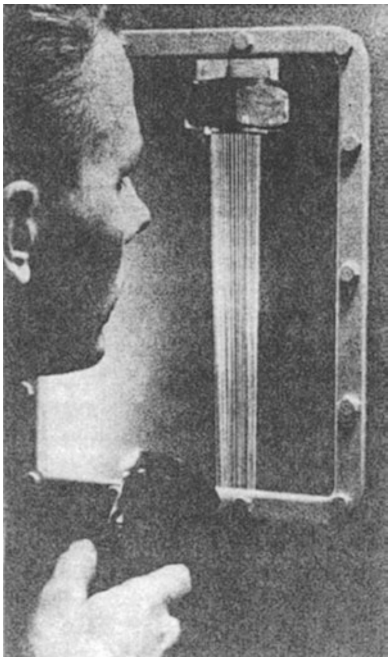


Fig. 12.12 Dry spinning of cellulose acetate (Courtesy of Tennessee Eastman Co.)

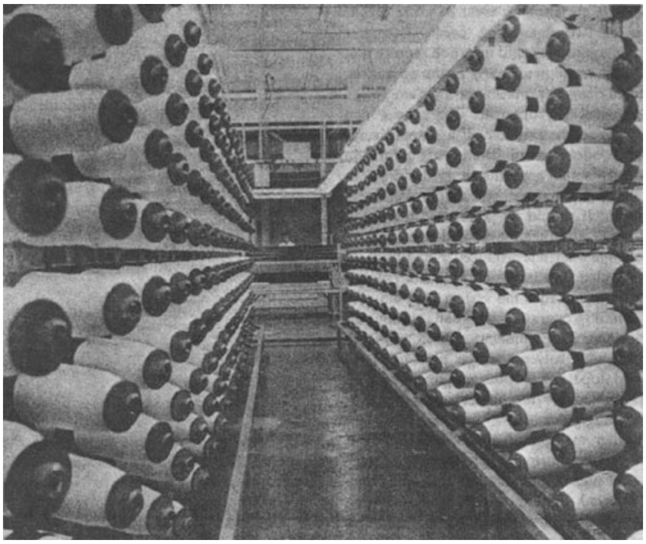


Fig. 12.13 Beaming cellulose acetate yarn from a reel holding about 800 packages of yarn (Courtesy of Tennessee Eastman Co.)

some of the heat needed for evaporating the acetone solvent (see Fig. 12.12).

The spinnerette is stainless steel, and because the filaments must be heated and prevented from sticking together and because space must be allowed for the escape of acetone vapor, the holes must be kept farther apart than those of the spinnerettes used for wet spinning. As the hot solution of cellulose acetate in acetone emerges downward into the spinning cabinet, an instantaneous loss of acetone takes place from the surface of the filaments, which tend to form a solid skin over the still liquid or plastic interior. A current of air, either in the direction the filaments are moving or countercurrent, heats the filaments, and as the acetone is diffused from the center through the more solid skin, each filament collapses to form the indented cross-sectional shape typical of cellulose acetate. The heated air removes the vaporized acetone. Each manufacturer uses a preferred updraft, downdraft, or mixed-draft operation, as needs dictate.

The cabinet through which the yarn passes vertically downward must be long enough to allow sufficient acetone to diffuse outward and evaporate from the surfaces of the filaments so that the latter will not stick to the first surface contacted or fuse to each other. The temperature of the air in the cabinet, the rate of flow, the length of the cabinet, the size and number of filaments, and the rate of travel are all interrelated in the spinning process. Because it is desirable to increase spinning speeds to the limit of the equipment, the tendency has been to construct longer spinning cabinets as each new plant is built. Present spinning speeds are of the order of 600 or more meters per minute, measured as the yarn emerges from the cabinet.

Other dry-spinning operations have followed essentially the same pattern. For example, the dry spinning of cellulose triacetate was identical to that for secondary acetate except that the acetone solvent had been replaced by a chlorinated hydrocarbon such as methylene chloride, the solubility of which was improved by the addition of a small amount of methanol (5–15%).

The acetate yarn emerging from the cabinet makes contact with an applicator that provides the lubricant required to reduce both friction and static formation in subsequent operations. With its surface lubricated, the yarn passes around a “feed” roll that determines the rate of withdrawal from the spinning cabinet and then to any of several desired packaging devices.

Unlike the packaging of rayon yarn, cellulose acetates are either “ring”-spun or wound into a package called a “disc,” “zero twist,” or “cam wound.” In the ring-spun package, the yarn carries a slight twist of less than one turn per inch, but it requires a relatively expensive bobbin. Since the trend is toward less twisting, such acetate yarn is “beamed” in the producer’s plant after little or no twisting, the heavy beams being shipped directly to knitters or weavers (see Fig. 12.13).

Filament yarns are twisted for two reasons. One is to supply certain esthetic characteristics such as touch, drapability, and elasticity. The other more fundamental reason is to provide physical integrity to the filament bundle so that it can be warped, woven, and knitted without excessive breakage or fraying of individual filaments.

The yarn just mentioned as having no twist imparted before beaming may have been subjected to intermingling just prior to windup after extrusion. In the intermingling process, yarn with no twist, and usually under low tension, is passed through a zone where it is impinged upon by a jet stream of compressed air. This causes the filaments to interlace or intermingle with each other, and they can become metastable in this configuration when tension is reapplied. In this condition, the yarn has the integrity of twisted yarn and will pass through several textile processing steps without difficulty, but with each handling, some of the intermingling is worked out.

Solvent Recovery. The air containing the acetone vapor is drawn out of the spinning cabinet and passed through beds of activated carbon that sorb the organic solvent. The acetone is recovered by steaming and then by separating it from the water by distillation. The efficiency of recovery is about 95%.

Dope Dying. As with viscose rayon, colored pigments or dyestuffs may be added to the spinning solution so that the yarn will be colored as it is produced, thus eliminating the need for dyeing the final fabric. As mentioned earlier, even in using titanium dioxide, a compromise must be made on the basis of two competing needs. Complete mixing, uniformity, and filtration require that the addition be made early in the operation; minimal cleaning problems during changeovers require just the opposite. There exist two solutions to the problem. If a manufacturer must produce a multitude of colors in relatively small amounts, it is desirable to premix individual batches of spinning dope. Each batch should be pretested on a small scale to ensure that the desired color will be acceptable when it is produced. Facilities must be provided to allow each batch of colored dope to be cut into the system very close to the spinning operation in order to minimize pipe cleaning. Permanent piping must be flushed with solvent or the new batch of colored dope; some of the equipment may be disassembled for mechanical cleaning after each change of color.

Another method of producing spun-dyed yarn involves using a group of "master" dopes of such color versatility that when they are injected by appropriate proportioning pumps into a mixer located near the spinning operation, they will produce the final desired color. The advantages of such an operation are obvious; the disadvantage lies in the public demand for an infinite number of colors. No small group of known pigments will produce final colors of every desired shade.

Protein Fibers

As previously mentioned, the use of naturally existing polymers to produce fibers has had a long history. In the case of cellulose, the results were fabulous. An initial investment of $930,000 produced net profits of $354,000,000 in 24 years for one rayon company [7]. On the other hand, efforts to use another family of natural polymers—proteins—have thus far resulted in failure or at best very limited production.

These regenerated proteins are obtained from milk (casein), soya beans, corn, and peanuts. More or less complex chemical separation and purification processes are required to isolate them from the parent materials. They may be dissolved in aqueous solutions of caustic, and wet-spun to form fibers, which usually require further chemical treatment as, for example, with formaldehyde. This reduces the tendency to swell or dissolve in subsequent wet-processing operations or final end uses. These fibers are characterized by a woollike feel, low strength, and ease of dyeing. Nevertheless, for economic and other reasons, they have not been able to compete successfully with either wool (after which they were modeled) or with other manufactured fibers.

Nylon

Historical

Nylon was the first direct product of the technological breakthrough achieved by W. H. Carothers of E. I. du Pont de Nemours and Co. Until he began his classic research on high polymers, the production of manufactured fibers was based almost completely on natural linear polymers. Such materials included rayon, cellulose acetate, and the proteins. His research showed that chemicals of low molecular weight could be reacted to form polymers of high molecular weight. By selecting reactants that produce linear molecules having great length in comparison with their cross section, fiber-forming polymers are obtained. With this discovery, the manufactured fiber industry entered a new and dramatic era.

Manufacture

Nylon 66. The word "nylon" was established as a generic name for polyamides, one class of the new high-molecular-weight linear polymers. The first of these, and the one still produced in the largest volume, was nylon 66 or polyhexamethylene adipamide. Numbers are used following the word "nylon" to indicate the number of carbon atoms contributed by the diamine and dicarboxylic acid constituents, in this case hexamethylenediamine and adipic acid, respectively.

To emphasize the fact that it does not depend on a naturally occurring polymer as a source of raw material, nylon often has been called a "truly synthetic fiber." To start the synthesis, benzene may be hydrogenated to cyclohexane

$$C_6H_6 + 3H_2 \xrightarrow{\text{Catalyst}} C_6H_{12}$$

or the cyclohexane may be obtained by fractionation of petroleum. The next step is oxidation to a cyclohexanol–cyclohexanone mixture by means of air:

$$xC_6H_{12} + O_2(\text{air}) \xrightarrow{\text{Catalyst}} yC_6H_{11}OH + zC_6H_{10}O$$

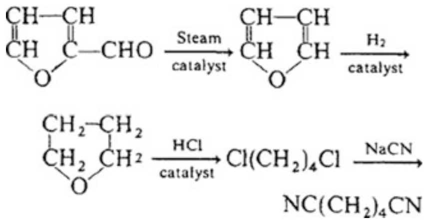
In turn, this mixture is oxidized by nitric acid to adipic acid:

$$C_6H_{11}OH + C_6H_{10}O + HNO_3 \xrightarrow{\text{Catalyst}} (CH_2)_4(COOH)_2$$

Adipic acid so obtained is both a reactant for the production of nylon and the raw material source for hexamethylenediamine, the other reactant. The adipic acid first is converted to adiponitrile by ammonolysis and then to hexamethylenediamine by hydrogenation:

$$(CH_2)_4(COOH)_2 + 2NH_3 \xrightarrow{\text{Catalyst}} (CH_2)_4(CN)_2 + 4H_2O$$
$$(CH_2)_4(CN)_2 + 4H_2 \xrightarrow{\text{Catalyst}} (CH_2)_6(NH_2)_2$$

Another approach is through the series of compounds furfural, furane, cyclotetramethylene oxide, 1,4-dichlorobutane, and adiponitrile, as illustrated below. The furfural is obtained from oat hulls and corn cobs.



Or 1,4-butadiene obtained from petroleum may be used as a starting raw material to make the adiponitrile via 1, 4-dichloro-2-butene and 1,4-dicyano-2-butene:

$$CH_2{=}CHCH{=}CH_2 \rightarrow ClCH_2CH{=}CHCH_2Cl$$
$$\xrightarrow[\text{Catalyst}]{HCN} NCCH_2CH{=}CHCH_2CN$$
$$\xrightarrow[\text{Catalyst}]{H_2} NC(CH_2)_4CN$$

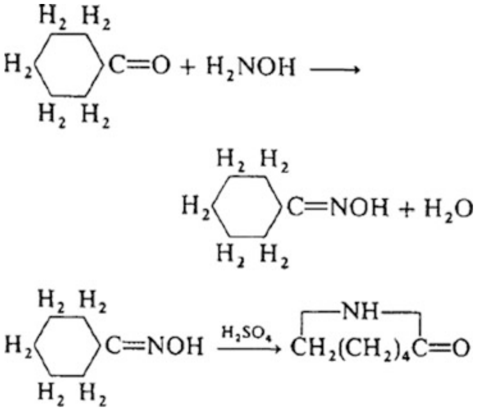
When hexamethylenediamine and adipic acid are mixed in solution in a one-to-one molar ratio, the "nylon salt" hexamethylenediammoniumadipate, the direct progenitor of the polymer, is precipitated. After purification, this nylon salt is polymerized to obtain a material of the desired molecular weight. It is heated to about 280°C under vacuum while being stirred in an autoclave for 2–3 h; a shorter holding period follows; and the process is finished off at 300°C. The molecular weight must be raised to a level high enough to provide a fiber-forming material, yet no higher. If it is too high, the corresponding viscosity in the subsequent spinning operation will require extremely high temperatures and pressure to make it flow. Accordingly, a small amount of acetic acid is added to terminate the growth of the long-chain molecules by reaction with the end amino groups.

The polymerized product is an extremely insoluble material and must be melt-spun, as discussed later. Therefore, should a delustered or precolored fiber be desired, it is necessary to add the titanium dioxide or colored pigment to the polymerization batch prior to solidification. For ease of handling, the batch of nylon polymer may be extruded from the autoclave to form a thin ribbon, which is easily broken down into chips after rapid cooling. But, whenever possible, the liquid polymer is pumped directly to the fiber melt-spinning operation (see Fig. 12.14).

Nylon 6. Nylon 6 is made from caprolactam and is known as Perlon® in Germany, where it was originally developed by Dr. Paul Schlack [8]. Its production has reached a very large volume in the United States in recent years.

Like nylon 66, nylon 6 uses benzene as raw material, which is converted through previously mentioned steps to cyclohexanone.

This compound is in turn converted to the corresponding oxime by reaction with hydroxylamine, and cyclohexanone oxime is made into caprolactam by the Beckmann rearrangement.



After purification, the lactam is polymerized by heating at elevated temperatures in an inert atmosphere. During self-condensation, the ring structure of the lactam is opened so that the monomer acts as an epsilon aminocaproic acid radical. Unlike that of nylon 66, the polymerization of caprolactam is reversible; the polymer remains in equilibrium

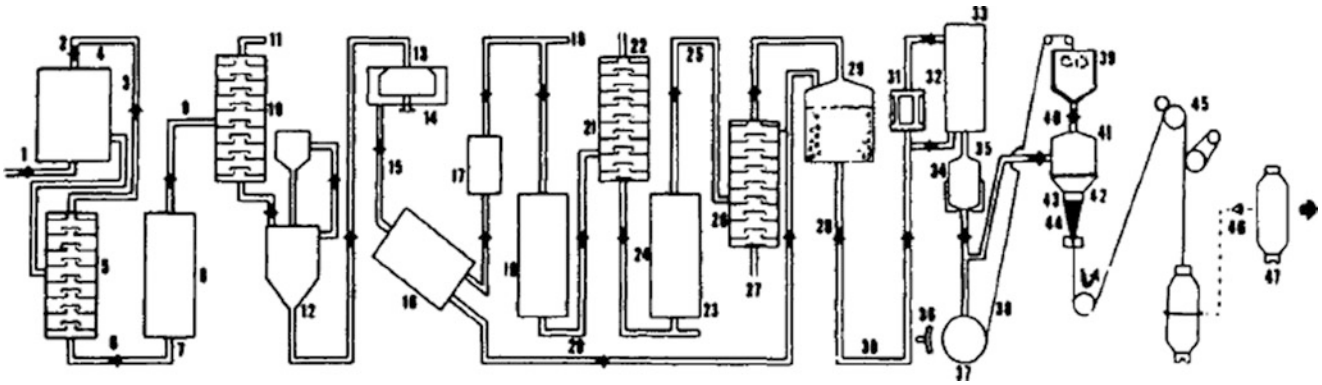


Fig. 12.14 Flow diagram for the manufacture of nylon 66 yarn: (1) air, (2) cyclohexane from petroleum, (3) reactor, (4) recycle cyclohexane, (5) still, (6) cyclohexanol–cyclohexanone, (7) nitric acid, (8) converter, (9) adipic acid solution, (10) still, (11) impurities, (12) crystallizer, (13) centrifuge, (14) impurities, (15) adipic acid crystals, (16) dryer, (17) vaporizer, (18) ammonia, (19) converter, (20) crude adiponitrile, (21) still, (22) impurities, (23) hydrogen, (24) converter, (25) crude diamine, (26) still, (27) impurities, (28) nylon salt solution, (29) reactor, (30) stabilizer, (31) calandria, (32) evaporator, (33) excess water, (34) autoclave, (35) delustrant, (36) water sprays, (37) casting wheel, (38) polymer ribbon, (39) grinder, (40) polymer flake, (41) spinning machine, (42) heating cells, (43) spinnerette, (44) air, (45) draw twisting, (46) inspection, (47) nylon bobbin (*Note*: Whenever the demand for liquid polymer at a spinnerette is large, as, for example, in the spinning of tire yarn, it is pumped directly from the autoclave)

with a small amount of monomer. As with nylon 66, nylon 6 is extruded in thin strands, quenched, and cut into chips for subsequent spinning, or the molten polymer is pumped directly to the spinning equipment.

Melt Spinning

Because of its extremely low solubility in low-boiling and inexpensive organic solvents, nylon 66 required a new technique for converting the solid polymer into fibers, hence the development of "melt" spinning, the third basic method for producing manufactured fibers. The following description refers essentially to nylon 66 because it was the first to use the method, but the process applies, in general, to all melt-spun manufactured fibers.

In the original production of nylon fiber by melt spinning, the chips of predried polymer were fed from a chamber onto a melting grid whose holes were so small that only passage of molten polymer was possible. Both solid and liquid were prevented from contacting oxygen by maintaining an inert nitrogen atmosphere over the polymer supply. The polymer melted in contact with the hot grid and dripped into a pool where it became the supply for the spinning itself. This melting operation has been entirely replaced by delivery of the molten polymer pumped directly from the polymerization stage or by "screw" melting. In the latter process, the solid polymer in chip form is fed into an extrusion-type screw contained in a heated tube. The depth and the helix angle of the grooves are engineered in such a way that melting takes place in the rear section, and the molten polymer is moved forward under increasing pressure to a uniformly heated chamber preceding the metering pump.

Whatever means is used to secure the molten polymer, it is moved forward to a gear-type pump that provides both high pressure and a constant rate of flow to the final filter and spinnerette. The filter consists of either sintered metal candle filters, several metal screens of increasing fineness, or graded sand arranged in such a way that the finest sand is at the bottom. After being filtered, the molten polymer at a pressure of several thousand pounds per square inch is extruded through the small capillaries in the heavily constructed spinnerette. It is necessary to maintain the temperature of the pool, pump, filter, spinnerette assembly, and spinnerette at about 20–30°C above the melting point of the nylon, which is about 264°C for nylon 66 and 220°C for nylon 6. Fibers having desired cross-sectional shapes can be produced by selecting spinnerettes containing holes of appropriate configuration. An example of a trilobal spinnerette capillary and the shape of the resulting trilobal fiber is given in Fig. 12.15.

The nylon production process requires that the extruded fibers emerge from the spinnerette face into a quench chamber where a crosscurrent of relatively cool quench air is provided to promote rapid solidification. The solid filaments then travel down a chimney to cool further, and a lubricant is applied before they make contact with the windup rolls in order to prevent static formation and to reduce friction in subsequent textile operations. The freshly spun yarn from the spinning chamber is taken up by a traversing winder onto a yarn package and "drawn" in a separate operation. In modern high-speed processes, drawing still is required, but in many cases, this is combined with spinning in a single operation, as will be described in what follows.

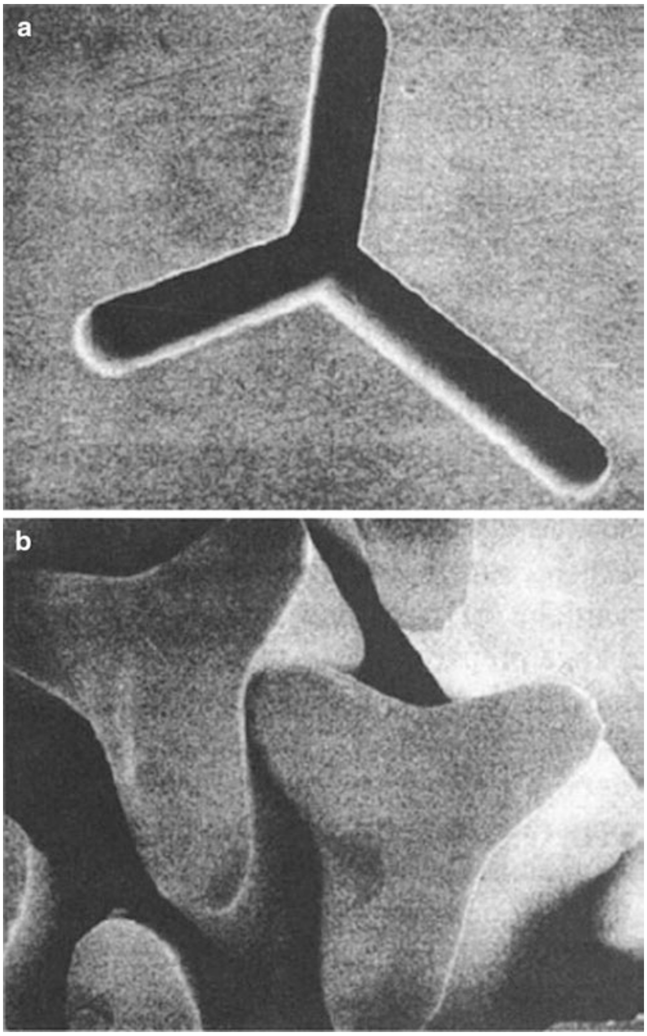


Fig. 12.15 Scanning electron micrographs of (**a**) trilobal-shaped spinnerette and (**b**) resulting cross sections of nylon fibers. The pictures were taken at different magnifications (Courtesy of BASF Corporation)

Drawing

It was learned early that the "as-spun" fibers made from nylon 66 could be extended to about four times their original length with very little effort, but that thereafter a marked resistance to extension took place. It was discovered that during this high extension, the entire length of fiber under stress did not extend uniformly. Rather, a "necking down" occurred at one or more points, and when the entire length under tension had passed through this phenomenon, a high-strength fiber was obtained. It also was found that when more than one necking down was allowed to take place in a given length of fiber, a discontinuity occurred at the point where the two came together. Accordingly, the drawing operation was aimed at forcing the drawing to occur at a single point as the yarn advanced from the supply to the take-up package.

Where still used, cold drawing consists essentially of removing the yarn from the package prepared in the melt-spinning operation and feeding it forward at a uniformly controlled rate under low tension. It is passed around a godet or roller that determines the supply rate and prevents slippage; for nylon 66, it then is wrapped several times around a stationary snubbing pin. From there, it goes to a second roller that rotates faster than the supply roller to produce the desired amount of stretch, usually about 400%. The necking down occurs at the pin. In the case of nylon 6, drawing may be effected satisfactorily without passing the yarn around such a snubbing pin.

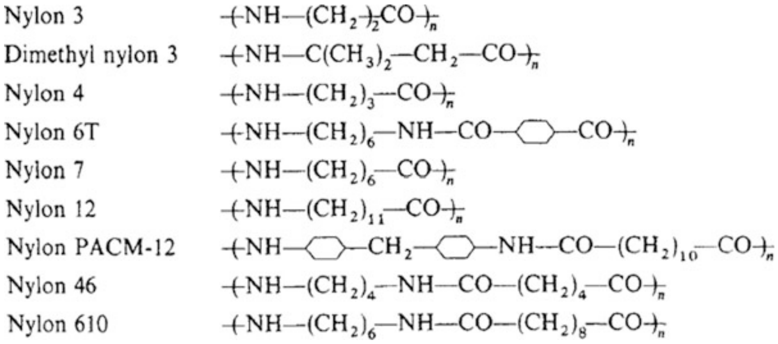
The long molecules of the nylon 66 or 6 polymer, which are randomly positioned in the molten polymer, when extruded from the spinnerette tend to form "crystalline" areas of molecular dimensions as the polymers solidify in the form of freshly spun fibers. In the drawing operation, both these more ordered portions and the amorphous areas tend to become oriented so that the lengthwise dimensions of the molecules become parallel to the long axis of the fiber, and additional intermolecular hydrogen bonding is facilitated. It is this orientation that converts the fiber having low resistance to stress into one of high strength.

By controlling the amount of drawing as well as the conditions under which this operation takes place, it is possible to vary the amount of orientation and the degree of crystallization. A minimal amount is preferable in the manufacture of yarns intended for textile applications wherein elongation of considerable magnitude and low modulus or stiffness are required rather than high strength. On the other hand, high strength and modulus are at a premium when fibers are to be used in tire cords and other industrial applications. High resistance to elongation is imperative if the tire is not to grow under conditions of use. In this connection, it should be noted that nylon tire cord that has been produced by twisting the original tire yarn and plying the ends of these twisted yarns together is hot stretched just before use at the tire plant to increase strength and reduce even further the tendency to elongate under tension.

The separate operations of spinning and drawing nylon presented a challenge whose object was combination of the two operations into a single continuous step. But the problem was obvious, for the operating speeds of the two separate steps already had been pushed as high as was thought to be possible. How then would it be possible to combine them into a continuous spin draw, wherein a stretching of about 400% could take place? The answer lay in the manner in which the cooling air was used and in the development of improved high-speed winding devices. By first cooling the emerging fibers by a concurrent flow of air and then cooling them further by a countercurrent flow, the vertical length of the cooling columns can be kept within reason. In-line drawing may occur in one or two stages, and relaxation may be induced if needed. The final yarn is said to be packaged at speeds of 6,000 m/min.

Other Nylons, Modifications, and New Developments

Although nylon 66 and 6 account for most of the polyamide fibers produced, a great many others have been experimentally synthesized and have been developed and manufactured in commercial amounts. Of these, some have been made into fibers, some with limited economic success. These nylons are identified by either the same numbering system used for nylon 66 or 6 or by a combination of numbers and letters, as follows:



Dimethyl nylon 3 is solution-spun because it tends to decompose during melt spinning. Nylon 4 has a moisture regain (mass water per unit mass of dry fiber, under standard atmospheric conditions of 20°C and 65% RH) of 6–9% and therefore is superior to other nylons for textile usages, being comparable to cotton [9]. Nylon 11 was developed in France and has been trademarked as Rilsan®. It has a moisture regain of 1.8% and density of 1.04 g/cc as compared with 4% and 1.14 g/cc, respectively, for nylon 66. Nylon 7 is made in the former Soviet Union and marketed under the name Enant. The fiber has better stability to heat and ultraviolet light than nylon 66 and 6. Nylon 6T, an aromatic polymer, has a much higher melting point (370°C), a higher density (1.21 g/cc), and slightly higher moisture regain (4.5%) than nylon 66. It also has superior resistance than nylon 66 against heat. This fiber has served as a precursor to the development of aramid fibers. Nylon PACM-12, formerly produced under the trade name Qiana® in the United States, is no longer in production.

A *chemical and engineering news* report [10] suggested that the most serious competition to nylon 66 and 6 will be provided by a new, still experimental fiber, nylon 46, which was being developed by DSM in the Netherlands, trade-named Stanyl®; this fiber results from the interaction of 1,4-diaminobutane and adipic acid. Better order in the structure in the fiber leads to greater crystallinity and, thus, to greater density (1.18 g/cc). The fiber has a melting point of about 300°C, a breaking stress or tenacity of 9.5 g force/denier (~1 GPa), modulus at 120°C of 20 g force/denier (2.1 GPa), and shrinkage in gas at 160°C of 3%.

Some of the outstanding characteristics of nylon that are responsible for its many uses in apparel, home furnishing, and industrial products are its high strength and toughness, elastic recovery, resilience, abrasion resistance, and low density. Among many applications of the fiber are such products as intimate apparel and foundation garments, sportswear, carpets, parachutes, tents, sleeping bags, and tire cords.

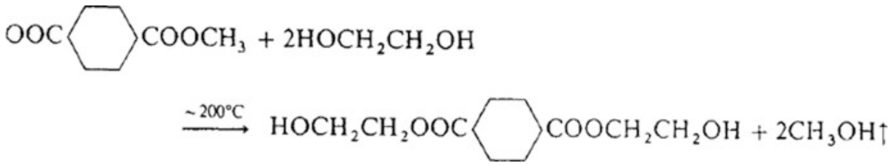
The world production of nylon has doubled in recent years, increasing from 3.8 billion lb in 1970 to about 7.5 billion lb in 1990. The fiber accounts for about 24% of the synthetic fibers produced worldwide. In the United States, the production of the fiber also doubled in two decades, increasing from 1.2 billion lb in 1970 to about 2.4 billion lb in 1990.

The newest activity connected with nylon is the effort at developing a more highly oriented stronger nylon than possible by the current technology. Because of the formation of hydrogen bonds between the chains, the normal polymer is restricted in terms of the maximum draw ratio by which it can be oriented. In the new technique [11], nylon 66 is dissolved in an agent such as gallium trichloride, which effectively breaks the hydrogen bonds. The solution is spun by the dry jet–wet spinning method. The $GaCl_3$/nylon 66 complex so obtained can now be stretched to very high draw ratios, levels as high as 40× have been possible. Once drawn, the structure is soaked in water to remove gallium trichloride, which allows the hydrogen bonds to reestablish and link the chains. In preliminary work done thus far, the strength and the modulus obtained exceed the values usually found in nylon 66.

Polyesters

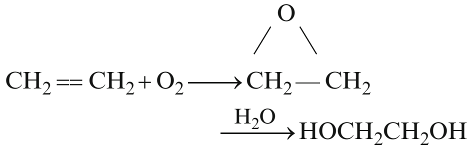
Historical

The stimulus for the development of polyester, as for nylon, was provided by the fundamental work of Carothers. Although his team's initial work was directed toward this material, because of greater promise shown by polyamides at the time, the developmental work on polyesters was temporarily set aside. The polymer, however, attracted interest in Great Britain, where J. T. Dickson and J. R. Whinfield experimented with it and developed a successful polyester fiber [7]. They found that a synthetic linear polymer could be produced by condensing ethylene glycol with terephthalic acid or by an ester exchange between the glycol and pure dimethyl terephthalate. The polymer thus obtained could be converted to fibers having valuable properties, including the absence of color. Like nylon, this material has been popularized under its generic name, polyester or just "poly." Those persons working with it commonly refer to it as polyethylene terephthalate (PET). It first appeared under the trade name Terylene® (Imperial Chemical Industries, Ltd.), in England, and was first commercialized in the United States in 1953 as Dacron® (E. I. du Pont de Nemours and Co.).



Manufacture

When the development of PET occurred, ethylene glycol already was being produced in large amounts from ethylene, a by-product of petroleum cracking, by the oxidation of ethylene to ethylene oxide and subsequent hydration to ethylene glycol, which, in a noncatalytic process, uses high pressure and temperature in the presence of excess water.



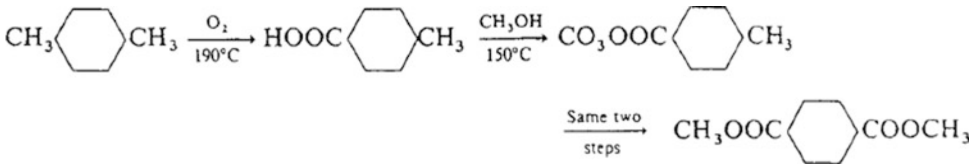
$$\mathrm{CH_2{=}CH_2 + O_2 \longrightarrow \overset{O}{CH_2{-}CH_2} \xrightarrow{H_2O} HOCH_2CH_2OH}$$

On the other hand, although *o*-phthalic acid, or rather its anhydride, had long been produced in enormous amounts for use in the manufacture of alkyd resins, the *para* derivative was less well known and not available on a large scale. The synthesis is a straightforward one, however, from *p*-xylene, which is oxidized to terephthalic acid, either by means of nitric acid in the older process or by air (catalyzed) in the newer one. In the early years, this compound then was converted to the easily purified dimethyl ester in order to obtain a colorless polymer adequate for the manufacture of commercially acceptable fibers.

Several other methods were developed for producing the desired dimethyl terephthalate. The Witten (Hercules) process goes from *p*-xylene to toluic acid by oxidation of one of the methyl groups on the ring, following which the carboxyl group is esterified with methanol. This process then is repeated with the second methyl group to secure the dimethyl ester of terephthalic acid.

Either phthalic anhydride or toluene, both in ample supply as raw materials, can be used in the Henkel processes. Use of phthalic anhydride depends only upon dry isomerization of the potassium salt of the *ortho* derivative to the *para* form at about 430°C and 20 atmospheric pressure, or toluene is oxidized to benzoic acid, whose potassium salt can be converted to benzene and the potassium salt of terephthalic acid by disproportionation.

The first step in the reaction of dimethyl terephthalate and ethylene glycol is *trans*esterification to form bis (*p*-hydroxyethyl) terephthalate (bis-HET) and eliminate methanol. This product then is polymerized in the presence of a catalyst to a low-molecular-weight compound, and the by-product glycol is eliminated. In a second stage, at a temperature of about 275°C and under a high vacuum, the molecular weight is raised to secure the melt viscosity desired for the particular material involved. Like nylon, this final material may be extruded, cooled, and cut into chips for storage and remelting, or it may be pumped directly to the spinning machines.

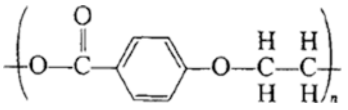


From the beginning, it was obvious that there would be considerable progress in industrial chemistry, to say nothing of cost reduction, if the process could be simplified by making it unnecessary to go through the dimethyl derivative to secure a product of adequate purity. This was accomplished in the early 1960s when methods of purifying the crude terephthalic acid were developed, and conditions and catalysts were found that made possible the continuous production of a color-free polymer. It is said that the selection of the catalyst is especially aimed at the prevention of ether linkages in the polymer chain due to intracondensation of the glycol end groups.

Two additional rather similar routes are known. Both depend upon the reaction between ethylene oxide, rather than ethylene glycol, and terephthalic acid to form the bis-HET monomer already mentioned. The difference between the two methods lies in the point where purification is done: in one case, it is the crude terephthalic acid; in the other, it is the bis-HET monomer. In both cases, this monomer is polymerized by known procedures to form a fiber-grade polyester. The titanium dioxide delustrant is added, as might be expected, early in the polymerizing process.

Another polyester that has reached long-term commercialization is now produced in limited volume as Kodel 200® by Tennessee Eastman Co. and is considered to be 1–4 cyclohexylene dimethylene terephthalate. The glycol that is used instead of ethylene glycol in this process exists in two isomeric forms, one melting at 43°C and the other at 67°C. This makes possible their separation by crystallization, to secure the desired ratio of the two forms for conversion to the polymer. This ratio determines the melting point of the polymer, a most important property for a material that is to be melt-spun. The polymer from the 100% *cis* form melts at 275°C and that from the 100% *trans* form at 318°C. Indications are that the commercial product is about 30/70 *cis–trans*.

In 1973, the Federal Trade Commission modified the generic definition of polyester to include in the polyester category materials that previously were polyester ethers or benzoate polyesters. As a result, the fiber known as poly (ethylene oxybenzoate) or PEB and manufactured under the trade name A-Tell in Japan came to be known as polyester. This material is made by reacting parahydroxybenzoic acid and ethylene oxide to give paraoxyethylenebenzoic acid, which is then polymerized to obtain PEB:



The fiber softens at about 200°C and melts at 225°C. It is said to have a silklike hand and appearance and other properties comparable to those of other polyesters.

Polyesters are melt-spun in equipment essentially the same as that used for nylon, already described. Wherever the volume is large and the stability of demand is adequate, the molten polymer is pumped directly from the final polymerization stage to the melt-spinning machine. The molten polymer is both metered and moved forward at high pressure by the use of an extruder coupled with a gear-type pump, through filters to the spinnerette, which contains capillaries of about 9 mil (230 μm) in diameter. Great care is taken to eliminate moisture and oxygen from the chips, if they are used, and from the spinning chamber. When the polyester fibers are destined to become staple, the emerging filaments from a number of spinnerettes are combined to form a tow, which can be further processed as a unit. Continuous-filament yarn is packaged for further processing such as drawing or texturing. Spin drawing, described later, has become commonplace today and represents major cost savings to the fiber manufacturer.

Drawing

Unlike nylon, which in the as-spun state contains a high amount of crystalline component, PET fibers are essentially amorphous as spun. In order to secure a usable textile yarn or staple fiber, this product must be drawn under conditions that will result in an increase in both molecular orientation and crystallinity. This is done by drawing at a temperature well above the glass transition point, T_g, which is about 80°C. Conditions of rate and temperature must be selected so that the amorphous areas are oriented, and crystallization will take place as the temperature of the drawn fibers drops to room temperature.

An appropriate contact-type hot plate or other device is used, and about 300–400% extension is effected [10]. Figure 12.16 shows a filament withdrawn and undrawn segments. As with nylon, the conditions of draw, especially the amount, determine the force–elongation properties of the

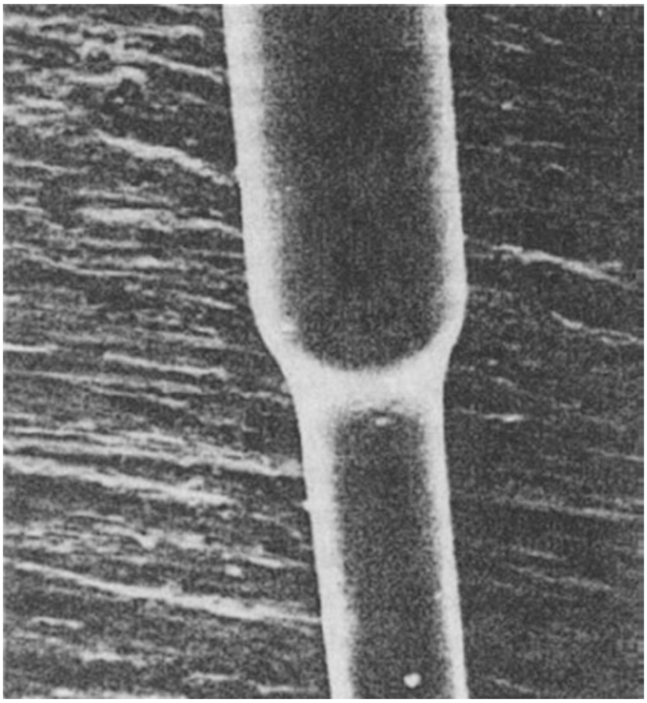


Fig. 12.16 Drawnecking in polyester single filament (Courtesy of E. I. du Pont de Nemours and Co.)

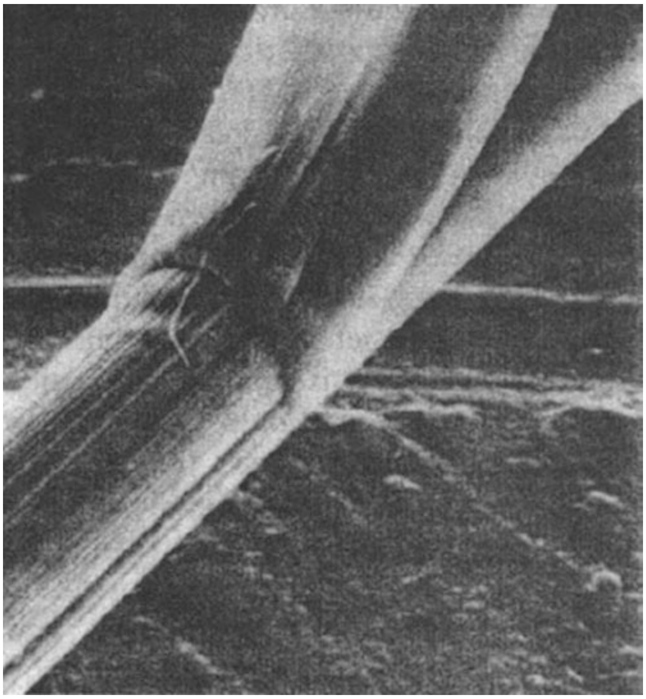


Fig. 12.17 Skin peeling in polyester showing fibrillar structure (Courtesy of E. I. du Pont de Nemours and Co.)

product. Industrial-type yarns, such as those intended to be used as tire cord, are more highly drawn than other yarns and have higher strength with less elongation. The fibers develop the much desired fibrillar morphology for such applications (Fig. 12.17).

Heat Setting

The ability of textile fibers to be "set" is not characteristic of manufactured fibers alone. Aided in many cases by the presence of starch, cotton fabrics can be ironed to a smooth and wrinkle-free condition; also, a sharp crease in wool trousers has been commonplace for generations. In other words, these fabrics were exposed to moisture at elevated temperatures while being held or pressed into desired geometrical configurations and then allowed to cool before being released from constraint. Such fabrics tend to remain unchanged while cool and dry, even though the fibers from which they are formed carry internal stresses, but reversion takes place upon washing or exposure to high relative humidity.

With the development of nylon, and especially polyesters, a durable kind of setting has become possible. When fabrics made from these fibers are shaped and then exposed to elevated temperatures either in the dry condition or, in the case of nylon particularly, in the presence of water vapor, thermoplastic relaxation of induced stresses in the fiber takes place and configurations at the molecular level adjust to a new and lower energy level. This depends on not only the temperature used but also the duration of the exposure. Thus a few seconds at 230°C will produce the same results as exposure for a considerably longer period at a temperature 50–75 ºC lower. The permanency of the setting, that is, the ability of a fabric or garment to return to its original configuration after temporary distortion even while exposed to moisture and raised temperatures, is a function of the severity of the heat setting. To impart true permanence, it is essential that the internal crystalline structure be annealed.

It is this property of polyamides and polyesters that has been the main factor contributing to "ease-of-care" and the "wash-and-wear" characteristics of garments made from these polymers. In turn, these garments have revolutionized both the textile and the apparel industries.

Textured Yarns

Fundamentally, the manufacture of "textured" yarns is closely related to the heat setting of fabrics, which must be composed of thermoplastic fibers such as nylon or polyester, the difference being that the individual filaments or bundle of filaments in textured yarns are distorted from an essentially straight rodlike form and then heat-set. In some instances, the fibers are distorted in a more or less random way; at other times, a regular pattern is introduced.

The first commercially successful textured yarn was produced by highly twisting nylon 66, heat-setting it as a full package of yarn, and then untwisting it through zero and a small amount of twist in the opposite direction. This process changed yarn from a close-packed structure to one that was voluminous because of mutual interference of distorted filaments. The technique of heat-setting the twisted yarn as a batch-unit operation now has been replaced by a continuous operation, using what is known as a "false-twisting" process. This is based upon the principle that if a length of yarn is prevented from rotating at both ends but is rotated on its axis at its center point, the resulting two sections will contain both "Z" and "S" twists in equal amounts. When this occurs with a moving yarn, any element in it will first receive a twist in one direction but, after passing the false-twisting point, must revert to zero twist. If it is then made to pass over a hot plate while in the twisted state and is heat-set in that configuration, even after returning to the untwisted condition, the individual filaments will tend to remain distorted when lengthwise stress is released. Because of the low mass and diameters of textile yarns or monofilaments, it is possible to false twist them at extremely high rotational speeds. Yarn forward speeds of about 1,000 m/min are currently obtainable by passing the yarn between, and in contact with, high-speed-friction twisting discs. (When attempts are made to secure higher rates, problems of twist control develop.) The same technique is now more commonly applied to unoriented (undrawn) or partially oriented yarn (POY) at the draw-texturing machine. The resulting yarn may be heat-set as part of the same continuous operation by passing it through a second heater under conditions of overfeed or little or no tension in order to secure both thermally stable geometric configurations in the individual distorted filaments that comprise the yarn and the degree of "stretchiness" and bulk desired in the final product.

Because these yarns are being made in one less step and also within the plants spinning the parent product, this latest development may be said to constitute another advance in the industrial chemical technology of manufactured textile products. This draw texturing appears to be especially applicable to polyester yarns intended for fabrics known as "double knits" and "textured wovens."

Yarn can be forced forward by means of "nip" rolls, although this may seem to be quite contrary to the old adage that one cannot push on an end of string. When this is done so that the yarn is jammed into a receiver (stuffer box) already full of the preceding materials, it collapses with sharp bends between very short lengths of straight sections. In this condition, heat is applied, usually in the form of superheated steam, to set it. In practice, the mass of such yarn is pushed through a heated tube until it escapes at the exit past a spring-loaded gate. During this passage, it is heat-set in a highly crimped configuration; then it is cooled before being straightened and wound onto a package. In another continuous process, the yarn or monofilament is pulled under tension over a hot sharp edge so that it is bent beyond its elastic limit and is heat-set in that condition. The process is known as "edge crimping," and the result is not unlike that produced by drawing a human hair over the thumbnail. The process is not used much today, but a yarn with similar crimp is produced by bicomponent spinning.

When such yarns are knitted or woven into fabric, the filaments tend to return to the configurations in which they were originally heat-set. Contraction takes place in the direction of the yarn axis, and this in turn converts the smooth flat fabric into a "stretch" fabric and gives the surface a textured appearance. These fabrics or the garments made from them, whatever the process used to produce the yarns, may be given additional heat treatment to secure stability in a desired geometric configuration. A degree of stretch may be retained, or a flat and stable textured surface may be produced. There are a number of variations of the texturing process, which, combined with the many possibilities of heat setting, impart considerable versatility to the final product (see Fig. 12.18).

The growth in the use of these products in the 1960s is well known. Carpeting also provides a significant market for them, as texture is one of the most important characteristics of soft floor coverings. Such products have been important to the successful use and expanded development of nylon and polyester yarns.

Staple Process

Unlike nylon, which is used mostly in the form of continuous-filament yarn, polyester is utilized both in staple and in continuous-filament form in large volumes. For producing staple fiber, the spinning machines consist of a series of packs, 10–30, each with 1,000–3,000 holes. The extrudates from different packs after solidification and application of finish are combined to form a subtow and collected in a can. Then several such cans are brought together in a creel area, and subtows from these cans are combined to feed a staple drawline. The latter may involve one or more stages of drawing and relaxation, one or more stuffer-box-type crimpers, a drying unit, and a cutter. A baling unit following the cutter collects and bales the cut fiber.

Continuous-Filament Yarn Process Variants

For producing continuous-filament yarn, several process routes are available [12], each of which involves the basic processes of polymer synthesis, extrusion, quenching,

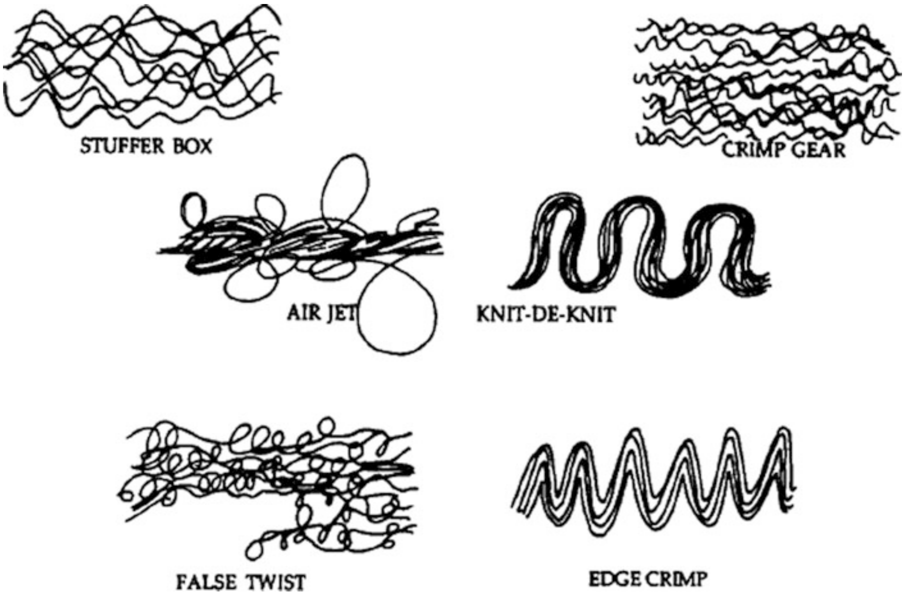


Fig. 12.18 Schematic comparison of the shapes of some textured yarns

stretching, and winding. In one process, specially meant for textile-type uses, spinnerettes with 20–100 holes typically are used. After solidification and finish application, the filaments may be split into smaller bundles, depending upon end-use applications, for downstream processing. The drawing can also be done in a single integrated process (spin draw) immediately after finish application to yield a fully oriented yarn (FOY), which, having no bulk or texture, is referred to as a flat yarn. Spinning speeds of the order of 1,000 m/min or more and winding speeds of about 4,000 m/min are used. Similar yarn may be produced on a separate drawing unit, but this process—called draw twisting because a certain amount of twist is also inserted prior to winding—is less commonly used today. It is more usual to see the separate drawing process done in conjunction with a texturing process—most commonly false twisting and called draw texturing—to give a textured yarn. The process leads to orientation and crystallization of the filament structure and bulking of the yarn. The feed yarn for such processes usually is produced at 3,000–4,000 m/min spinning speed as POY. The latter is designed to be run on draw-texturing machines at speeds ranging from 800 to 1,200 m/min.

For industrial applications, polyester filaments having high strength low shrinkage (HSLS), low creep properties, and high glass transition temperatures are targeted. To produce such a filament, more severe processing conditions and higher-molecular-weight polymer are generally used. The filament is spun at low speeds (500–1,000 m/min), sometimes with retarded quench, to obtain minimum orientation. The drawing can be achieved in the more common integrated, or spin draw, process or in a separate draw–twist operation. High strength is achieved by drawing the filament to several times its length over very high-temperature rolls and then heat-setting and relaxing the structure prior to winding. Low shrinkage properties are obtained with a relaxing step at high temperature.

Modifications and New Developments

As was the case with the nylon fibers, the potential the polyester fibers offered in apparel, home furnishing, and industrial applications was judged to be enormous. For this potential to be realized in practice, however, some characteristics had to be improved, and others had to be engineered for specific end uses. Thus, fibers of different cross-sectional shapes were developed in order to impart antisoiling, reflective, and resilient characteristics for rug and carpet applications. A difficulty associated with the early polyester fiber that restricted its applications was its lack of ability to take on dyes through one of many methods available for dyeing. This problem was overcome by introducing chemicals that added sulfonate groups to the molecule and by substituting in some cases isophthalic acid for a small portion of terephthalic acid. These changes allowed fibers to be dyed by cationic and disperse dyes, the dyes most frequently used for polyester. Another area of modification has been the development of inherently flame-resistant fiber. One process involves copolymerizing a derivative of phosphoric acid with PET. An exciting new development in polyester filament yarn for apparel uses is the production of microdenier fiber (denier per filament less than one), discussed later in a separate section. The introduction of finer denier yarns opened up a whole new field for developing fabrics with special esthetic and performance characteristics that were not possible earlier.

With the use of fibers in conveyor belts, tires, and composites, fibers of greater strength and modulus and lower extensibility have been needed. Much effort was directed in the 1970s and, later, to developing such fibers from polyester. The composition and the properties of wholly aromatic polyamides or aramids are discussed in a later section. When both the diacid and diamine components are *para*-substituted aromatic compounds, the resulting polymer is capable of forming lyotropic liquid–crystalline solutions. These solutions can be dry- or wet-spun into fibers with unusually high tensile strength and tensile modulus. When a similar strategy is tried to make polyester fiber from a homopolymer of a *para*-substituted aromatic diacid and a *para*-substituted aromatic diol, only infusible and intractable materials are obtained. A solution to this problem has been found in the development of polyester copolymers that give thermotropic liquid–crystalline melts over a useful temperature range and have viscosities suitable for melt extrusion into fibers or films having high levels of orientation. Spin-line stretch factors of the order of several hundred percent are used to achieve orientation, and physical properties are developed further by heat treatment at temperatures approaching melting conditions.

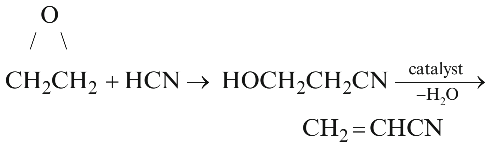
The first fibers from a thermotropic liquid–crystalline melt whose properties were reported were spun from a copolyester of *para*-hydroxybenzoic acid (PHB) and PET by workers at Tennessee Eastman Co. The preparation of the copolymer proceeds in two stages. First, *para*-acetoxybenzoic acid is reacted with PET in an acidolysis step to give a copolyester prepolymer, which in the second step is condensed further to a higher degree of polymerization suitable for fiber formation.

Among melt-spun fibers, those based on thermotropic liquid–crystalline melts have the highest strength and rigidity reported to date and appear comparable to polyamides spun from lyotropic liquid–crystalline solutions. This was a very active field of research in the 1970s, and later, and many comonomers have been reported. Obviously, these compositions must contain three components at a minimum, but many have four or five components. Some frequently used constituents, in addition to those mentioned above, are 2,6-naphthalenedicarboxylic acid, hydroquinone, 4,4′-biphenol, isophthalic acid, and 4,4′-dihydroxy-diphenyl ether.

Acrylics

Polymer Manufacture

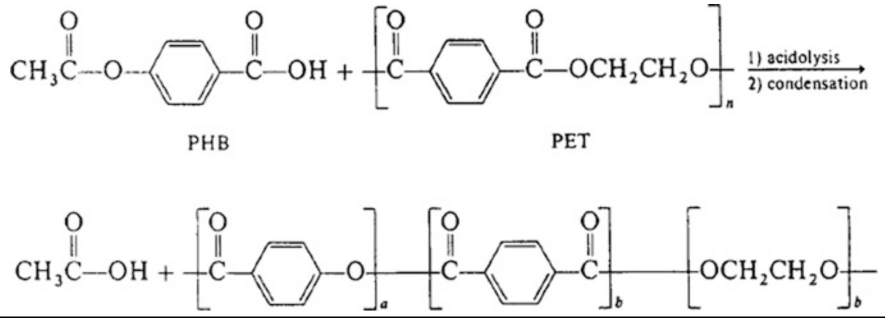
Acrylic fibers are spun from polymers that are made from monomers containing a minimum of 85% acrylonitrile. This compound may be made from hydrogen cyanide and ethylene oxide through the intermediate ethylene cyanohydrin.



$$\overset{\overset{\text{O}}{/\ \ \backslash}}{\text{CH}_2\text{CH}_2} + \text{HCN} \rightarrow \text{HOCH}_2\text{CH}_2\text{CN} \xrightarrow[-\text{H}_2\text{O}]{\text{catalyst}} \text{CH}_2{=}\text{CHCN}$$

It also may be made directly from acetylene and hydrogen cyanide:

$$\text{CH} \equiv \text{CH} + \text{HCN} \rightarrow \text{CH}_2 = \text{CHCN}$$



When the mol.% of PHB in the copolymer exceeds about 30–40%, a liquid–crystalline melt is obtained. Up to about 60 mol.%, order in the melt increases and melt viscosity decreases. Compositions containing about 60 mol.% PHB can be melt-spun into fibers using standard extrusion techniques. It is the unusual combination of properties that makes this class of materials valuable for the formation of high-strength fibers and plastics.

But the reaction that currently is preferred uses propylene, ammonia, and air:

$$\begin{aligned} &3\text{CH}_2 = \text{CHCH}_3 + 3\text{NH}_3 + 7\text{O}_2(\text{air}) \\ &\xrightarrow[<500\,^\circ\text{C}]{\text{Catalyst}} \text{CH}_2 = \text{CHCN} + 2\text{CO} + \text{CO}_2 \\ &\qquad + \text{CH}_3\text{CN} + \text{HCN} + 10\text{H}_2\text{O} \end{aligned}$$

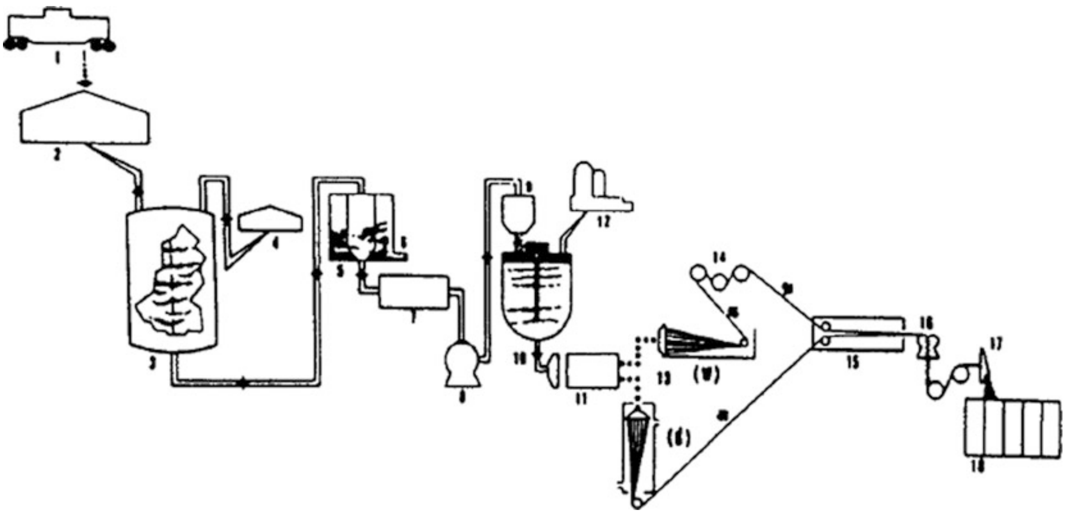


Fig. 12.19 Flow diagram for the manufacture of acrylic fiber: (1) acrylonitrile, (2) tank farm, (3) polymerizer, (4) comonomer and catalyst, (5) centrifuge, (6) waste liquid, (7) dried polymer, (8) grinding, (9) polymer storage, (10) dissolver, (11) filter, (12) solvent plant, (13) spinnerette, (13w) wet spinning, (13d) dry spinning, (14) roller dryer, (15) additional treatment, (16) crimper, (17) cutter, (18) acrylic fiber bale

Pure acrylonitrile may polymerize at room temperature to polyacrylonitrile (PAN), a compound that, unlike polyamides and polyesters, does not melt at elevated temperatures but only softens and finally discolors and decomposes. Nor is it soluble in inexpensive low-boiling organic solvents. Because fibers made from it resist the dyeing operations commonly used in the textile industry, the usual practice is to modify it by copolymerization with other monomers, for example, vinyl acetate, styrene, acrylic esters, acrylamide, or vinyl pyridine in amounts up to 15% of the total weight (beyond which the final product may not be termed an acrylic fiber). The choice of modifier depends on the characteristics that a given manufacturer considers important in a fiber, the availability and cost of the raw materials in the manufacturer's particular area of production, and the patent situation.

In copolymerizing acrylonitrile with another monomer, conditions must be controlled in such a way that the reaction produces a polymer having the desired chain structure and length. The reaction takes place in the presence of substances capable of producing free radicals. In addition, certain trace metals that have been found to increase reaction rates offer a means of controlling chain length. When polymerization is carried out in solution, after an induction period, the reaction is rapid and liberates a considerable amount of heat. Furthermore, because the polymer is not soluble in the monomer, a thick paste is formed. These facts limit the usefulness of such a process. Carrying out the polymerization in the presence of a large amount of water (water/monomer of 2/1 to 3/1) is a convenient method and the one most generally used. In this case, the polymer forms a slurry, and the water provides a means for removing the heat from the site of the reaction. Moreover, most of the common redox catalyst systems are water soluble. Polymerization may be carried out batchwise or by a continuous process.

In the batch method, the monomers and catalyst solutions are fed slowly into an agitated vessel containing a quantity of water. The heat of reaction is removed either by circulating cold water through the jacket surrounding the vessel or by operating the reaction mixture at reflux temperature and eliminating the heat through the condenser water. The monomer and catalyst feeds are stopped when the desired amounts have been added, and polymerization is allowed to continue until there is only a small amount of monomer remaining in the reaction mixture. Then the slurry is dumped from the reaction vessel, filtered, washed, and dried (see Fig. 12.19).

In the continuous overflow method, rather than stopping the monomer and catalyst feed when the reaction vessel is full, the slurry is simply allowed to overflow; the solids are removed by filtration, washed, and dried. The filtrate contains a certain amount of unreacted monomer, which is recovered by steam distillation after the trace metal present has been chelated to stop the redox reaction and thus further polymerization. The dried polymer is the raw material from which fibers are spun.

As mentioned, PAN polymer has two major weaknesses: (1) extremely poor dyeing characteristics due to highly ordered structure and (2) very low solubility in most of the solvents. To overcome these problems, comonomers are incorporated into the chains. Comonomers have been used to improve hydrophilicity, dyeability, flame retardancy, heat resistance, and so on. For obtaining hydrophilic acrylic fiber,

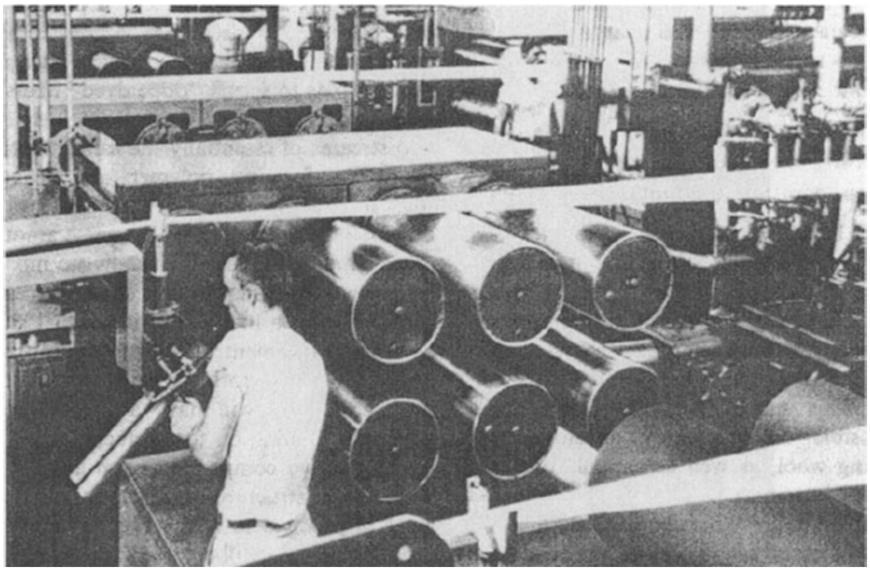


Fig. 12.20 Wet spinning of acrylic tow (Courtesy of Monsanto Co.)

comonomers containing hydrophilic functional groups such as hydroxyl, ester, carboxyl, amide, and substituted amide are used. To make the fiber base dyeable, comonomer containing carboxylic and sulfonic acid groups is introduced, and to obtain acid-dyeable fiber, comonomers containing pyridine, aliphatic amine, or quaternary ammonium salt are used. Antistatic acrylic fibers can be made by incorporating in the chain polyethylene oxide, polyalkylene derivatives, polyethylene glycol, acrylates, or methacrylates as the comonomers. For improved flame retardancy, halogen-containing vinyl comonomers are used.

Spinning

As already indicated, pure PAN softens at elevated temperatures, and thermal decomposition starts before the molten state is reached. The same is true of the copolymers commonly used to produce fibers. Accordingly, melt spinning is not possible; spinning must be done from a solution of the polymer. Both dry and wet spinning are carried out in current commercial operations.

The operations used to either wet- or dry-spin acrylics are essentially the same as those already described for rayon and acetate, respectively. The polymer must be completely dissolved in solvent and the solution filtered to remove any impurities that would cause spinnerette blockage. Because acrylic polymers are not soluble in common nonpolar solvents, polar substances such as dimethylformamide, dimethylacetamide, or aqueous solutions of inorganic salts such as zinc chloride or sodium thiocyanate are required. Only wet spinning is possible with the latter. Dimethyl formamide boils at 152.8°C and exerts a vapor pressure of 3.7 mm of Hg at 25°C compared with acetone (used in dry spinning of cellulose acetate), which has a vapor pressure of 228.2 mm of Hg at 25°C. It follows that, unlike acetone which requires an activated-carbon system for recovery, dimethylformamide may be condensed directly from the gas stream used to evaporate the solvent from the forming fiber.

In order to obtain the desired characteristics of modulus, rupture tenacity, and rupture elongation, acrylics, like rayon, require stretching which is usually carried out after the fiber has been spun, either as part of the fiber washing operation or after the fiber has been dried. These same properties are influenced by spinning speeds and the temperature of the drying air, if they are dry-spun, or the temperature and the composition of the bath, if wet-spun. The multitude of combinations made possible by the use of various comonomers and the flexibility of the fiber-forming operations furnish the different manufacturers with versatility and the users with a variety of acrylic fibers. Figure 12.20 shows a wet-spinning operation for acrylic tow.

Acrylic fibers possess a property that made it possible for them in the late 1950s and early 1960s to find immediate, even spectacular, acceptance in the knitted sweater field, until then dominated by wool. When acrylic fibers, normally in the form of a heavy tow, are hot stretched (e.g., by being drawn over a hot plate and then cooled under tension), they are converted to a labile state. Upon immersion in hot water, such fibers will contract considerably, but not to their prior unstretched length. In practice, this characteristic is used to produce a bulky yarn resembling the woolen yarns long accepted for use in sweaters. The process is described briefly below.

Using “stretch-break” equipment, the stretched labile fibers are further cold stretched to the breaking point so that the fiber breaks at different points leading to a distribution of fiber lengths, similar to the lengths found in wool. These are crimped and then mixed with thermally stable acrylic fibers that have been stretched and relaxed and have about the same length and degree of crimp. The blend is converted to a spun yarn by the same process used in making woolen yarns, and in turn, this yarn is knitted into sweaters and other similar products. When such garments are dyed in hot water, the labile fibers, intimately blended with stable ones, contract lengthwise individually. In the process, segments of the stable units tend to be carried along physically by entrapment and friction, but because such fibers do not change their overall length, the yarn as a whole decreases in length. Lateral displacement of the large volume of stable fibers results in the formation of a more voluminous structure known as “hi-bulk” yarn.

Bicomponent or Conjugate Spun Fibers

As will be shown, it should be theoretically possible to make any of the common manufactured fibers in bicomponent forms. However, acrylics have received the most attention for quite good reasons. Their general characteristics have tended to make them competitive with wool. This means that they should be processible on machinery developed for handling wool, as well as capable of being accepted into markets previously dominated by an animal hair fiber. It follows that because the natural fiber possesses crimp which produces the cohesion that determines its behavior in processing and in part its appearance and “hand” in usage, a similar crimp was desired for acrylics.

The principle that is the basis for bicomponent fibers usually is likened to that which underlies the bicomponent metal strips often used in temperature controllers. With the latter, differential thermal expansion of the two joined components results in a bending of the thermal element. With fibers, moisture usually is the agent that acts upon the two side-by-side portions. Differential swelling or shrinkage causes the fiber to be brought into a crimped, or preferably, a spirally distorted condition. As such, the side-by-side structure exists naturally in wool.

The combination of small size and large number of holes in a spinnerette might lead one to conclude that it would be almost impossible to design a spinnerette assembly that could bring two streams of polymer or polymer solutions together at each such hole and extrude them side by side to form a single filament. Such designs have, in fact, been made, but solutions of fiber-forming polymers fortunately possess properties that encourage laminar flow and thus make other approaches possible. This phenomenon was remarked upon earlier in connection with dope dyeing; when a suspension of a colored pigment is injected into a dope stream, a considerable problem must be overcome to achieve adequate mixing so as to secure “dope-dyed” fibers of a uniform color. Thus, it was known that when two streams of essentially the same solution of a fiber-forming polymer are brought together, side by side, and moved forward down a pipe or channel by the same amount of pressure behind each, virtually no mixing takes place. By bringing these streams to each spinnerette hole in such an individual side-by-side arrangement and using appropriate mechanical separators, the extruded filament from each hole will have a bicomponent structure. In addition to producing fibers in which the two components form a bilateral symmetrical structure, an ingenious arrangement of predividers of the two streams can produce from the full complement of holes in a single spinnerette a selected group of fibers wherein the amount and the position of each of the two components are randomly distributed throughout their cross sections. It follows that curls of uniform or random geometry may be produced to meet the required needs.

The worldwide production of acrylic fiber has declined significantly over the years because of the environmental concerns associated with the solution-spun process. In view of this, scientists have sought over the years a method that could render the high acrylics melt spinnable. Such a method would not only be economical and environmentally friendly but also allow for engineering the fiber with a wider range of morphologies and properties. In 1997, British Petroleum patented a polymerization process in which the two components usually used in developing spinnable acrylic copolymer were redistributed to allow the resulting material to be melt processable [13, 14]. Preliminary findings show that the polymer can be melt-spun into reasonable fine denier fibers with mechanical properties expected of the usual solution-spun material [15].

Vinyl and Modacrylic Fibers

Vinyls

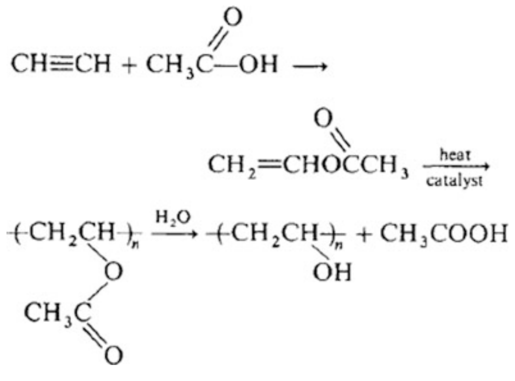
When nylon 66 was developed, it was described as being “synthetic” or “fully synthetic” in order to differentiate it from rayon and acetate. This was no small act of courage, as the word “synthetic,” in that period just following the repeal of Prohibition in the United States, was often associated in the public mind with the least palatable kind of alcoholic beverages. In due time, what is known in the advertising business as “puffing” led it to be known as the “first fully synthetic fiber,” which was an anachronism. It so happens that fibers based upon polyvinyl chloride (PVC) predated nylon by several years.

About 1931, the production of fibers from PVC was accomplished by dry spinning from a solution in cyclohexanone. But by chlorinating the polymer, it was possible to secure solubility in acetone, which has the advantage of possessing a boiling point about 100°C lower than that of cyclohexanone. Several million pounds per year of this fiber were produced in Germany during World War II to relieve the shortages of other materials. Unfortunately, PVC begins to soften at about 65°C, and in the fibrous state, it shrinks disastrously upon heating. Because of its low softening point, it cannot be dyed at the temperatures commonly used for this purpose, and, furthermore, it resists dyeing.

Modifications of PVC have been produced by copolymerization with other monomers. The first successful one consisted of 90% vinyl chloride copolymerized with 10% vinyl acetate. It was dry-spun from acetone and given the trade name Vinyon by its producer, Union Carbide Corporation. (In 1960, vinyon was accepted as a generic name for fibers containing not less than 85% vinyl chloride.) It has never been produced in large volume; it is used for heat-sealable compositions.

A copolymer of vinyl chloride with vinylidene chloride was used for a number of years to produce melt-spun, heavy monofilaments, which found use in heavy fabrics, where the chemical inertness of the polymer was needed, in outdoor furniture, and in upholstery for seats in public-transportation vehicles.

Another vinyl-based fiber, polyvinyl alcohol, or vinal, was developed in Japan but has not been produced or used in the United States. As such, it illustrates the importance of both relative availability of raw materials and differences in markets, in the success of a chemical product. Acetylene made from calcium carbide is converted to vinyl acetate, which, following polymerization, is saponified to polyvinyl alcohol.



The polyvinyl alcohol is soluble in hot water, and the solution is wet-spun into a coagulating bath consisting of a concentrated solution of sodium sulfate. The fibers are heat-treated to provide temporary stability so that they may be converted to the formal derivative by treatment with an aqueous solution of formaldehyde and sulfuric acid. This final product resists hydrolysis up to the boiling point of water. It seems reasonable to assume that it contains hemiacetal groups and some unreacted hydroxyls on the polymer chain, as well as cross-linking acetyl groups between the adjacent molecules.

Under the trade name Kuralon® (Kuraray Co., Ltd.), it achieved a production level of about 180 million lb in 1970, but production dropped to 16 million lb of continuous filament and 87 million lb of staple in 1980. The former has been mainly used in industrial rubber products, and the latter has been used mostly for uniforms, nonwoven and coated fabrics, and filters.

Modacrylics

In the United States, the modification of PVC has moved in the direction of copolymerizing vinyl chloride with acrylonitrile, or perhaps, it should be said that PAN has been modified by copolymerizing the acrylonitrile with chlorine-containing vinyl compounds. In any case, one modacrylic fiber is currently produced in the United States, a modacrylic being defined as containing at least 35% but not over 85% acrylonitrile.

The first two modacrylic fibers ever introduced in the United States were Dynel® (by Union Carbide) in 1949 and Verel® (by Tennessee Eastman Co.) in 1956. The former was a copolymer of 60% vinyl chloride and 40% acrylonitrile, and the latter was said to be a 50–50 copolymer of vinylidene chloride and acrylonitrile with perhaps a third component graft copolymerized onto the primary material to secure dyeability. SEF® and its version for wigs, Elura®, were introduced by Monsanto Fibers in 1972. A few foreign manufacturers are making modacrylic fibers, but the only modacrylic fiber currently in production in the United States is SEF®.

Modacrylic fibers, like acrylic, require after stretching and heat stabilization in order to develop the necessary properties. It is thought that the stretching is of the order of 900–1300% and that, in a separate operation, shrinkage of about 15–25% is allowed during the time that the fibers are heat stabilized.

The modacrylic fibers, like vinyon and unlike the acrylic fibers, have not become general-purpose fibers. They can be dyed satisfactorily and thus are acceptable in many normal textile products, but their nonflammability tends to place them in uses where that property is important, even vital. Blended with other fibers, they are used in carpets, but their largest market is in deep-pile products, such as “fake furs,” or in doll hair, where a fire hazard cannot be tolerated.

Elastomeric Fibers

The well-known elastic properties of natural rubber early led to processes for preparing it in forms that could be incorporated into fabrics for garments. One such process uses standard rubber technology. A raw rubber of high quality is compounded with sulfur and other necessary chemicals, calendered as a uniform thin sheet onto a large metal drum, and vulcanized under water. The resulting skin is spirally cut into strips that may be as narrow as they are thick, for example, 0.010 in. by 0.010 in. square in cross section. These strips are desulfurized, washed, dried, and packaged. Larger cross sections are easier to make. This product, coming out of the rubber rather than the textile industry, is known as a thread.

Another method produces a monofilament known as a latex thread. As the name would indicate, rubber latex is the raw material, and because extrusion through small holes is required, the purity of the material must be of a high order. With proper stabilization, the latex solution may be shipped from the rubber plantation to the plant, where it is compounded with sulfur and other chemicals needed for curing, as well as with pigments, antioxidants, and similar additives. This is followed by "precuring" to convert the latex to a form that will coagulate upon extrusion into a precipitating bath of dilute acetic acid and will form a filament having sufficient strength for subsequent operations. It passes out of the bath and is washed, dried, vulcanized in one or two stages, and packaged.

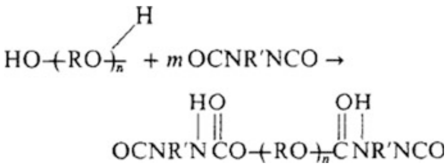
The rubber threads manufactured by either process can be used as such in combination with normal nonelastomeric yarns in fabrics made by weaving or knitting, but most of them, especially those made by a latex process, first are covered by a spiral winding of natural or manufactured yarns. Often, two layers are applied in opposite directions to minimize the effects of torque. Such coverings have two purposes. The first is to replace the less desirable "feel" of rubber on human skin by that of the more acceptable "hard" fiber. The second concerns the engineering of desired properties into the product to be woven or knitted into fabric. As an elastomeric material begins to recover from a state of high elongation, it supplies a high stress, but as it approaches its original unstretched condition, the stress drops to a very low order. When wound in an elongated state with a yarn having high initial modulus and strength, the elastomeric component cannot retract completely because its lateral expansion is limited, and jamming of the winding yarn occurs. Thus, the combination of such materials can be made to provide stretch and recovery characteristics needed for a broad spectrum of applications.

The traditional elastomeric threads have been subject to certain inherent limitations, however. The presence of unreacted double bonds makes them sensitive to oxidation, especially with exposure to the ultraviolet radiation of direct sunlight. They also have low resistance to laundry and household bleaches and dry-cleaning fluids.

During recent years, elastomeric yarns or threads have been used to impart comfort, fit, and shape retention to a variety of garments such as women's hosiery and swimwear. Such garments must be thin and highly effective per unit of weight. The materials of which they are composed must be compatible with these requirements. Thus, it was not unexpected that the producers of manufactured fibers, already eminently successful in meeting the needs of the marketplace, should look to the field of elastomeric fibers for new possibilities. Given the limitations of rubbers, both natural and synthetic, as well as the relationships between molecular structure and behavior of fiber-forming linear polymers, the scientists faced new challenges.

As an oversimplification, it can be said that within limits, a rubberlike material can be stretched relatively easily but reaches a state where crystallization tends to occur. The structure produced in this manner resists further extension, and the modulus rises sharply. In contrast to the conditions that occur when the manufactured fibers discussed earlier such as nylon or polyester are drawn to form fibers of stable geometry in the crystalline and oriented states, the crystalline state of the elastomeric fibers is labile unless the temperature is lowered materially. Thus, to improve on the chemical sensitivity of rubber, new approaches were necessary. The solution was found in developing linear block copolymers containing "soft" liquid-like sections that impart elasticity, connected with "hard" components that act as tie points to hold the structure together.

The soft, flexible, and low melting part is commonly an aliphatic polyether or a polyester with hydroxyl end groups and molecular weight in the range of 500–4,000. The hard portion is derived from an aromatic diisocyanate supplied in an amount that will react with both end groups of the polyether or polyester to form urethane groups. The product, an intermediate known as a prepolymer, is a thick liquid composed essentially of molecules carrying active isocyanate groups at each end. For example,



$$\mathrm{HO{+}RO{+}_{n}H} + m\,\mathrm{OCNR'NCO} \rightarrow \mathrm{OCNR'\overset{H}{\underset{}{N}}\overset{\overset{O}{\|}}{C}O{+}RO{+}_{n}\overset{\overset{O}{\|}}{C}\overset{H}{N}R'NCO}$$

where –(RO)– is an aliphatic polyether chain, R′ is one of several commonly available ring structures, *n* ~ 10–30, and *m* ~ 1.5–2.

The elastomeric polymer is obtained by "extending" the prepolymer through its reaction with short-chain diols such as butanediol or diamines such as ethylene diamine, thus completing the formation of hard groups between soft, flexible chains. When amines are used, the final step is typically done in a polar solvent such as dimethylacetamide. The conversion of these polymers into usable fibers may be accomplished by wet-, dry-, or melt-spinning operations, depending on the polymer. Additives to impart whiteness or improve resistance to ultraviolet radiation and oxidation may be incorporated in the spinning solutions or in the melts.

The development of elastomeric fibers has resulted in a variant of wet spinning called "reaction" or "chemical" spinning. In point of fact, rayon, the first wet-spun material, might properly be said to be produced by "reaction wet spinning" or "chemical wet spinning" because complex chemical reactions always have been involved in that operation. In any case, it has been found that the prepolymer of an elastomeric fiber may be extruded into a bath containing a highly reactive diamine so that the chemical conversion from liquid to solid occurs there.

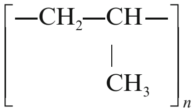
The elastomeric fibers produced in this fashion are based upon segmented polyurethanes and by definition are known generically as spandex yarns. Each manufacturer uses a trade name, for the usual commercial reasons. Perhaps the most noteworthy aspect from the standpoint of industrial chemistry is the multitude of options available to the manufacturer through the ingenious use of various chemicals for soft segments, hard units, chain extenders, and conditions of chemical reaction, followed by numerous possibilities for extrusion and aftertreatments. In the United States, there are two main producers of spandex fibers: DuPont (Lycra®) and Globe Rubber Co. (Cleerspan®, Glospan®). There are numerous worldwide producers, including Bayer, Germany (Dorlastan®); Asahi, Japan (Roica®); Nisshinbo, Japan (Mobilon®); and Taekwang, Korea (Acelan®).

Polyolefin Fibers

Polypropylene

Although polyethylene was considered a source of useful fibers at an early date, its low melting point (110–120°C) as well as other limitations precluded active development during the period when production of other fibers based upon the petrochemical industry expanded enormously. The higher melting point of high-density polyethylene (HDPE) gave some promise, but it was overshadowed by the introduction of polypropylene (PP) around 1958–1959. Great expectations were held for the latter as a quick competitor with the polyamides and the polyesters, already successful, as well as the acrylics, which then were entering the fiber field in volume. PP was thought to have several advantages. The raw material costs were low, only a few cents a pound; also, there was a high level of sophistication in the spinning and processing of fibers and a presumption that this would readily lead to the development of means for converting the polymer to fibers; and, finally, there was the belief that the American consumer would be ready to accept, and perhaps even demand, something new and different, which this polymer offered. However, the limitations of PP fibers, such as lack of dyeability, low melting temperature, low heat stability, and poor light stability, combined with the lower prices and the greater versatility of the already-established fibers, dashed the hopes for quick success. However, all of these deficiencies except the low melting temperature and lack of dyeability now have been overcome. The fiber has found an increasingly important place, and its properties have led to new techniques of manufacture and specialized uses

The structural formula of PP is as given below, where $100{,}000 < n < 600{,}000$ for chips or granules and $50{,}000 < n < 250{,}000$ for fibers:



$$\left[-CH_2-\underset{\displaystyle CH_3}{\underset{|}{CH}}- \right]_n$$

The steric configuration is extremely important in the polymer. Only isotactic polypropylene (iPP) has the properties necessary for forming fibers. The molecules are cross-linked only by van der Waals forces, so it is important that they pack as closely as possible. The isotactic molecules form a 3_1 helix, as shown in Fig. 12.21 [16], and exhibit a high crystallization rate.

The atactic molecules, shown in the figure, do not pack well, and although the syndiotactic molecules can pack better and crystallize, this configuration is not a normal product of commonly used catalyst systems. Some properties of isotactic, syndiotactic, and atactic PP are listed in Table 12.2 [17]. The insolubility of iPP in hydrocarbon solvents at room temperature often is used to separate iPP from atactic polypropylene (aPP). The insolubility of iPP in hydrocarbon solvents at room temperature often is used to separate iPP from aPP.

Early in the manufacture of PP, a concept was developed for dry spinning directly from the solution obtained in the polymerization operation. Had it been feasible, it would have been the realization of a chemical engineer's dream: the gaseous olefin fed into one end of the equipment and the packaged fiber, ready for shipment to a textile mill, coming out the other end. But it did not turn out that way, and today, melt spinning is the accepted technique for the production of staple fibers, monofilament, and multifilament yarns. To this usual method have been added the fibrillation and the "slit film" procedures for producing yarns.

The PP materials are completely resistant to bacterial attack, are chemically inert, and are unaffected by water.

Monofilaments can be produced that possess high strength, low elongation under stress, and dimensional stability at normal atmospheric temperatures.

PP monofilaments have found broad application in cordage and fishing nets (which float), and if highly stabilized, they are woven into fabrics used for outdoor furniture, tarpaulins, and similar applications. Large filament denier staple is used widely in "indoor–outdoor" carpets. Also, staple fibers have found major applications in tufted indoor carpets and nonwovens used for diaper, filtration, and civil engineering fabrics.

Synthesis. The early PP plants used a slurry process adopted from polyethylene technology. An inert liquid hydrocarbon diluent, such as hexane, was stirred in an autoclave at temperatures and pressures sufficient to keep 10–20% of the propylene monomer concentrated in the liquid phase. The traditional catalyst system was the crystalline, violet form of $TiCl_3$ and $AlCl(C_2H_5)_2$. Isotactic polymer particles that were formed remained in suspension and were removed as a 20–40% solid slurry while the atactic portion remained as a solution in the liquid hydrocarbon. The catalyst was deactivated and solubilized by adding HCl and alcohol. The iPP was removed by centrifugation, filtration, or aqueous extraction, and the atactic portion was recovered by evaporation of the solvent. The first plants were inefficient because of low catalyst productivity and low crystalline yields. With some modifications to the catalyst system, basically the same process is in use today.

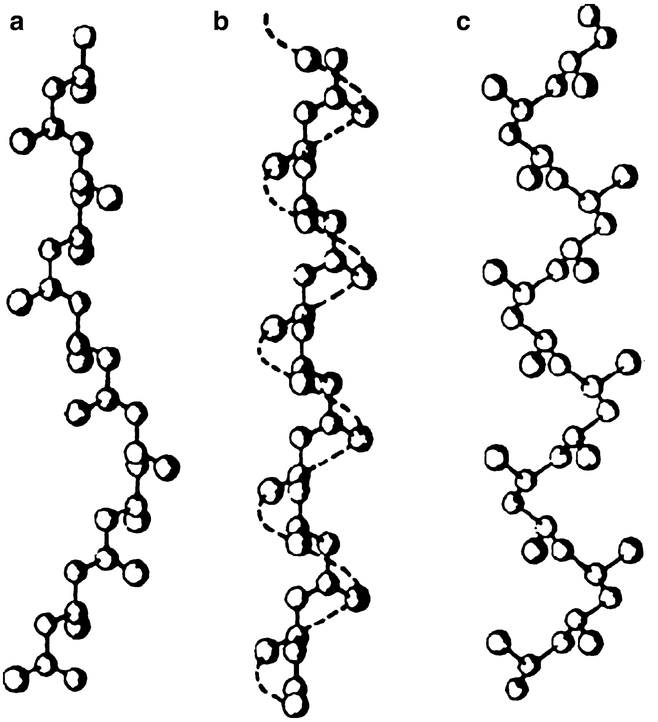


Fig. 12.21 Diagrams showing (**a**) irregular atactic, (**b**) stereoregular isotactic, and (**c**) stereoregular syndiotactic configurations in polypropylene polymer (*Source*: Ahmed [16]) (Copyright © M. Ahmed with permission)

In 1963, liquid polymerization was introduced in which liquid propylene, catalysts, and hydrogen were pumped continuously into the reactor, while polypropylene slurry was transferred to a cyclone separator. The unconverted monomer gas was removed, compressed, condensed, and recycled, and the polymer was treated to reduce the catalyst residue. This system also suffered from a poor catalyst yield, and the polymer produced lacked the required stereospecificity, so that it was necessary to remove the atactic portion of the polymer.

In the mid-1960s, a gas-phase process was introduced for production of the polymer. The monomer was pumped over adsorbing beds and entered the reactor with the catalyst system. These feed streams of monomer and catalyst, together with a mechanical stirrer, created a turbulent bed of powdered polymer. Periodically, the polymer powder was vented off in a carrier gas to extrusion storage hoppers. Meanwhile, the heat of polymerization was removed by condensing the unreacted monomer in a cooling loop and returning it to the reactor, where it immediately vaporized. This process eventually led to the production of highly crystalline products and was adopted by several companies in the United States.

Most processes in use today rely on a combination of these technologies. Montedison's introduction in 1975 of third-generation catalysts gave high yields and allowed polymerization to take place at 60–80°C and 2.5–3.5 MPa (362–507 psi). This was welcome news during the energy crisis, but the resulting polymer was not stereospecific enough to eliminate the need for removal of aPP. Real progress came with the discovery of superactive

Table 12.2 Properties of isotactic, syndiotactic, and atactic polypropylene

Property	Isotactic	Syndiotactic	Atactic
Density (g/cm^3)	0.92–0.94	0.89–0.91	0.85–0.90
Melting point (°C)	165	135	
Solubility in hydrocarbons at 20 °C	None	Medium	High
Yield strength	High	Medium	Very low

Source: Lieberman and Barbe [17] (Copyright © John Wiley and Sons and reproduced by permission of the copyright owner) third-generation catalysts, which gave both the optimal yield and stereospecificity [18].

Production. Classical melt spinning, which was developed for the production of nylon filaments, is widely used to produce PP fibers today. It involves a high-speed process (2,000–3,000 m/min) that is particularly suitable for long production runs. The average molecular weight of polypropylene polymer, like that of other addition (olefin) polymers, is relatively high compared to that of other polymers. This results in a high melt viscosity, so, unlike the case of other polymers, its extrusion temperatures are 70–100°C above its melting point. Single-screw extruders are used for melting and homogenizing the polymer. The screw diameters are from 45 to 200 mm, and screw lengths are 24, 30, or 36 times the diameter. The polymer granules are fed into the extruder hopper, where they are melted and homogenized. Chips carrying pigments can be fed into and blended with the main charge of the extruder if colored fibers are desired. The molten polymer is forced through the spinnerette via a screen pack to eliminate any contaminant particles. The spinnerette hole determines the shape of the filaments, and the flow rate and the take-up speed determine the size. The polymer has high specific heat and low thermal conductivity, so the cooling zones must be longer than those for polyester and nylon. For filament yarn production, the filaments are drawn at high speed and wound on packages. For staple fiber production, the filaments are collected in the form of tow and then are drawn, crimped, and cut. The multifilament yarns are often textured to improve bulk and appearance. The false twist method is generally used to texture finer yarns, with the stuffer box used for coarser yarns.

The short spinning method used to produce staple fiber is considerably slower than high-speed spinning. The lower spinning speeds (30–150 m/min) would have a negative effect on productivity, but this is counteracted in industry by the use of spinnerettes with a large number of holes (up to 55,000). The required cooling zones are much shorter because of the lower speeds and the use of higher volumes of quench air, which gives this method its name. Because the drawing units can match these low speeds, the two can be fed directly and continuously from the spinning machine to the draw frame to the texturing chamber or the crimper to produce bulked continuous-filament (BCF) yarn or staple fiber, respectively. The short spinning method is used to produce high-tenacity fibers. Many PP yarns are produced using the slit film method. The film extruder is almost identical to a filament extruder. In it, the molten polymer is forced through a film die that converts the melt into film, where the thickness of the film can be controlled by adjusting one of the die lips. The takeoff unit is either a chill roll that removes the film uniformly and cools it below T_m or a water bath followed by nip rolls. The cooled film is slit into separate tapes using a slitter bar that contains a large number of special knives separated by spacers. The film tapes are heated and drawn to their final length.

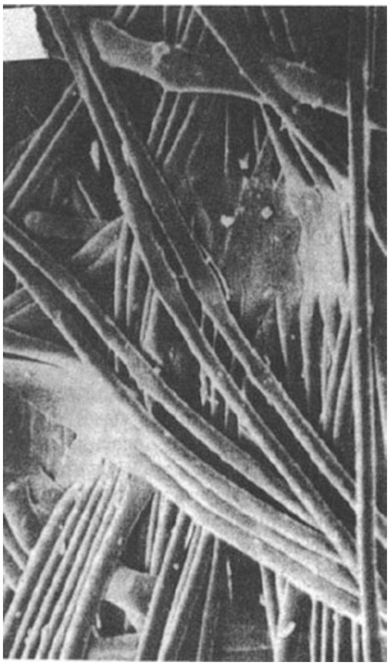


Fig. 12.22 Spunbonded polypropylene showing interfiber bonding, which binds the structure (Courtesy of E.I. du Pont de Nemours and Co.)

Several methods are used to produce fibrillated film. They are produced in much the same way as slit film, but these techniques take advantage of the tendency of PP to fibrillate. In one method, a profiled tape is extruded and hot-drawn. It is drawn again to achieve a 10:1 ratio, and the film splits into separate filaments. In another method, called roll embossing, the film is hot-drawn, and then embossed using profiled rollers. The profiled film is drawn again, and fibrillation results. In the pin-roller technique, the film is drawn and cut by knives or pins on a rotating cylinder. This method can produce individual fibers or a controlled web network.

PP nonwovens are created by forming a staple fiber web and then consolidating the latter. The fibers can be entangled by a needling machine, an air jet, an air jet, or a water jet. Another method uses a single-stage process in which melt-spun fibers are drawn through an air aspirator jet and deposited randomly on a conveyor. The fibers then are bonded by fusion under heat and pressure to give a "spunbonded" nonwoven fabric (see Fig. 12.22).

Yet another method, known as melt blowing, can be used in producing a nonwoven web having special characteristics. The melt-blown (MB) fibers are characterized as ultrafine fibers because of their size relative to other fibers. In the process, molten polymer is forced through a melt-blowing

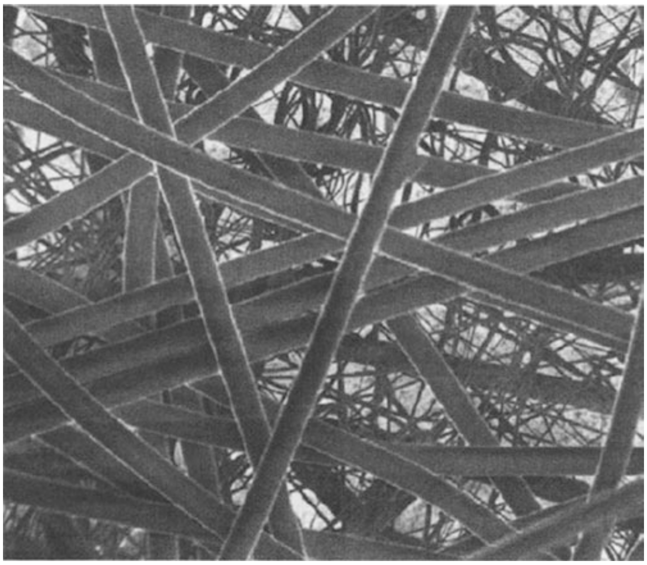


Fig. 12.23 Spunbond/melt-blown composite, SMMS, containing two layers of melt-blown microfibers sandwiched between two layers of spunbonded material and bonded by hot calendaring (*Source*: Madsen [20])

die and die-tip orifice that are finer than most orifices used to make manufactured fibers; it is attenuated by a jet of high-velocity hot air. This allows the polymer to stay in a molten state but attenuate until broken. The fibers then come in contact with cool quench air, which causes the former to solidify. The fibers deposit on a collector screen and form a MB web. These webs, composed of very fine-sized fibers, can be engineered for applications in areas where barrier (such as against bacteria), insulative, and absorptive (such as for oil) characteristics are important.

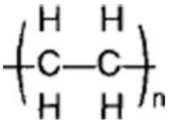
One of the most recent developments in the use of polyolefin fibers is in composites of spunbond (SB) and melt-blown systems. Some examples of structures being made are SB/MB, known as SM, SB/MB/SB, or SMS, and other combinations of SB and MB. The production and properties of these are particularly enhanced by the use of PP/polyethylene bicomponent fibers in the preparation of the MB layers. Such to composites are particularly suited for use as lightweight coversheets or facings for diapers and sanitary products. The spunbond layers on the top and bottom provide the needed high strength, whereas the layers of MB microfibers sandwiched between them provide the greatly enhanced visual uniformity and barrier properties [19]. Figure 12.23 illustrates a low-weight (15 g/m^2) SMMS fabric comprised of MB fibers accounting for less than 15% of the structure by weight [20].

The production figures of PP show impressive growth. The volume, which was less than 10 million lb in 1965, reached over 1.5 billion lb in 1990. This included approximately 17% filament yarn, 23% staple fiber, 18% spunbonded fabrics, and 38% split film products.

It is recognized that despite its many advantages compared with other synthetic fibers, the use of the fiber is restricted by the fact that most dyestuffs lack affinity for the material. This makes the fiber unsuitable for conventional exhaustion dyeing techniques using disperse or acid dyes. A number of attempts have been made in recent years to improve dyeability of the material [21]. In one of the most recent studies, an alloy fiber of polypropylene (PP) and nylon 6 (N6) has been produced by melt spinning. The fiber contained PP grafted with maleic anhydride (MAH) as the third component. The MAH units reacted with the amine end groups of N6 to form block or graft copolymers that acted as a compatibilizer for the blend. Addition of 10% N6 produced a fiber that could be dyed conventionally using disperse dyes [22].

High-Molecular-Weight Polyethylene

Polyethylene is probably the material one encounters most abundantly in daily life, such as in the form of grocery bags, shampoo bottles, and toys, but now, one also finds it as the material in such superhigh-performance products as bullet-proof vests. For such a versatile fibrous material, it has the simplest structure of all known polymers:



$$\left[\begin{array}{cc} \mathrm{H} & \mathrm{H} \\ | & | \\ \mathrm{C} - & \mathrm{C} \\ | & | \\ \mathrm{H} & \mathrm{H} \end{array} \right]_n$$

However, during the general addition polymerization process, some of the carbon atoms, instead of having hydrogen attached to them, have segments of polyethylene chains grown on them, leading to a branched or low-density polyethylene (LDPE). If branching is eliminated or greatly minimized, such as found with the use of the Ziegler–Natta polymerization process, utilizing special catalysts, one can get linear chains capable of greater packing. This provides an HDPE material suitable for fiber use. Molecular weights of the order of one million have been achieved resulting in ultrahigh-molecular-weight polymer (UHMWPE). Accordingly, a recent addition to the group of high-performance fibers is the ultrahigh-molecular-weight, extended-chain, linear polyethylene. Although fundamental work in the area of developing fully oriented and crystalline structures in polyethylene polymer had been going on since the mid-1960s, it was not until the late 1970s that the possibility of producing such materials on a commercial scale became evident. Presently, three companies worldwide are manufacturing extended-chain polyethylene fibers, which have very high-strength and high-modulus characteristics.

Routes to High Performance. The achievement of ultimate strength and high modulus in fibers has been a subject of great interest to material scientists and fiber producers. In an attempt to identify the preferred structures for high performance, many theoretical analyses have been conducted to calculate the limiting values. The basis of such calculations is the assumption that in a fiber in which all polymer chains of infinitely long dimensions are extended and oriented parallel to the axis, rupture will occur only when the stress exerted exceeds the intramolecular bond strength. Such calculations show that the limiting theoretical values for conventional fibers usually are several times higher than those obtained in actual practice. The reasons for this difference lie in one or more of the following explanations: (1) the molecular weights are not high enough, (2) the chains are not fully extended, and (3) the chains are not fully oriented. For a given polymer system and molecular weight, tensile strength and modulus can be enhanced by extending and orienting the chains. Practically, many modern techniques, including zone stretching, multiple-step stretching, and state-of-the-art high-speed extrusion methods, have been adopted to achieve such results.

Still, the results of studies on conventional fibers show that modifications in physical processing alone cannot lead to values that even approach the theoretical maximum to within an order of magnitude. Flexible chains, which characterize the bulk of the commercial polymers, tend to conform to a random coil or folded chain structure in an as-spun material, and are very difficult, to reorganize into an extended-chain structure by known methods. Accordingly, in order to achieve ultimate properties, either novel spinning methods are needed, which allow flexible chains to be fully extended, or the chains used must be so configured that they have high intrinsic stiffness and would remain extended in solution or melt. Both of these goals have been met and are exemplified in the production of (1) high-strength polyethylene fibers, discussed in this section, and (2) aramid fibers, discussed in the next section.

Extended-chain polyethylene fiber became available commercially in 1984 when DSM, a Dutch firm in the Netherlands, introduced Dyneema® and Mitsui Petrochemicals in Japan announced Tekmilon®. AlliedSignal of the United States entered the field in 1985 when it introduced Spectra® fibers. These materials are characterized by very high strength and modulus, which are achieved by the use of ultrahigh-molecular-weight polyethylene spun by the gel-spinning method into fibers having extended-chain structures and near-perfect orientation.

Gel Spinning. In general, the purpose for which the gel-spinning method is used is to produce an as-spun fiber that contains a loose network of chains with few entanglements, which then can be drawn out to ultrahigh levels to yield a highly oriented structure. The surface-growth method of Pennings [23], which uses a Couette-type apparatus, and the gel-spinning method of Smith–Lemstra [24], which uses a more conventional spinning apparatus, led to the achievement of this goal. In the former method, a polymer solution is stirred between two counterrotating cylinders that provides the elongational flow necessary for initial chain alignment. Essentially the same result is achieved by passing the polymer solution through a constriction prior to spinning in the latter method.

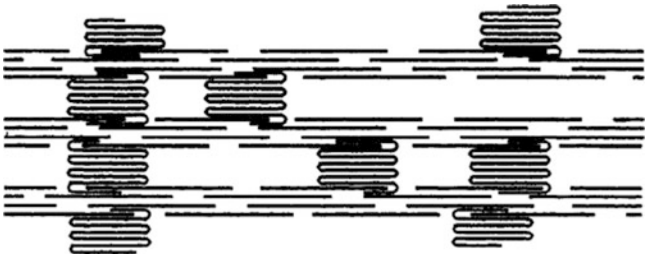


Fig. 12.24 "Shish-kebab" structure showing extended-chain crystals with lamellar overgrowths

In the surface-growth technique, polymer solution between the inner rotating cylinder and the outer stationary cylinder is maintained within a certain temperature range above the polymer crystallization temperature. A fiber seed is immersed through an opening in the outer cylinder, and its tip is made to attach to the polymer layers absorbed on the inner rotating cylinder. By pulling on the other end of the fiber and winding it onto a bobbin, while at the same time replenishing the solution in the gap between the cylinders, a fiber can be produced continuously. This process results in a fibrous precipitate with a "shish-kebab" morphology (Fig. 12.24).

One can conclude, then, that the flow field that is developed extends the chains, which then crystallize in a fibrillar form. Chains that are left unextended, those that are below a certain length for the speed, use the preformed clusters (shishes) as nuclei and crystallize as overgrowths in the form of chain-folded lamellae (kebabs). Fibers formed in this way showed remarkable mechanical properties with modulus in excess of 1,200 g force/denier (102 GPa) and tensile strength in the neighborhood of 30 g force/denier (2.6 GPa).

Production. Recognition that the shish-kebab fibers produced by the surface-growth procedure result from the deformation of a gel-like entangled network layer at the rotor surface led to the development of gel-spun polyethylene fibers. The fiber is made by the solution spinning method. The polymer is dissolved in a high-temperature solvent, such as decalin, at a low concentration, 10% or less, and extruded into cold water. Initially, the fiber is formed as a gel that still contains the spinning solvent.

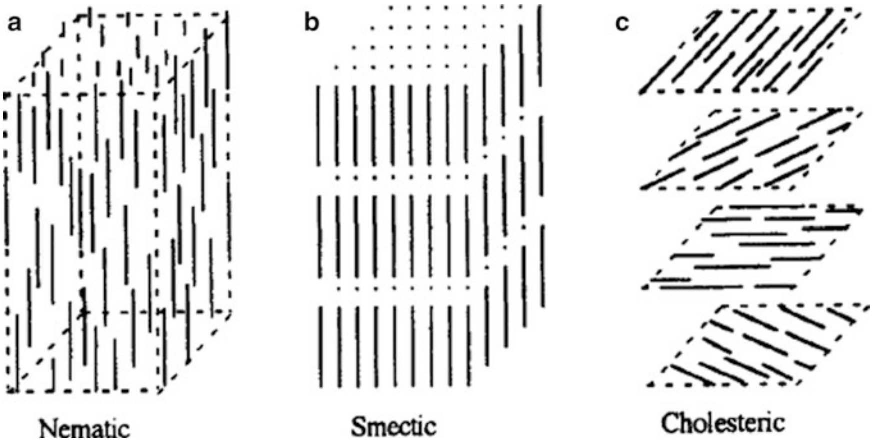


Fig. 12.25 Structure of liquid crystalline polymers showing (**a**) nematic, (**b**) smectic, and (**c**) cholesteric orders

It can be oriented by drawing at remarkably high draw ratios of 30 or greater, either before or after solvent removal. A method of spinning polyethylene fiber is described in the patent by Kavesh and Prevorsek [25]. The factors that govern the properties of fibers produced by this method are the polymer molecular weight, the concentration of polymer in the solvent, the type of solvent, the solution temperature, and the stage and the extent of drawing.

The high-performance behavior in these fibers is obtained from their having a very high molecular weight (1–5 million) combined with a very high degree of extended-chain crystal continuity. The fiber can approach a 100% crystalline structure with a theoretical maximum density (~0.97 g/cc). Their use, of course, is limited by the melting point of the polymer, which even in the extended form is only about 150°C. Although higher than that of ordinary polyethylene, it is still much lower (80°C or more) than those of commercial textile and other higher-performance industrial fibers. There also has been concern about the creep that occurs in these fibers, although significant improvement has been made in this regard since the introduction of the first fiber. In spite of the limitations mentioned, the unique combination of strength, stiffness and lightness properties have allowed these fibers to prove highly successful in a number of applications, such as sailcloth, body armor, medical implants, fishing net, and sports equipment.

Aramids

Introduction

As pointed out in the preceding section, a second route for developing fibers having properties approaching the ultimate values is the use of polymer chains that have high intrinsic stiffness and will remain extended in solution or melt. The development of aramid organic fibers based on aromatic polyamides met these requirements and added another chapter to the history of the development of synthetic fibers. Nomex® aramid, a thermally resistant fiber based on a *meta*-oriented structure, was commercialized by the DuPont company in 1962.

Following the technological breakthroughs which led to the discovery of (1) the liquid crystalline behavior of *para*-oriented aramids [26] and (2) a novel method for spinning anisotropic liquid crystalline polymer solutions [27], Kevlar® aramid fiber was produced and commercialized by the DuPont company in 1972. Other fibers based on aromatic polyamide compositions, which were produced and commercialized by other companies, were Technora® (Teijin, Japan), Teijinconex® (Teijin, Japan), and Twaron® (Akzo, The Netherlands). Additionally, SVM is a fiber produced in the former Soviet Union, and it was announced in 1990 that a new aramid fiber had been introduced by Hoechst, in Germany.

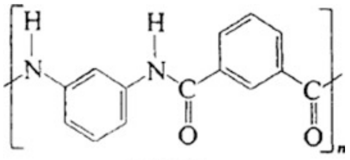
The rigid rod chains in *para*-aramids tend to form so-called liquid crystals when they are dissolved in polar solvents or heated to certain temperatures. The polymers showing liquid crystalline behavior in melts are called the thermotropic liquid crystalline polymers, and those showing similar properties in solution are called the lyotropic liquid crystalline polymers. These liquid crystals exist in three distinct phases according to their specific structures (Fig. 12.25).

A one-dimensional arrangement of rigid rods is called a nematic structure; a two-dimensional ordered arrangement represents a smectic structure; and the cholesteric structure is formed in terms of rotating oriented sheets so that rigid rods align parallel to each other in every layer, but the directional vectors in each layer are different. The preferential phase considered for fiber spinning is nematic. Nematic solutions or melts are easy to develop into oriented structures through shearing and elongational flow during extrusion because of the rigidity of the polymer chains.

Thus high orientation can be obtained in the as-spun fibers without much posttreatment.

Manufacture

In 1973, the Federal Trade Commission recognized aramid as a distinctly different generic material and defined it as "a long-chain synthetic polyamide in which at least 85% of the amide linkages are attached directly to two aromatic rings." This distinguishes aramids from nylon, which was redefined as a polyamide with less than 85% of the amide linkages attached to two aromatic rings. The first aramid fiber produced in the United States was Nomex®, the reactants being *m*-phenylene diamine and isophthaloyl chloride to give poly (*m*-phenylene-isophthalamide) (MPD-I).

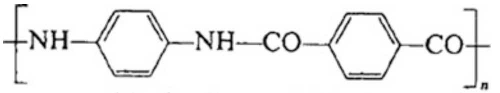


MPD-1

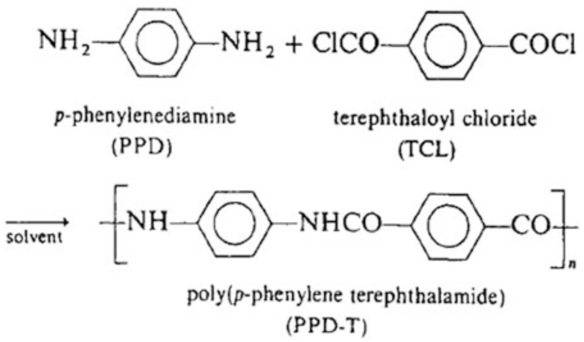
This polymer could not be melted without decomposition, so the preferred fiber formation route was solution spinning. Patent literature suggests that the fiber is spun from a solvent system composed of dimethylformamide and lithium chloride. The final properties are achieved by stretching in steam after washing to remove residual solvent.

The physical and chemical properties of this fiber are not remarkably different from those of other strong polyamides, but it does have excellent heat and flame resistance that makes it particularly suited for use in protective clothing and in specified industrial end uses. Military flight suits, firefighter uniforms, and hot gas filtration are a few of its many possible applications.

The other important fiber in this category, which also was first produced commercially in the United States, is Kevlar®, introduced in 1971 as fiber B and later coded as Kevlar® 29. It was produced from poly(*p*-phenylene terephthalamide) (PPD-T).



poly(*p*-phenylene terephthalamide)



Later, a higher modulus version, Kevlar® 49, believed to be made by heat annealing of Kevlar® 29, was introduced.

Poly(*p*-phenylene terephthalamide) PPD-T, also called *para*-aramid, can be polymerized to a fiber-forming molecular weight by polycondensation of terephthaloyl chloride and 1,4-phenylene diamine.

Poly(1,4-phenylene terephthalamide) (PPD-T) of high molecular weight can be prepared by low-temperature solution polymerization techniques. This polymer is less soluble in amide-type solvents than is poly(*p*-benzamide). The most successful conditions required hexamethylene phosphoramide (HMPA), the original solvent, alone or mixed with *N*-methyl-2-pyrrolidone (NMP), although other mixtures such as NMP containing $CaCl_2$ also could produce a fiber-forming polymer. During polymerization, the molecular weight increases rapidly within the first few seconds of the reaction. The critical molecular weight or viscosity is reached, and the stir opalescence typical of lyotropic solutions is observed. Although gelation of the reaction mixture occurs quickly, polymerization continues but at a greatly reduced rate. With the choice of a suitable solvent system, gelation can be delayed until the desired higher molecular weight is reached. Large-scale manufacture of a polymer requires continuous polymerization to minimize cost. In the case of PPD-T, special problems that had to be accommodated included the rapid gelation of the reaction solution with increasing molecular weight, the need to control the temperature of a vigorous exothermic reaction, and the handling of solvent HMPA, which is suspected to be a carcinogen.

Although melt spinning would be preferred from the standpoint of process simplicity and conversion cost, aramids must be spun from solutions, by wet, dry, or dry jet–wet methods because they decompose before or during melting, ruling out melt spinning. Dry spinning is used to produce Nomex® fiber, where a dope (20%) of the polymer in solvent is converted to yarn. In wet spinning, the

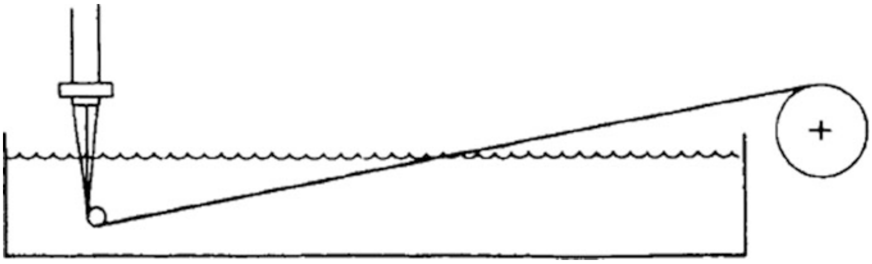


Fig. 12.26 Schematic of dry jet–wet spinning method

polymer dope is extruded into a nonsolvent where the fiber coagulates. The coagulated fiber then is washed and often drawn to develop desired fiber properties. In wet spinning, the spinnerette is in the coagulation bath, hence the dope temperature and the coagulation temperature are the same.

In dry jet–wet spinning (Fig. 12.26), as used for Kevlar®, on the other hand, the spinnerette is separated from the coagulation bath, allowing independent control of the dope extrusion and coagulation temperatures. The extrusion jet is placed a small distance above the coagulation bath, and the nascent fibers descend into the liquid, pass under a guide, and proceed in the bath while undergoing stretch; then they are withdrawn from the bath and wound up. A subsequent washing step may be required to remove residual acid solvent. In systems where the dope must be hot to lower viscosities to extrudable levels, and the bath cold for developing specific structures during coagulation, as is preferred with Kevlar®, dry jet–wet spinning is the option. For production of Kevlar® fibers, a PPD-T/concentrated sulfuric acid solution, containing approximately 20% polymer, is extruded at 90°C into a cold water bath (~1°C). Following extrusion, the fiber is washed, dried, and given posttreatment, depending upon the properties desired. Kevlar® is reported to have about twice the breaking strength (23–27 g force/denier or 2.9–3.4 GPa) of high-tenacity nylon and polyester, but its most outstanding physical property is its high stiffness (550–1,000 g force/denier or 70–127 GPa), which can range up to an order of magnitude greater than that of standard polyester. This property has led to high-volume usage of the fiber as reinforcement in composite materials such as belts in radial tires and aerospace structures. The world production figure (1999) of this fiber was estimated to be over 60 million lb.

High-Temperature-Resistant Fibers

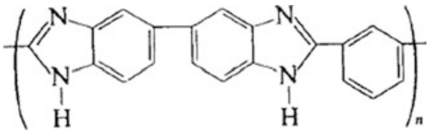
The need for high-temperature-resistant fibers has arisen from demands of a number of industrial applications, as well as applications in aerospace programs. In many of these applications, the usual characteristics of organic-based fibers are desired, but the high temperature resistance of inorganic fibers is required. Thus, the fibers are expected to retain their structural integrity at temperatures of 300°C and above for considerable periods of time, while their other physical and mechanical properties resemble those of the more common manufactured textile fibers. The two leading groups of fibers in this area are the *meta*-aramids and PBI.

Meta-Aramid

The major fibers in this group, based on sales volume, are the products made from poly(*m*-phenylene-isophthalamide), which were introduced by DuPont in 1962 as HT-1 nylon (later known as Nomex®) and by Teijin in 1972 as Conex®. General manufacture of the fiber was described earlier under the heading "Aramids." Other manufacturers now are entering the field with products of similar chemical structure. Although these products are made and spun by different processes, their chemical and physical properties are similar. The fiber is usually utilized in the form of cut staple, which is amenable to conversion on traditional spinning and weaving machinery. Its mechanical properties (tenacity 4–5 g force/denier, ultimate strain 25–30%, and initial modulus 90–100 g force/denier) compare favorably with those of other textile fibers. The main utility of the fiber lies in the resistance it offers to combustion: it has a limiting oxygen index (LOI) of about 0.29, a melting point above decomposition temperature, an ignition point above 600°C, and a flash point about 800°C [28]. The LOI gives a relative measure of flame resistance; the higher the number, the lower the flammability. The fiber Kermel® from Rhone Poulenc also is classified as an aramid. It is, chemically, a polyamide-imide fiber and has an LOI of about 0.31. A major application of these fibers is in protective clothing. In order to reduce their cost, they often are blended with other, less expensive, flame-retardant fibers such as those based on cotton, rayon, and wool.

PBI

Poly-2-2′-(*m*-phenylene)-5,5′-bibenzimidazole, commonly called PBI, was developed under the aegis of the U.S. Air Force Materials Laboratory in cooperation with the then-existing Celanese Corporation. The fiber went into commercial production in the United States in 1983. It is a condensation polymer obtained from the reaction of tetraaminobiphenyl and diphenylisophthalate in a nitrogen atmosphere at temperatures that may reach 400°C in the final stages [29]. The structure of a repeating unit is shown below:



The polymer is dissolved at a high temperature under nitrogen pressure in dimethylacetamide, to which a small amount of lithium chloride may be added to increase the stability of the solution. Then, it is dry-spun in an atmosphere of heated nitrogen (about 200°C), from which the solvent is recovered; next, it is stretched slightly in steam and washed. Drawing and relaxing are done in an inert atmosphere, as might be expected, because temperatures up to 250°C or higher are used. The fiber then is given a stabilization treatment in a sequence of steps and made into staple fibers using conventional crimping and cutting techniques. The stabilization treatment involves reaction with sulfuric acid and heating at high temperatures (~475°C) for short periods of time. The process, known as sulfonation, yields a product that has a significantly lower shrinkage than the unstabilized material.

The final yarn is golden yellow, and because this color appears to be an intrinsic property of the polymer, it may have some limitation as far as the civilian market is concerned. This material originally suffered from high shrinkage on exposure to flame; however, further developments including the sulfonation treatment have reduced the shrinkage to only about 5–10% at 600°C [30]. The fiber is capable of retaining about one half of its original strength (~3 g force/denier) upon exposure to air for 18 h at 350°C or 1 h at 425°C; it has an LOI of 0.41, which is well above that of the aramids; and it has high resistance to inorganic acids and bases and organic chemicals. Further, with tensile and moisture regain properties comparable to those of many textile fibers, the PBI fiber is well suited for blending with other fibers and conversion into final products using conventional spinning and weaving or nonwoven equipment. Applications include high-performance protective apparel, flight suits, and aircraft furnishings.

Polytetrafluoroethylene

Historical

Polytetrafluoroethylene (PTFE) was discovered in 1938 when Dr. Roy J. Plunkett and his assistants working on new nontoxic, nonflammable refrigerants at E. I. du Pont de Nemours and Co. found that one cylinder, which was supposed to contain tetrafluoroethylene (TFE), ceased to release the gaseous material. Upon opening the cylinder, they discovered that the inside was covered with a white powder. Polymerization of TFE had taken place. The result was PTFE or Teflon®, which is the trade name applied to that polymer by DuPont. The company tested the polymer and found that it was virtually inert to all known solvents, acids, and bases, a characteristic that was unique at that time. It also found that this material was resistant to high temperature and had the lowest coefficient of friction of any known solid. The research on PTFE was intensified during World War II because of military demands for improved materials for products such as gaskets, packing, and linings for containers for handling corrosive materials. When the war was over, the release of classified information caused a booming interest in the polymer. Since that time, PTFE has found itself in many different applications.

Manufacture

In one process, the manufacture of the monomer, TFE, involves the following reactions: hydrogen fluoride is made by reacting calcium fluoride with sulfuric acid:

$$CaF_2 + H_2SO_4 \rightarrow CaSO_4 + 2HF$$

Chloroform is reacted with hydrogen fluoride in the presence of antimony trifluoride as the catalyst:

$$CHCl_3 + 2HF \rightarrow CHClF_2 + 2HCl$$

TFE is obtained by the thermal decomposition of this monochlorodifluoromethane (known as Freon) in a continuous noncatalytic gas-phase reaction, carried out at or below atmospheric pressure at temperatures from 600 to 900°C:

$$2CHClF_2 \rightarrow C_2F_4 + 2HCl$$

Numerous side products are generated in this process. Many of them are present in trace amounts, but the highly toxic perfluoroisobutylene, $CF_2{=}C(CF_3)_2$, requires special precautions.

TFE also may be manufactured by the reaction of zinc and tetrafluorodichloroethane

$$ClF_2C - CF_2Cl + Zn \rightarrow C_2F_4 + ZnCl_2$$

or by the reaction of tetrafluoromethane molecules in an electric arc furnace

$$2CF_4 \rightarrow C_2F_4 + 2F_2$$

TFE is a colorless, tasteless, odorless, and nontoxic gas. To avoid any undesired reactions during storage, inhibitors must be added. The polymerization is carried out by an addition-type reaction in an aqueous emulsion medium and in the presence of initiators such as benzoyl peroxide, hydrogen peroxide, and persulfates. The monomer is fed into a cooled emulsion medium and then heated to a temperature of 70–80°C, at which the polymerization takes place:

$$CF_2 = CF_2 \rightarrow \left(CF_2 - CF_2 \right)_n$$

The pressure may range from 40 to 100 atmospheres. After removal of the unreacted material, the polymer is washed, pressed, and dried. The degree of polymerization can be quite high, of the order of 50,000. PTFE is manufactured in four different forms, granular, fine powder, aqueous dispersion, and micropowder, and in a variety of grades, each differing in properties. Fillers such as glass fibers, asbestos, graphite, or powdered metals may be added to the granules in order to modify properties.

Fiber Manufacture

Because Teflon® is not soluble, it cannot be wet- or dry-spun, and because it is thermally unstable at its melting point of about 400°C, this combination would seem to pose an impossible problem for the production of fibers. Research into the fundamental characteristics of the polymer, however, revealed that the submicroscopic particles precipitated from the polymerization reaction were about 100 times as long as they were thick.

In one manufacturing process, an aqueous dispersion of PTFE is mixed with a solution of ripened cellulose xanthate, from which a fiber is obtained by a wet-spinning process, after which the cellulose is completely decomposed by heating. The remaining PTFE is sintered into continuous fibers by transporting them over heated metal rolls, followed by stretching to achieve the desired diameters and physical structure. In a process known as “paste extrusion,” the powder is mixed with an organic plasticizer and compressed at 300–500 psi to make a preform or billet. The latter is extruded into filaments which are then dried to evaporate the lubricant. In another process, films are produced that are slit into strips of very small widths, which then are stretched and sintered. In yet another, tapes are extruded, which are converted to fibers by stretching while being twisted to a very small cross-sectional area.

Properties

The unique combination of properties, including chemical resistance, thermal oxidation resistance, high lubricity, electrical and thermal insulation, low flammability, and excellent weatherability, is derived primarily from two factors, namely, the molecular structure and the molecular weight of PTFE. The structure consists of a core chain of carbon atoms with a fluorine sheath, which essentially completely shields the core. The fluorine atoms are so tightly packed that steric interactions cause a slight rotation of the carbon chain from the normal planar zigzag to a helical conformation. This dense shield has a low surface energy and a very smooth surface with no side chains or imperfections. As a result, the interchain forces are low, and individual molecules are able to slide past each other with relative ease. Compared with other polymers, therefore, creep tends to be high.

The polymer chains pack themselves very closely and regularly to give cylindrical packing which consequently leads to very high crystallinity (~90%). The material has a high melting point which is about 330°C. The smooth surface when combined with low surface energy makes the surface so neutral that it resists sticking to any material. Accordingly, PTFE has the unique property of the lowest coefficient of friction (~0.007), with essentially no stick–slip character. The material has no affinity for water and, therefore, is totally hydrophobic.

The chemical inertness and the thermal stability of this polymer are so great that, in spite of its high price ($4–10/lb, depending on the resin type), it is used in chemical operations where drastic conditions exist and no other organic material is suitable. Its low friction allows it to be used as nonstick coating for metals, work surfaces, and cooking utensils. The polymer is regarded as biocompatible and tends to be well accepted by the body. This has allowed PTFE to be explored as a material for surgical implants.

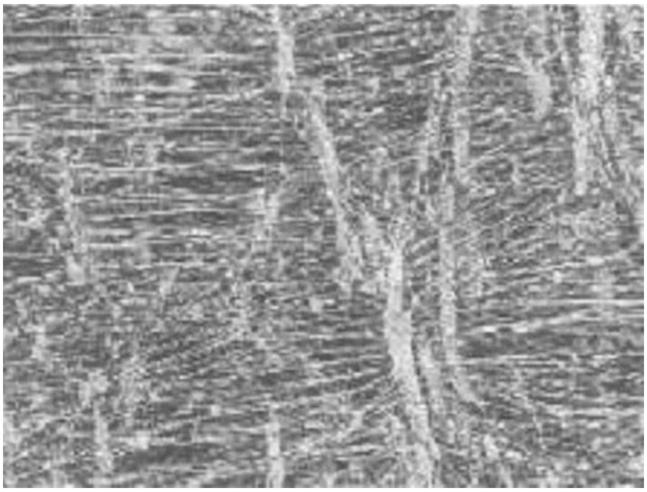


Fig. 12.27 Expanded PTFE micrograph (Microtex® Membrane at ×500) showing a porous structure containing nodes interconnected by fibrils (Courtesy of Menardi-MikroPul, LLC)

Expanded PTFE

In early 1970s, Dr. Robert W. Gore invented a process by which PTFE could be expanded and gave the trade name of Gore-Tex® to the product obtained. In this process, the specific gravity is reduced by the introduction of micropores, but much of the original properties of the polymer are retained. The new material is a hydrophobic but porous membrane of PTFE that is used as a protective layer in a number of applications.

In the process used, in general, a paste is formed of the PTFE polymer and a plasticizer and shaped into an article. It is then expanded by stretching in one or more directions, and while it is held in the stretched form, it is stabilized by heating to high temperature (~327°C) and cooling. The porosity that is produced by the process is retained in the final product [31]. The structure formed by the process consists of "nodes" and "fibrils." The nodes, that vary in size from 5 to 500 μm and are always found perpendicular to the direction of expansion, are interconnected by fibrils. This is seen in Fig. 12.27, which shows a micrograph of Mikrotex®, an expanded PTFE (ePTFE) product [32]. An ePTFE material may have as many as a billion or more randomly spaced pores per square centimeter. These pores are unique in size: they are three or more orders of magnitude smaller than the size of a water droplet but two or more orders of magnitude larger than a water vapor molecule. Thus, when used as a rain wear, the product allows the perspiration vapors to escape but blocks out the liquid water from penetrating. The product can likewise serve as a barrier against chemicals and microbes.

Glass and Carbon Fibers

Glass

Among the manufactured inorganic fibers, glass is produced in by far the largest volume. There has been a rapid increase in the use of glass fibers in both the textile and the industrial grades. Some of the key applications are in belts, air filters, thermal insulation (glass wool), and reinforced plastics.

Glass possesses obvious and well-known characteristics which have largely determined the methods used to form it into large objects. It flows readily when molten and can be drawn into filaments, whose extreme fineness appears to be limited only by the drawing speed. The method used in producing textile-grade glass fibers follows this principle (see Fig. 12.28).

In the commercial operation, the molten glass, produced either directly from raw materials or by remelting of marbles, is held at a uniform temperature in a vessel, whose bottom carries a bushing containing small uniform holes. The molten glass flows through these holes as tiny streams that are attenuated into filaments at speeds on the order of 3,000 m/min; these filaments are coated with a lubricant, gathered into groups to form yarns, and wound up. For a particular glass viscosity, the size of the individual filaments is determined by the combination of the hole size and the speed of attenuation. Because of the inherently high modulus of glass, very fine filaments are required in order to approach the required properties of textile materials. Thus, the diameter of glass filaments falls in the range of 3.8–7.6 μm, whereas the average diameter of the finest organic fibers is about twice as large. The fiber and yarn numbering system is based on nomenclature used in the glass industry and differs from the traditional systems accepted in the textile and organic fiber industries.

The method of manufacture of glass staple fibers differs from those used to produce the corresponding organic materials, all of which are based upon cutting the continuous-filament product. Air jets, directed in the same line of flow as the emerging streams of glass, attenuate the streams and break the solid glass into the lengths desired for further processing, which are gathered on an appropriate vacuum drum and delivered as slivers or a matte. To produce fibers that may be coarser and considerably less uniform in length, to be used for the production of filters, paper, or thermal insulation, large streams of molten glass are cross-blown by blasts of hot air, steam, or burning gas.

As might be expected from the nature of glass, the conversion of glass fibers into final products has required the development of new lubricants, finishes, and processing

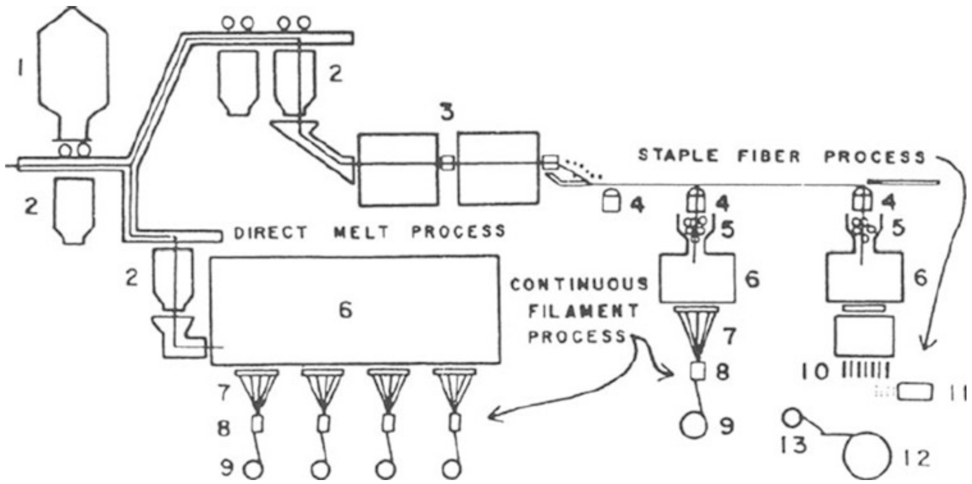


Fig. 12.28 Flow diagram for manufacture of textile glass fiber: (1) glass batch, (2) batch cans, (3) marble forming, (4) cullet cans, (5) marbles, (6) melting furnaces, (7) filament yarn formation, (8) gathering and sizing, (9) yarn packaging, (10) air jets, (11) lubricant spray, (12) collection for staple fibers, (13) staple fiber packaging (Courtesy of Owens Corning Fiberglass Corp.)

techniques. For example, because glass fabrics cannot be dyed directly or printed with the colors demanded for their acceptance as draperies, the colorant must be applied to a resin coating. But before applying the coating, it is necessary to remove the lubricant that was placed on the fibers to permit their conversion into a fabric, which is done by burning. The elevated temperature resulting from this operation also relaxes the internal strains developed in the glass fibers during the steps of the textile operations and sets the yarns in the required geometry. The fabric then is resin-treated, cured, and dyed or printed.

Another inherent property of glass is the tendency of unprotected fiber surfaces to abrade each other to destruction under the action of very little mechanical working. When it was first considered for rubber-reinforcing purposes, its poor adhesion to rubber and the inadequacy of the then-existing bonding agents frustrated attempts by manufacturers to take advantage of the very high tensile strength, completely elastic behavior, high modulus, and lack of moisture sensitivity of glass fibers. However, it has been possible to modify the fiber surfaces so that satisfactory adhesion is achievable, and the impregnant can be applied in such a way that fiber-to-fiber contact is prevented. With the adhesion problems solved, glass in cord form could effectively enter markets in belt-type tire construction and in all kinds of power-transmitting rubber belts. It is estimated that the total glass fiber production in the United States in 2000 was approximately 2.5 billion lb.

Carbon and Graphite Fibers

Following World War II, the development of jet aircraft and rockets created demands for fibers having thermal resistance, strength, and modulus far beyond what could be obtained in existing organic fibers. Much of this need was for reinforcing materials that could be embedded in matrices of one type or another. As a result, techniques have been developed for preparing fibers from a good many metals and refractory inorganic compounds. Although these materials are essential for certain uses, the volume of production still is low, and the prices are correspondingly high (as much as $1,000/lb).

Carbon and graphite fibers are made from rayon and acrylic precursors by driving off virtually all of the hydrogen and oxygen contained in them. The principle is essentially the same as that which brought about the formation of coal or, citing a more recently discovered and dramatic example, the conversion of the original wooden beams of buildings in Herculaneum, buried by a flow of mud from Vesuvius in AD 67, to what appears to be charcoal. In the present commercial process, the starting material is selected so as to produce a final product of the desired size and properties. Cotton, bamboo, and other natural fibers were the earliest materials used as precursors. In the 1950s, rayon was used for this purpose, but the first attempts led to fibers with poor strength. Union Carbide, in the mid-1960s, made the first strong carbon fiber from rayon and extended its work to using PAN as the precursor. The use of mesophase pitch as the starting material began in the 1970s. Carbon fibers may be divided into three types: (1) highly graphitized with a high modulus, (2) heat-treated at a lower temperature to produce fibers with lower modulus but high strength, and (3) randomly oriented crystallites with relatively low modulus, and low strength, but most importantly, low cost

The preparation of fibers generally consists of heat-treating the precursor at a low temperature (200–350°C), usually in air, which gives a stable fiber for higher temperature processing. During this step, extensive decomposition occurs, and a percentage of the initial weight is lost, which is related to the cleavage of the C–C and C=O bonds and expulsion of H_2O, CO, and CO_2 (additionally HCN in the case of PAN precursor) as gases. This step is followed by carbonization in an inert atmosphere at 1,000–2,000°C, which is said to collapse the cyclized structure into a stacked ring carbon fiber structure. Almost all noncarbon elements are removed as volatiles. This is followed by graphitization, which usually is carried out at temperatures above 2,500°C for short periods in argon or nitrogen. The process increases the purity, removes the defects, and further improves the order in the structure. Thus, the difference between the so-called carbon and graphite fibers lies in differences in the ranges of temperatures at which the last step is carried out, the degree of carbon content (97% for carbon and 99.6% for graphite), and the mechanical properties, which are superior in the graphite fibers. In usual discussions, the term "carbon fibers" covers both materials. The majority of carbon fibers produced today are made from a PAN precursor. PAN fibers are finer and have a higher degree of molecular order and a higher decomposition point than those from rayon. This precursor also leads to a greater carbon yield (45%, as opposed to 24% from rayon), but the fibers are more expensive to produce than those based on rayon. Recent commercial developments have allowed the production of carbon fibers from low-cost petroleum or coal-tar pitch, instead of synthetic fibers. These precursors lead to a higher yield (90%) of carbon, improved lubricity of fiber products, and higher production rates. However, the pitch-based fibers may be more brittle and harder to handle, have a higher specific gravity (2.0, as opposed to 1.8 for those from PAN and 1.66 from rayon) and lower compatibility with some matrix materials.

The fibers have a diversity of applications. One major application is in composites, where they are used for reinforcing resins and metals to provide structural materials with high strength, high modulus, and lightweight. The resulting composites are used in the aircraft and aerospace (the largest users), automotive, and sports industries. The fibers also find uses in protective garments, electrical devices, insulation, and filtration. The use of the fibers in the world was about 25 million lb in 2000. The prices dropped significantly during the past decades, but in 2000, the majority remained in the $15–70/lb range. The price of the ultra high modulus carbon fibers, however, can be as high as $1,500/lb or greater.

Sulfar

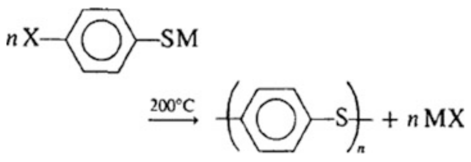
Historical

Sulfar fibers are extruded from poly(phenylene sulfide) or PPS by the melt-spinning process. The first PPS polymer was made in 1897 by the Friedel–Crafts reaction of sulfur and benzene. Researchers at Dow Chemical, in the early 1950s, succeeded in producing high-molecular-weight linear PPS by means of the Ullmann condensation of alkali metal salts of *p*-bromothiophenol.

In 1973, Phillips Petroleum Company introduced linear and branched products under the trade name Ryton® by reacting 1,4-dichlorobezene with sodium sulfide in a dipolar aprotic solvent. In 1983, the same company succeeded in stable melt spinning of PPS. In 1986, the Federal Trade Commission gave the fiber the generic name sulfar, defined as "a manufactured fiber in which the fiber forming substance is a long chain synthetic polysulfide in which at least 85% of the sulfide (–S–) linkages are attached directly to two aromatic rings."

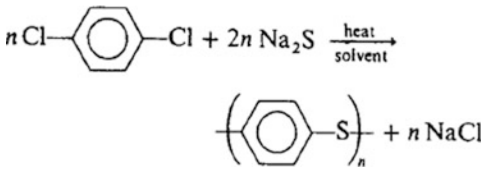
Manufacture

In one process, synthesis was carried out by self-condensation of a metal salt of a *p*-halothiophenol



where *X* is a halogen, preferably Br, and *M* is a metal (Na, Li, K, or Cu, preferably Cu). This reaction was carried out under nitrogen in the solid state or in the presence of materials such as pyridine as reaction media. Considerable difficulty was encountered in removing the by-product, copper bromide, from polymers made by this process. The current commercial synthesis of the polymer is carried out by reaction between *p*-dichlorobenzene and sodium sulfide in a polar solvent.

The process discovered by workers in the laboratories of Phillips Petroleum Co. marked a significant departure from prior processes and made it possible to prepare a variety of arylene sulfide polymers from the readily available starting materials [33].



Melt spinning of PPS involves problems such as plugging of the filter and the spinnerette. In order to prevent cross-linking and gel formation during the process of spinning, the company introduced cure retarders comprising group IIA or group IIB metal salts of fatty acids, which improved the heat stability of PPS:

$$(CH_3(CH_2)_nCOO-)_2M,$$

where M is a group IIA or IIB metal and n is an integer from 8 to 18. Representative compounds of the type described above include calcium stearate, calcium laurate, calcium caparate, and calcium palmitate. In general, the cure-retardant additives are employed in an amount within the range, of about 0.1–5, preferably about 0.5–2, weight percent based on the weight of the PPS.

Sulfar fibers, sold under the trade name Ryton® in the United States, are characterized by high heat resistance, inherent flame retardancy, excellent chemical resistance, low friction coefficient, good abrasion resistance, and good electrical properties [34]. Physical characteristics include medium tenacity (3.5 g force/denier or 423 MPa) and elongation (25–40%) and low shrinkage (<5% at 100°C). The fibers find application in a number of industrial products, including filter fabric for coal-fired boiler bag houses, paper maker's felts, materials for electrical insulation, high-performance composites, gaskets, and packings.

Several Japanese and European companies have begun the production of PPS, some with a U.S. partner. The new decade should see the introduction of a number of new fibers and fiber products based on the PPS polymer.

Microdenier Fibers

One of the more important developments in the field of fiber technology in recent years has been the production of fine denier fibers, with worldwide activity in developing new products and applications for them. Improvements in the quality of polymers, coupled with new technology for extrusion of fibers, have led to the production of fibers with sizes ranging from 1 to 0.1 denier or even lower. Fabrics produced from such fibers have novel and unique properties, and they are finding applications in a wide variety of apparel, and industrial products, including high-fashion fabrics with silk-like texture, synthetic suede, breathable porous but rainproof or bacteria-proof fabrics, wipes for oil and other spills, clean room materials, and cloths for cleaning camera, microscope, and spectacle lenses and compact discs.

Unfortunately, there is no universally accepted definition for microdenier, and companies have been free to use terms they choose. In general, there seems to be a consensus that the term "fine" may be used for a denier of 1 or less and "micro" for a denier less than 0.5. The production of such fibers is difficult and expensive because the throughput rate must be reduced in order to obtain fine denier filaments, and there is a limit to how fine in denier a regular process could be made to go. Generally, the technology of production of fine and microdenier filaments can be divided into five categories. First, conventional spinning technology can be used by using fine-size dies and adjusting the throughput rate and quenching and drawing parameters to obtain fine denier polyester filaments. This method has been used by most fiber manufacturers to produce such fibers in limited quantities. In another method, alkali reduction or surface etching is used to dissolve the surface layers, in an effort to reduce the cross-sectional size of filaments in a polyester fabric. A weight loss of as much as 25% has been achieved by this process. In the MB process, molten polymer is forced through a melt-blowing die and die-tip orifices, and the emerging stream is attenuated by a jet of high-velocity hot air until broken. Then the broken fiber is forced into a stream of cold air, where the fiber solidifies. The fiber is collected on a wire screen or apron with other fibers, and a homogeneous MB web is produced. This method is used extensively with PP materials, but other polymers such as nylon, polyester, and polyethylene also have been used. Fibers as fine as 0.5 μm and finer have been produced.

In another approach, filaments containing two polymers that do not adhere to each other are spun and then split. One may, for example, spin a bicomponent fiber of nylon that has several filaments of polyester embedded. After a fabric containing bicomponent filaments has been woven, it is treated to split the components, thus converting the original filament into several smaller filaments. Deniers of the order of 0.1 can be achieved by this ingenious method. In the last method, instead of splitting the two components as in the previous example, one component is dissolved away chemically, leaving bundles of very fine fibers in the fabric.

The majority of the technology for producing fine and microdenier fibers is new and thus expensive, so efforts in the future can be expected to be directed toward optimizing process parameters in existing methods and discovering faster and cheaper methods for manufacturing these esthetically attractive and functionally very promising fibers.

Nanofibers

An extreme example of the microdenier-size fibers discussed above is the nanofibers that are one or more orders of magnitude smaller in diameter. The fiber with diameter at submicron or nanometer level is spun by a process known as electrospinning, in which the fiber is spun in a field involving high electrostatic forces [35].

The electrospinning setup essentially consists of a capillary tube or a needle attached to a syringe filled with a polymer solution or melt, a grounded collector screen, and a high-voltage power supply (Fig. 12.29). The collector

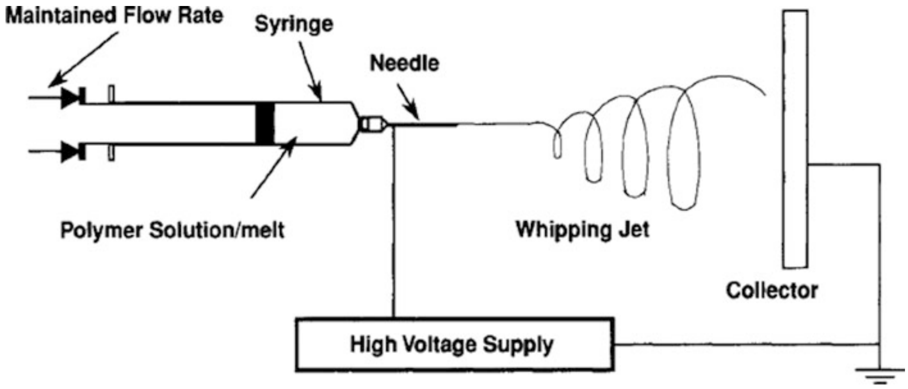


Fig. 12.29 Electrospinning setup in horizontal configuration

typically is a metal plate, an aluminum foil, or a metal grid. Other forms of the collector, such as a rotating drum, have been used. When the potential (1–30 kV) is applied, the pendant droplet at the end of the needle becomes charged, and two opposite forces, namely, the surface tension of the droplet itself and the electrostatic force due to applied field, act on it. As the potential is increased, at a certain point, electrostatic force overcomes the surface tension of the droplet, and a cone-like structure, commonly known as a "Taylor cone," is formed at the tip of the droplet, and a jet of polymer solution or melt emanates from it which is accelerated toward the collector [36]. In this process, the polymer jet gets drawn to submicron level while the solvent evaporates or the melt cools down to form the fibers that deposit on the collector and form a nonwoven web.

The electrospun nanofiber webs, with very high surface area to volume ratio and high porosity with interconnected pore networks, have been explored for applications in a number of highly specialized areas such as the scaffolding for tissue engineering, wound dressing, drug delivery, nanosensors, and nanoelectronics. More than 50 different polymers have been electrospun successfully in a solution or melt form that include synthetic nondegradable (polyesters, polyamides, polyurethanes, polycarbonates, etc.), natural (collagen, elastin, chitosan, silk, etc.), and synthetic biodegradable (polyglycolic acid, polycaprolactone, polylactic acid, polylactide-*co*-glycolide, etc.) polymers [37].

Fiber Variants

Introduction

In a previous section, data and plots were given showing the rapid rise in consumption and production of commonly used manufactured textile fibers at the expense of natural fibers. The principal reason for this has been the wide range of manufactured fiber variants that can be produced from a single fiber-forming polymer. The wide range of polymers available, each with its particular properties, adds yet another dimension. This is not to say that there is only one type of cotton, wool, silk, or asbestos fiber; there are many varieties of natural fibers, but their supply is limited by natural factors such as climate and genetics. The relative availabilities of manufactured fiber types can be altered by controlled chemical process changes, whereas the amount and the quality of a desired cotton type that can be grown is determined to a great extent by climatic conditions, which humans have not yet learned to control. Another factor that has aided the growth of manufactured fibers is their consistent quality and properties. Again, the grade and the quality of natural fibers are subject to the vagaries of nature.

For the purposes of this discussion, fiber variants will be divided into two types: chemical and physical. Chemical variants will be those involving a small but significant change in composition, whereas physical changes will be those involving a change in either the dimensions of the fiber or its stress/strain or stability features. The definitions of the two variants also could be based on modification of either esthetics or functionality.

Physical Variants

Most manufactured fibers are available as staple, tow, and filament. Natural fibers are available only in the characteristic forms in which they occur, with filament silk and cotton staple as examples. All manufactured fibers are formed initially as filament yarns. The German adjective *endlos* (literally translated "endless") is very descriptive, as filament yarns are continuous strands consisting of one or more members that for most practical purposes are infinite in length. Fine filament yarns (40–100 denier) are used in producing lightweight apparel fabrics, whereas coarse filament yarns (800–1,200 denier) are found as reinforcement in tires or conveyor belts. These examples are chosen as extremes to show the range or applicability of manufactured fibers and represent only a small fraction of the actual range of end uses.

When many filament yarns are collected into a bundle immediately after formation, the resulting structure is called a tow. Tows may range from 10,000 to over a million in total denier. In a next step, the tow may be crimped by the process previously described, which imparts what is usually a sawtooth appearance along the length of all the filaments. For some end uses, the crimped tow itself is provided by the fiber maker, as are, for example, the previously discussed acrylic tows, which are converted to staple as part of the spinning of yarns with a woollike character. Another example is the cellulose acetate tows used to form cigarette filters or the ink reservoirs for marker pens. In the latter case, the compact tow bundle first is treated to separate the individual filaments, giving a voluminous structure, which then is gathered into a continuous rod, wrapped in paper, and cut into appropriate lengths.

In the fiber-making plant, tow may be cut into short lengths of staple, ranging in length from $\frac{1}{16}$ to 6 in., depending upon the end use. For instance, the very short staples are used either in making flocked structures or in the production of papers containing blends of natural cellulose and manufactured fiber. The longest lengths are used in spinning heavy yarns for carpets or cordage. But most staple is $1\frac{1}{2} - - 3$ in. length and is used to form blends with cotton, rayon, or wool in the yarns employed in standard apparel fabrics. The staple length of the manufactured fiber is chosen to match that of the other blend component; otherwise, uneven yarn of poor quality results.

The size of a manufactured fiber can be altered by changing the size of the hole through which it is extruded and maintaining a constant take-up speed. Commercial fibers range from about 1.25 to 25 denier, corresponding to average diameters of about 5–50 μm. The very large ones would be used to make doll's hair or wigs, but the majority of manufactured fiber staples are made in the 1.5–6.0 denier range, corresponding to cotton-blending fiber at the low denier end and coarse-wool-blending fiber at the high end. Staple and filament yarns used in carpets are in the 12–16 denier range, whereas industrial filament yarns such as tire cord are about 6 denier per filament. The size of a fiber is a determinant of its stiffness, which in turn influences the draping, quality, and the surface feel (often called hand in the textile industry) of a fabric made from it.

Although fibers normally are extruded through circular jet holes, the use of noncircular holes in the jets has led to the availability of a wide variety of cross-sectional shapes. In the case of fibers dry- or wet-spun from solutions, most of the mass exiting from the jet hole is not the polymer. For example, cellulose acetate fiber is made from a 25% solution of cellulose acetate in acetone containing a small amount of water. After the fiber leaves the jet face, solvent begins to evaporate, and as a result, the area of the fiber cross section decreases. The final result is a fiber of roughly circular cross section, but with a serrated edge, and much smaller in area than the parent jet hole. When cellulose acetate is extruded through a triangular jet hole, the end result is a fiber of "Y" cross section due to shrinkage from the original triangular shape.

In the melt-spinning process, there is no solvent loss to influence final cross-sectional shape. In the case of a triangular jet hole in melt spinning, the molten fiber leaves the jet face with a triangular cross section but, being molten, immediately tends to return to a circular cross section due to surface tension forces. It is necessary, therefore, to quench or cool the fiber as soon as possible in order to maintain the desired cross-sectional definition. Some loss of this definition is unavoidable during the drawing step that normally follows extrusion in a melt-spinning process.

Fibers of noncircular cross section can modify and change both functional and esthetic properties in textile structures. The triangular cross section is typical in those respects; its shape leads to a stiffer fiber than circular fiber of the same cross-sectional area, and in a fabric, this results in less drapability and a crisper surface feel. Also, the flat surfaces reflect light in a different way than do curved surfaces and can create desirable lustrous effects. These optical effects are subject to many subtle influences having to do with the size of the reflecting surface and the amount of internal reflection that takes place. Triangular or "Y" cross-sectional fibers have greater specific surface areas per unit weight than their circular counterparts, which accounts for their use in aerosol filtration, where surface area is a major factor in efficiency. Yarns made with triangular cross section fibers are more voluminous than those from round cross section fibers. Thus, when the two types are made into fabrics of equal weight, the variant cross-sectional fabric will transmit less light and be less permeable to air than that made with round cross section fibers. In the spinning of blended staple yarns, maximum strength is obtained if the blend components have similar load/elongation characteristics. The stiffness of polyester staple fibers can be varied mainly by changing the draw ratio used to orient the fiber after extrusion and thus producing fibers suitable for blending with cotton, rayon, acrylic, or wool fibers, which possess markedly different properties. In many industrial applications, fibers and yarns are used under conditions where they bear a load while hot, as, for example, in tires and power transmission belts, where it is important that they do not grow or stretch significantly under these conditions. Accordingly, industrial yarns are drawn to a greater extent than other yarns to reduce as much as possible the stretch of the final products. Relatively, these yarns have high strengths, high stiffnesses, and low elongations at rupture.

Although in most instances it is desirable to have fibers that are dimensionally stable, in some structures, it is advantageous to use mixtures of fibers that are stable with fibers that do shrink upon exposure to heat or steam. High shrinkage potential usually is built into a fiber by stretching it and not giving it a stabilization treatment. The yarn bulking that occurs with blend yarns of high- and low-shrinkage acrylic staples has already been described in the section on that fiber. Using the same principle, feltlike structures can be made by heat treatment of nonwoven battings containing low- and high-shrink staple.

Chemical Variants

The use of titanium dioxide as an additive in the delustering or dulling of manufactured fibers was discussed in the section on viscose rayon. The addition of pigment particles influences the processing and performance of the fiber, along with changing its appearance. Because of the whiteness of the delustered fiber, it requires more dyestuff to reach a given shade than that needed in the case of a bright fiber. The sliding friction of a delustered fiber is lower because the pigment particles protruding from the fiber surface reduce the contact area between a fiber and the surface it slides against—a guide, for example. By the same token, some pigments can accelerate the wear of contacted surfaces. It has been suggested that the drawing and consequent orientation of delustered fibers proceeds more smoothly because the pigment particles act as nucleation sites where molecular motion is initiated. Unless specially treated, the surface of anatase—one of the crystalline forms of titanium dioxide—can accelerate the ultraviolet light degradation of acetate or nylon fibers. It is postulated that the crystal surface catalyzes the formation of peroxides from the water and oxygen under the influence of ultraviolet light, and that peroxides are the active species in the resulting polymer degradation. For this reason, rutile—the other common crystalline form of titanium dioxide—is used to deluster fibers when improved sunlight resistance is needed.

The degree of polymerization or the molecular weight of the polymers on which manufactured fibers are based can be controlled as part of one of the early steps in the process. The polymer molecular weight chosen for a fiber has a strong influence on process economics, ease of conversion to fiber, and end-use performance. Commercially acceptable fibers are based on the best balance of these factors. However, in fabrics that are open in texture and made from standard polyester staple spun yarns of low twist, a condition known as "pilling" will develop as a result of wearing. The pills are made up of fiber ends that have worked loose from the yarn bundles as a result of surface rubbing and have wrapped around themselves. In the case of fabrics from natural fibers, which are generally less wear resistant than manufactured fibers, these pills or fiber bundles will be lost by attrition with continuing wear. Because the wear resistance of a manufactured fiber is related to its molecular weight in a general way, the pilling tendency of a polyester staple can be reduced by lowering its degree of polymerization. This compromises the tensile characteristics of the fiber only to a small extent, and all the other desirable properties such as minimum care characteristics are essentially unaffected. When the molecular weight of a polyester or a nylon polymer is increased above the standard, the resultant fibers will have increased tensile properties and fatigue resistance. The filament polyester and the nylon yarns used in end products such as tires and conveyor belts are based on such polymers.

The technology for "dope dyeing" or mass coloration of fiber as part of the fiber manufacturing operation was described in the sections on viscose rayon and cellulose acetate. The use of this technology has decreased largely because of problems of profitably managing the required inventory of colors in rapidly changing fashion markets, and for this reason, it is used extensively only if the fiber cannot be dyed by any other means. PP fiber for use in outdoor carpets is a good example of such a situation. However, if one regards "white" as a color, one finds that a substantial portion of the polyester staple fiber produced for blending with cotton or rayon contains an optical brightener or a fluorescing agent. This is needed to overcome the yellowing tendency of polyester following the absorption of hydrophobic soils. The cotton or rayon fibers in a blend are continually rewhitened by the fluorescing agents added to laundry detergents for that purpose, but these agents are without effect on the polyester blend component. Because the polyester component usually is at least half of the blend, the spun-in optical brightener it contains is vital for the maintenance of overall whiteness.

In the cases of polyester, nylon, and acrylic fibers, their manufacturers have developed fiber variants with a wide range of dyeing behavior referred to as dye variant fibers. Polyester fibers usually are dyed with what are described as disperse dyes. These dyes are only slightly soluble in boiling water and are used in the form of dispersions. The dye in aqueous solution is assumed to be in monomolecular form and is absorbed from the dyebath into the polyester fiber by a process often called solid solutioning. As this occurs, more solid dye is dissolved to replenish that entering the fiber. The overall rate of dyeing is a very complex phenomenon but in part is determined by the molecular structure of the fiber through which the dye must diffuse. Generally, the more that the polymer molecules have been organized by drawing and annealing into more geometrically perfect domains or crystallites, the slower is the dyeing rate. The ability of the molecules to be thus organized into compact ordered structures can be reduced by polymerizing a small amount of a foreign dibasic acid or glycol into the polymer. Usually, 5–10 mol% is sufficient to prevent this regularity. Adipic acid, isophthalic acid, and polyethylene glycols are used to produce some of the comonomers for fast- or deep-dyeing polyesters. In this way, these fibers are made more economical to dye or print because special dyebath additives, high dyeing temperatures, and high-pressure steam-print fixation usually are not required.

Polyester fibers can be given an additional mechanism for dyeing if an ionic comonomer is added during polymerization. A common additive is an alkali metal salt of dimethyl-5-sulfo-isophthalate, which gives sulfonic (anionic) groups as part of the polymer structure. These groups allow the fiber to absorb basic (cationic) dyes by a specific ionic mechanism. The amount of cationic dye that can be absorbed by the fiber is stoichiometrically related to the number of anionic sites present in the fiber; this is quite distinct from the general solid solutioning that takes place with disperse dyes and polyester. A cationic dyeable polyester is useful for two main reasons in fabric coloration. First, cationic dyes give brighter, clearer shades than disperse dyes, and this can be important for both solid-dyed shades and prints during ever-changing fashion cycles. Second, fabrics containing arrangements of unmodified and cationic dyeable polyesters can be dyed in the piece to a variety of color/white combinations by selection of dyestuffs, as day-to-day changes in demand may require. This is more economical than dyeing yarns to different colors, holding them in inventory, and then weaving them into fabrics, if and when they are required.

Nylon 66 can be dyed either with disperse dyes or with acid (anionic) dyes, the former carried out by essentially the same steps as described above for polyester, except that nylon absorbs these dyes much more readily. The ability of nylon to absorb acid (anionic) dyes is the result of a significant number of accessible free amine (cationic) end groups being present in the polymer. The dyeing of nylon with acid dyes is analogous to the dyeing of wool or to the dyeing of modified polyester with basic dyes, except that the polarities of the interacting groups are reversed in the latter case.

For best fastness to light and washing, nylon 66 is dyed with acid dyes, and nylon dye variants thus are based on manipulation of the level of acid dye uptake. By adding a monobasic acid such as acetic acid to the reaction mix near the end of the polymerization process, the amino end groups are converted to amide groups which have no affinity for acid dyes under normal dyeing conditions. This technique creates light acid dyeing or acid reserve dye variants that can have some capacity to absorb basic dyes at the carboxyl end groups. Nylon dye variants with increased acid dye uptake can be made by using a slight excess of diamine in the polymerization. In this way, there are no free carboxylic acid end groups. Nylon dye variants have found the greatest acceptance in floor coverings where attractive patterns can be piece-dyed using controlled dyebath conditions and selected acid dyestuffs.

Acrylic fibers are dyed most frequently with basic dyes. This is made possible by copolymerizing acrylonitrile with an acidic monomer such as styrene-*para*-sulfonic acid. Acrylic fiber suitable for acid dyeing can be made by using a basic comonomer such as a vinyl pyridine or a vinyl pyrrolidone.

Fibers not having inherent flame resistance often can be given this property by incorporation of a suitable additive. This may be done by copolymerization of the additive into the polymer reaction of the additive with the polymer after polymerization or by applying a polymeric or monomeric noninflammable finish to the surface as a coating. These additives usually contain bromine, nitrogen, or phosphorus or a combination of these elements. Great care must be taken in choosing the additive and its level of addition in order to prevent loss of other desirable fiber properties and to avoid any harmful effects to processors or ultimate consumers.

A wide variety of special durable surface treatments have been used on manufactured fibers. These include treatments for imparting such characteristics as soil resistance, antistatic behavior, and wearer comfort through moisture wicking and transport. Fiber finishes also have been used successfully in promoting adhesion between two materials, as, for example, between polyester tire cord and rubber, and between glass fiber and polyester resin.

Polymers, Fibers, and Textiles in Medicine

Introduction

Over the past six decades, polymers, fibers, and fabrics have found numerous applications in medicine. Starting from their use as sutures over 4,000 years ago to their application as arterial, skin, hernia repair meshes, and other prostheses in the present century, fibrous materials continue to be selected and explored for new and novel products in the healthcare field. This trend has its basis in a fiber's ability to be modified chemically and physically and easily converted into a wide array of desired end products. Textiles are unique materials with unique properties. They are known to be lightweight and flexible, and, what is commonly not realized is that on a weight-to-weight basis, they are very strong, that is, usually two or more times stronger than steel. Since the majority of fibers are extruded products of synthetic or regenerated polymers, they can be manufactured in different cross-sectional shapes and sizes to meet the design needs of different applications. In addition, a given polymer can be alloyed with another material and/or chemically or physically modified to further vary its characteristics. Accordingly, because of the outstanding inherent properties of fibrous polymers, essentially unlimited potential one has for tailoring their behaviors, and those of several different types of assembly structures made from them, textiles have been sought for designing and used in manufacturing a wide range of medical products.

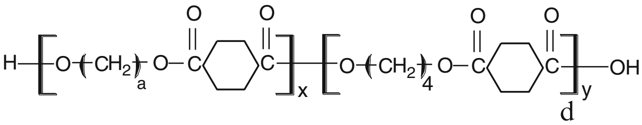
Examples of Medical Products and Fibers Used to Produce Them

Broadly, the medical products available include absorbent materials for blood and other fluids resulting from injury, surgery, or incontinence (cotton, rayon); protective gowns, masks, and covers for patients, nurses, and doctors (cotton, polyester, polyolefin); dressings and bandages for caring for and managing wounds (cotton, rayon, chitosan, alginate); fabrics for repair of fractures and rehabilitation of injured (fiber glass, polyester, polyolefin, cotton); and, most noteworthy, many different types of implants, which include sutures and ligatures for closing wounds and prostheses for replacing or bypassing injured or failing organs. Polymer-based fibers and textiles are particularly suited for making implants as body tissues are themselves polymers and have structures and properties that closely match those found in fibers. Among the implants, some relatively well known are surgical sutures (PET or polyester, polypropylene, ePTFE, nylon, polybutester, and the aliphatic or biodegradable polyesters, mentioned and discussed in the next section), arterial and stent grafts (polyester, ePTFE), hernia and prolapsed repair meshes (polypropylene, polyester), and tendon and ligament prostheses (polyester). The examples of products given are of textile-based structures that must be largely flexible and able to stretch, bend, and twist. Used as implants are also strictly polymer-based devices such as those used for correcting vision (ocular lenses), filtering blood (in kidney dialysis devices), and repairing or replacing joints and bones. Additionally, fibrous polymers have drawn interest for use as scaffolds for engineering tissues.

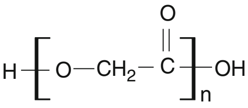
Special Fibers Developed Primarily for Use in Medicine

Clearly, many different types of fibers (natural and manufactured) have been used in medicine, some of these (manufactured) have already been covered in detail earlier in this chapter. There are several others that are more specific to medicine and were in fact largely developed for this application. These, along with their chemical structures, are listed below. Out of these, the first six are polyesters and the last two polysaccharides. Of these, the first, that is, polybutester, is an aromatic block copolymer, which is nonabsorbable and has properties that are a compromise between fibers and elastomers. The next five (polyglycolic acid, polylactic acid, polyglactin, polydiaxanone, and polycarbonate) are aliphatic and biodegradable. The intermolecular links being weak and the molecular weights being low, these materials break down by the hydrolytic processes in the body and get absorbed. Some (e.g., PGA) are absorbed faster while some others (e.g., polyglyconate) are absorbed more slowly, the rate controlled by the chemical constitution, molecular weight, and the degree of packing [38–41]. Although these materials were initially developed for use as biodegradable sutures, they are now also being explored in constructing implant devices. The general idea behind this trend being that as the material biodegrades, the natural tissues will infiltrate and fill the void, leading to a shaped biological implant. The exploratory work along this direction is continuing with some success [42, 43].

Polybutester (Novafil):

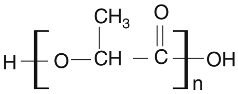


PGA (Dexon):



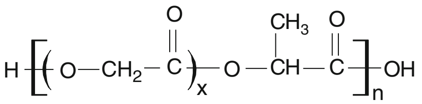
$$\mathrm{H}\left[\mathrm{O}-\mathrm{CH_2}-\overset{\overset{\displaystyle \mathrm{O}}{\|}}{\mathrm{C}}\right]_n\mathrm{OH}$$

PLA:



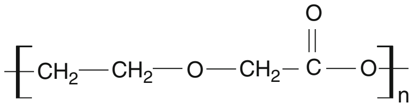
$$\mathrm{H}\left[\mathrm{O}-\overset{\overset{\displaystyle \mathrm{CH_3}}{|}}{\mathrm{CH}}-\overset{\overset{\displaystyle \mathrm{O}}{\|}}{\mathrm{C}}\right]_n\mathrm{OH}$$

Polyglactin acid (Vicryl):



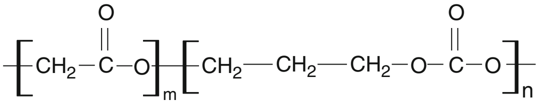
$$H\left[\left(O-CH_2-\overset{\overset{O}{\|}}{C}\right)_x-O-\overset{\overset{CH_3}{|}}{CH}-\overset{\overset{O}{\|}}{C}\right]_n OH$$

Polydiaxanone (Maxon):



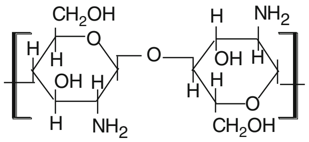
$$\left[CH_2-CH_2-O-CH_2-\overset{\overset{O}{\|}}{C}-O\right]_n$$

Polyglyconate (Maxon):

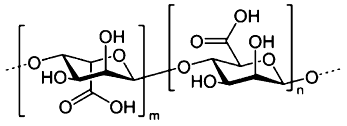


$$\left[CH_2-\overset{\overset{O}{\|}}{C}-O\right]_m\left[CH_2-CH_2-CH_2-O-\overset{\overset{O}{\|}}{C}-O\right]_n$$

Chitosan:



Alginate:



The last two materials in the above list, that is, chitosan and alginate, are polysaccharides. Both, derived from natural polymer sources and biodegradable, are primarily used as wound dressings for complex or bleeding wounds. Chitosan is made from chitin by converting the pendant acetyl group into amino group with a degree of substitution of about 80%. This conversion makes the material highly reactive and soluble. Due to its high hemostatic nature and ability to promote healing, chitosan has been considered as a material of choice for dressings for bleeding wounds [44]. Alginate, which is a block copolymer of mannuronic and glucuronic acids, is derived from seaweeds. Sodium alginate, extracted from the crushed and washed raw material, is solubilized in water and extruded into a calcium bath to produce calcium alginate fiber. When a product, such as a web, comes into contact with blood or body exudates, the sodium ions replace the calcium ions, which makes the material highly absorbent and soluble [45]. Desirable healing and antimicrobial compounds could be incorporated during the production of the fiber. Accordingly, a primary use of the material is in dressings for exuding and contaminated wounds.

Assembly Structures Used in Implants

The textiles used in body are in the form of either one-dimensional structures (threads), plane or two-dimensional structures (sheets), or three-dimensional structures (primarily tubes) [46]. The first, or the thread, is usually braided and used as a suture or ligature. The plane sheet (woven, knitted, or braided) is used as a patch or skin graft, a part of hernia or prolapsed repair device, cover for sewing ring for heart valve, and base material for developing other products (e.g., heart sac) that require a flexible and porous textiles but for which a tubular structure will not serve the purpose. The tubular structure (produced primarily by weaving or warp knitting) is used as a bypass or stent graft for alleviating blood flow disorders in natural vessels. Such tubular devices could be straight, bifurcated, or even trifurcated, depending on the region where corrective action is required. Except the suture or ligature materials, the other textile products are generally required to be porous for healing. The porosity in a fabric is determined by the size of the thread used in constructing the fabric and the thread counts in the long and the cross directions of the fabric. If the constituent thread is a multifilament yarn, then the porosity can also be influenced by the size of the fiber and the bulk given to the yarn, such as by a texturizing process. In ePTFE materials, porosity is greatly influenced by the heating and the stretching conditions used in forming the product.

A vast majority of the textile implants are made by the conventional technologies of weaving, knitting, and braiding, which produce desirably flexible (for conformation and comfort) and porous (for tissue integration and healing) structures. In some applications, however, such as the lenses used for correcting vision and the dental products, tissue integration is not required. For such uses, a molded or machine-shaped product could be made directly from a polymer.

Acknowledgements The author gratefully acknowledges the assistance he has received from associates both from within the College of Textiles and from outside, including several fiber-producing companies, in preparing this chapter.

References

1. Time-Life Books (1970) Seven centuries of art. Time-Life Books, New York
2. Mark H, Whitby GS (1940) Collected papers of W. H. Carothers. Wiley, New York
3. Turbak A (1985) Rayon. In: Klingsberg A, Baldwin T (eds) Encyclopedia of polymer science and engineering, vol 14, 2nd edn. Wiley, New York, p 55
4. Gupta BS, Hong CJ (1995) Absorbent Characteristics of Nonwovens Containing Cellulosic Fibers. INJ 7(1):38
5. Davis S (1989) All you need to know about tencel. Text Horizons 9(2):62
6. Albrecht W, Reintjes M, Wulfhorst B (1997) Lyocell (alternative regenerated cellulose fibers). Chem Fibers Int 47:298
7. Markham JW (1952) Competition in the Rayon industry. Harvard University Press, Cambridge, MA, p 16
8. Robinson JS (1980) Fiber-forming polymers: recent advances. Noyes Data Corporation, Park Ridge, NJ
9. Barnes CE (1987) Nylon 4-development and commercialization. Lenz Ber 62:62–66
10. O'Sullivan D (1984) Conventional Nylons Encounter Strong New Competitor in Nylon 46. Chem Eng News 62(21):33
11. Jung D-W, Kotek R, Vasanthan N, Tonelli AE (2004) High modulus Nylon 66 fibers through Lewis acid–base complexation to control hydrogen bonding and enhance drawing behavior. American Chemical Society. Polym Mater Sci Eng Div Preprints 91:354–355
12. Davis GW, Everage AE, Talbot JR (1984) Polyester fibers: Variants. Fiber Producer 12(6):45
13. Smierciak RC, Wardlow E, Lawrence B (1997) US Patent 5,602,222
14. Smierciak RC, Wardlow E, Lawrence B (1997) US Patent 5,618,901
15. Hutchinson SR (2005) Thermoplastic polyacrylonitrile. M.S. Thesis, North Carolina State University, Raleigh
16. Ahmed M (1982) Polypropylene fibers-science and technology, textile science and technology, vol 5. Elsevier, New York, p 16
17. Lieberman RB, Barbe PC (1990) Propylene polymers. In: Kroschwitz JI (ed) Concise encyclopedia of polymer science and engineering. Wiley, New York, p 916
18. Hogan JP, Banks RL (1986) History of crystalline polypropylene. In: Seymour RB, Cheng T (eds) History of polyolefins. Reidel, Boston, p 103
19. Gupta BS, Smith DK (2002) Nonwovens in absorbent materials. In: Chatterjee PK, Gupta BS (eds) Absorbent technology. Elsevier, Amsterdam, p 378
20. Madsen JB (2001) New generation of spunmelt composites. Nonwovens World 69, Aug–Sept 2001
21. Sekar N (2000) Dyeable polypropylene fibres: On the research front - review of developments. Colourage 47(2):33
22. Kotek R, Afshari M, Gupta B, Kish MH, Jung D (2004) Polypropylene alloy filaments dyeable with disperse dyes. Color Technol 120:26
23. Zwijnenburg A, Pennings AJ (1976) Longitudinal growth of polymer crystals from flowing solutions III. Polyethylene crystals in Couette flow. Colloid Polym Sci 254:868
24. Smith P, Lemstra PJ (1980) Ultra-high-strength polyethylene filaments by solution spinning/drawing. J Mater Sci 15:505
25. Kavesh S, Prevorsek D (1983) US Patent 4,413,110, to Allied Chemical
26. Kwolek DL (1971) US Patent 3,600,350, to E. I. duPont de Nemours and Company
27. Blades H (1973) US Patent 3,767,756, to E. I. duPont de Nemours and Company
28. McIntyre E (1988) High Performance for Industrial fibers. Text Horizons 8(10):43
29. Chenevey EC, Conciatori AB (1970) US Patent 3,549,603, to Celanese Company
30. Coffin DR, Serad GA, Hicks HL, Montgomery RT (1982) Properties and Applications of Celanese PBI - Polybenzimidazole Fiber. Text Res J 52:466
31. Gore RW (1973) US Patent 3,953,566, to W. L. Gore & Associates, 27 April 1973
32. Menardi-MikroPul L. L. C. www.mikropul.com/products/media/mikrotex.html
33. Edmonds JT Jr, Hill HW Jr (1967) US Patent 3,354,129, to Phillips Petroleum Company
34. Scruggs JG, Reed JO (1985) Polyphenylene sulfide fibers. In: Lewin M, Preston J (eds) High technology fibers, Part A. Marcel Dekker, New York
35. Formhals A (1934) Process and apparatus for preparing artificial threads. US Patent 1,975,504
36. Reneker DH, Chun I (1996) Nanometre diameter fibres of polymer, produced by electrospinning. Nanotechnology 7:216–223
37. Huang ZM, Zhang YZ, Kotaki M, Ramakrishna S (2003) A review on polymer nanofibers by electrospinning and their applications in nanocomposites. Compos Sci Technol 63:2223–2253
38. Greisler HP (1991) Biohybrids-biological coatings in vascular grafts. In: New biologic and synthetic vascular prostheses. R. G. Landes Company, Austin, pp 33–46
39. Hakkarainen M (2002) Aliphatic polyesters: abiotic and biotic degradation and degradation products. Adv Polym Sci 157:113–138
40. Xue L, Greisler HP (2003) Biomaterials in the development and future of vascular grafts. J Vasc Surg 37:472–480
41. Hutmacher DW (2000) Scaffolds in tissue engineering bone and cartilage. Biomaterials 21:2529–2543
42. Kim SH, Kwon JH, Chung MS, Chung E, Jung Y, Kim SH, Kim YH (2006) Fabrication of a new tubular fibrous PLCL scaffold for vascular tissue engineering. J Biomater Sci Polym Ed 17 (12):1359–1374
43. Haslauer CM, Moghe AK, Osborne JA, Gupta BS, Loboa EG (2010) Collagen-PCL sheath-core bicomponent electrospun scaffolds increase osteogenic differentiation and calcium accretion of human adipose-derived stem cells. J Biomater Sci Polym Ed 22 (13):1695–1712
44. Smart G, Miraftab M, Kennedy J, Groocock M (2005) Chitosan: crawling from crab shells to wound dressings. Medical textiles and biomaterials for healthcare. Woodhead Publishing, Cambridge
45. Wang L, Khor E, Wee A, Lim L (2002) Chitosan-alginate PEC membrane as a wound dressing: assessment of incisional wound healing. J Biomed Mater Res 63(5):610–618
46. Gupta BS (1998) Medical textile structures: an overview. Med Plast Biomater 5(1):16–30

Suggested Reading

The reader is referred to the four encyclopedias listed below for additional information. They contain enormous quantities of information on manufactured fibers as well as comprehensive bibliographies.

Concise encyclopedia of polymer science and engineering. Wiley, New York, 1990

Encyclopedia of polymer science and engineering, 2nd edn. Wiley, New York, 1985 (17 volumes, index volume, and supplement volume)

Encyclopedia of polymer science and technology. Interscience Publishers, New York (16 volumes)

Kirk-Othmer encyclopedia of chemical technology, 3rd edn. Interscience Publishers, New York (21 volumes and a supplement, 3rd edn; to date, 16 volumes)

The following books contain broad discussions of manufactured textile fibers.

Baer E, Moet A (eds) (1991) High performance polymers. Hanser, New York

Billmeyer FW (1984) Textbook of polymer science. Wiley, New York

Ciferri A, Ward IM (eds) (1979) Ultra-high modulus polymers. Applied Science, London

Datye KV (1984) Chemical processing of synthetic fibers and blends. Wiley, New York

Hearle JWS, Peters RH (eds) (1963) Fibre structure. The Textile Institute, Manchester, Butterworths

Lewin M, Preston J (eds) (1983) Handbook of fiber science and technology: high technology fibers, vol III. Marcel Dekker, New York

Mark HF, Atlas SM, Cernia E (eds) Man-made fibers; Science and technology, vols I, II, and III. Wiley, New York, 1967, 1968, and 1968

Moncrieff RW (1975) Man-made fibres, 6th edn. Wiley, New York

Morton WE, Hearle JWS (1993) Physical properties of textile fibres. The Textile Institute, Manchester, Butterworths

Peters RH (1967) Textile chemistry; The chemistry of fibers, vol I and Impurities in fibers; Purification of fibers, vol II. Elsevier, New York, 1963 and 1967

Dye Application, Manufacture of Dye Intermediates and Dyes

13

H.S. Freeman and G.N. Mock

Introduction

Dyeing

It is difficult if not impossible to determine when mankind first systematically applied color to a textile substrate. The first colored fabrics were probably nonwoven felts painted in imitation of animal skins. The first dyeings were probably actually little more than stains from the juice of berries. Ancient Greek writers described painted fabrics worn by the tribes of Asia Minor. But just where did the ancient craft have its origins? Was there one original birthplace or were there a number of simultaneous beginnings around the world?

Although it is difficult to determine just when each respective civilization began to use dyes, it is possible to date textile fragments and temple paintings, which have survived the ensuing centuries. The ancient Egyptians wove linen as early as 5000 BC, and paintings on tomb walls infer that colored wall hangings were in use by 3000 BC. By 2500 BC, dyer's thistle and safflower were used to produce red and yellow shades. Egyptian dyers developed a full range of colors by 1450 BC.

Another cradle of civilization was the Indian subcontinent where religious and social records dating to 2500 BC refer to dyed silk and woven brocades of dyed yarn. Cotton, first cultivated in the Indus valley of Pakistan was woven as early as 2000 BC. A book written around 300 BC included a chapter on dyes. It is believed that systematic dyeing occurred in China as early as 3000 BC near the city of Xian in the Huang (Yellow) River Valley, although there is no conclusive proof. Empress Si-Ling-Chi is credited with the discovery of silk about 2640 BC. Kermes and indigo were used as dyes as early as 2000 BC. Fragments of silk have been found in the corrosive patina of bronze swords of the Shang dynasty (1523–1027 BC), but most assuredly these samples are not the oldest [1].

The New World was similarly active in developing the textile art. With help from the desert climate in the high Andes of Peru, dyed samples of wool have been preserved and recovered from burial sites. These fragments have been dated to the millennium before the Christian era. The western and southwestern regions of the United States provided homes for the Anasazi, or ancient ones, who dwelt in the region of Mesa Verde National Park in southwestern Colorado, northern Arizona and New Mexico, and eastern Utah. Again the dry climate has helped to preserve samples from these early civilizations.

Very little in the nature of large, intact textile samples has survived in Europe. Remains of a large woolen robe, the Thorsberg Robe, found in northern Germany and dated prior to 750 BC, indicate a highly developed dyeing and weaving technology [2]. Indirect evidence is more plentiful: for example, a tombstone of a purpurarius, a Roman purple dyer, was found near Parma, northern Italy, and a dyer's workshop excavated in Pompeii. This great center of the Roman Empire was destroyed by the eruption of Mount Vesuvius in 79 AD. Similar stone vats for dyeing have been excavated in the tells of Israel and in present-day Turkey.

The Dark Ages following the fall of the Roman Empire were dark indeed, with little development of the dyer's art. The robes of a number of the monastic orders were brown and black, surely a dark age. By the end of the 1300s, however, civilization began making the swift and certain strides that have led to our present level of development. In 1371, the dyers of Florence, a city famous for its Renaissance art, formed a guild, or association of like merchants and craftsmen, which lasted for 11 years. Other guilds were being formed in other centers across Europe. Some of these guilds exist to this day. The Worshipful Company of Dyers was formed in 1471 in London. One of the legacies of this guild is a Publications Trust, which has underwritten the

---

H.S. Freeman (✉) • G.N. Mock
College of Textiles, North Carolina State University,
Raleigh, NC, USA
e-mail: hfreeman@ncsu.edu

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_13, © Springer Science+Business Media New York 2012

publication of a number of books on dyeing in cooperation with The Society of Dyers and Colourists (SDC).

The art and craft of dyeing was largely passed down from father to son or from craftsman to apprentice by word of mouth and example until the early 1500s. The *Plictho* of Gioanventura Rosetti, a Venetian armory superintendent, is believed to be the first published book on dyeing. It certainly is the oldest surviving European text to have come down to us in the twenty-first century. Five known Italian editions were published between 1548 and 1672. A French edition appeared in 1716. It is interesting that no known English translation was made until 1968 [3] when Sidney Edelstein of the Dexter Chemical Company and Hector Borghetty collaborated to reproduce a facsimile of the original 1548 edition along with a complete translation into English. During his extensive travels, Rosetti collected dyeing recipes and processes used in the flourishing city states of Venice, Genoa, and Florence. He published: *Plictho de L'arte de Tentori che insegna tenger pani telle banbasi et sede si per larthe magiore come per la comune* or *Instructions in the Art of the Dyers*, *Which Teaches the Dyeing of Woolen Cloths*, *Linens*, *Cottons*, *and Silk by the Great Art as Well as by the Common* or simply *Instruction in the Art of Dyeing*. The book was divided into four sections: the first and second sections were devoted to the dyeing of wool, cotton, and linen; the third to the dyeing of silk and the use of fugitive colors; and the fourth to the dyeing of leather and skins. Approximately 160 complete recipes were preserved in the first three sections. Edelstein and Borghetty labored diligently in determining the meanings of terms in recipes written in the dialect of sixteenth century Italy.

Synthetic Dyes

The father of modern synthetic dyes was William Henry Perkin (1838–1907), who synthesized mauve, or aniline purple, in 1856. The story behind this great story bears telling. William's father was a builder who wanted him to become an architect, but like many others, Perkin did not follow his father's chosen profession. Perkin studied at the City of London School where he became interested in chemistry at the age of 12. A teacher, Mr. Hall, gave him work in the laboratory, which in turn, inspired Perkin to follow his natural curiosity. At age 15, Perkin entered the Royal College of Science and listened to the lectures of the great German chemist, August Wilhelm von Hofmann (1818–1892). He was granted an assistantship under von Hofmann at age 17. Because his work did not allow time for his own research, he set up a separate laboratory at home and it was there that he discovered aniline purple, the first dyestuff to be commercially produced. Another dye, based on naphthalene, and prepared in collaboration with Arthur H. Church, actually preceded aniline purple, but was not commercially produced before aniline purple. Aniline purple was discovered at this home during Easter vacation while looking for quinine, an antimalarial drug. After oxidizing aniline with potassium dichromate and getting a black precipitate, extraction with ethanol gave a brilliant purple solution. Almost immediately, he sent a sample of this dye to a dyer in Perth with a request to dye silk fabric. The dyer's report read: "If your discovery does not make the goods too expensive, it is decidedly one of the most valuable that has come out for a long time." Trials on cotton were not as successful because the need for a mordant was not realized. Perkin later reported, "The value of mauve was first realized in France, in 1859. English and Scotch calico printers did not show any interest in it until it appeared on French patterns, although some of them had printed cloth for me with that colour." [4]

Table 13.1 Dyes in order of discovery

Basic: Mauve or Aniline Purple, Perkin, 1856; Fuchsin, Verguin, 1859
Acid: Alkali Blue, Nicholson, 1862
Vat: Alizarin, Gräbe and Liebermann, 1868
Indigo, von Bayer, 1880, discovered structure and synthesized indigo
Indanthrene, Böhn, 1901
Direct: Congo Red, Böttiger, 1884
Direct Developed: Primuline coupled with beta-naphthol, Green, 1887
Sulfur: Vidal, 1893
Azoics: Zitscher and Laske, 1911
Disperse: A simple azo dye, 1920s
Phthalocyanine: Linstead and Diesbach, 1928–1929
Reactive: commercialized, I.C.I., Rattee and Stevens, 1956

Since that beginning, thousands of dyes have been synthesized; some 1,500–2,000 are commercially successful today. Until 1884, however, all synthetic dyes required a mordant to give acceptable wash fastness on the textile substrate. In 1884, Böttiger produced Congo Red, which could dye cotton directly without a mordant. These dyes were commonly called direct dyes. In order to improve washfastness, the path taken in some synthetic dye chemistry was to build the dye from two or more components, directly in the fiber, or in situ. By building a large molecule without solubilizing groups within the fiber, washfastness was markedly improved. The first practical development along these lines was by A. G. Green who synthesized primuline, a dye that because of poor light fastness was not commercially important but later led the way to many important commercial dyes. Table 13.1 lists a number of classes of dyes along with the date of discovery. Worldwide, 80% of all dyes go into textiles and 20% into paper, leather, food, and the like.

The Development of the US Dyestuff Industry

The natural dyes industry was more than just a cottage industry in Colonial America. Indigo was a very important cash crop in South Carolina among the coastal islands and for some distance inland. Plantations existed well into the early 1900s despite the growth of the synthetic dyestuff industry.

The modern synthetic dye industry in the United States dates from World War I. However, in 1864, Thomas Holliday of Great Britain, and in 1868, the Albany (NY) Aniline Company with participation of Bayer of Germany began coal-tar dye manufacture. In the early 1900s, most synthetic dyes used in the United States were imported from Germany and Switzerland. With the outbreak of World War I, the British naval blockade of Germany prevented export of dyes from Europe. In spite of the blockade, the German submarine, *Deutschland*, ran the British blockade and sailed into American ports twice with dyestuffs and drugs. The Germans needed critical war material and export moneys; the United States and others needed dyes. Ironically, in 1914, German dyes were used by French dyers to dye the official French Army uniforms. The outbreak of war and ensuing blockade showed the United States how important dyes were to the American economy. Several companies began investigative work that would lead to dye synthesis; they found that dyestuffs were very difficult to make; the chemistry was much more complex than imagined. A real boost to the US industry came after World War I, when the German patents were given over to the Allies via the Alien Property Custodian. According to Lehner, DuPont reportedly spend $43 million, a tremendous sum of money in the early twentieth century, before ever showing a profit [5]. Obviously, only financially strong companies could afford to enter the business. The early pioneers included Allied (formed by merging five companies), American Cyanamid, and DuPont, to name only a few who survived to become major factors later in the twentieth century. In 1938, others included Dow, German-owned General Aniline and Film (GAF), and Swiss-owned Cincinnati Chemical Company (Ciba, Geigy, and Sandoz). In the 1960s, 50–60% of all US manufacturing was in the hands of four principal US companies:

- Allied Chemical (later sold to Bayer of Germany and to independent investors as Buffalo Color, 1977)
- American Cyanamid
- GAF, the result of the break-up of the German cartel, I. G. Farben, which was nationalized during World War II and sold to BASF in 1978
- DuPont (sold in 1980–1981 with various lines going to Crompton & Knowles, Ciba, and Blackman-Uhler)

Today, there are at least 42 dyestuff manufacturers, distributors, and repackaging agents in the United States [6]. Of the major companies, Swiss and German-based companies tend to dominate the US market. Those companies include:

- Ciba, formerly Ciba-Geigy, Switzerland
- Clariant (split off from Sandoz in 1995), Switzerland
- DyStar, formed by the merger of Hoechst and Bayer (1995), acquisition of BASF textile colors which include the former Zeneca, and Mitsubishi of Japan, Germany
- Crompton & Knowles sold to Yorkshire Group Plc and is now Yorkshire Americas

There is no major surviving the US-based company. Raghavan [7] and Mock [8] give an interesting description of these mergers and the reasons behind them.

Most of the international companies have limited manufacturing facilities in the United States and major facilities in other countries where environmental laws are not as stringent or where the parent companies have a modern integrated low-pollution facility. These facilities in the USA minimize the tariffs paid and also allow quicker response to the marketplace. Ciba has a manufacturing facility in St. Gabriel, LA; Clariant in Martin, SC; DyStar near Charleston, SC; and Yorkshire Americas at Lowell, NC.

Today well over 1,500 dyes are produced in commercial quantities, although only a select handful in each class are the true “workhorse” colors found in virtually every dyehouse dyeing a particular substrate for a particular end-use. Approximately two-thirds of the dyes and pigments consumed in the United States are used by the textile industry. One-sixth of the dyes and pigments are used for coloring paper, and the rest are used chiefly in the production of organic pigments and in the dyeing of leather and plastics.

Dyes are catalogued and grouped under a set of rules established by the Colour Index committee, consisting of representatives from the SDC, Bradford, England, and the American Association of Textile Chemists and Colorists (AATCC), Research Triangle Park, NC. Table 13.2 shows how over 9,000 dyes are enumerated in the current Index [9]. The Colour Index, now in its fourth edition, is updated periodically with newly released information and is available in book form and on CD-ROM. Volumes 1–3, published in 1971, contain the C. I. name and number, chemical class, fastness properties, hue indication, application, and usage. Volume 4, also published in 1971, contains the structures of all disclosed structures. Volume 5, last published as part of Volume 9 in 1993, contains the commercial names of all known dyes and pigments. Volumes 6–8 are supplements with updates to information in Volumes 1–4 up to 1976, 1981, and 1987, respectively. An online version of the Colour Index containing a significant number of new entries was made available in 2002.

In addition to the Colour Index, AATCC publishes a *Buyer’s Guide* annually in July [6]. Part A lists dyes, pigments, and resin-bonded pigment colors available from companies who choose to list this information.

The textile industry uses a large number of dyestuffs from each of the dye categories, the choice depending on the shade, fiber, and dyeing process, end-use of the textile product, requirements for fastness, and economic considerations.

Table 13.2 Dyes listed in the Color Index

Shade	Yellow	Orange	Red	Violet	Blue	Green	Brown	Black	Sum
Acid	255	178	439	131	356	121	453	236	2,169
Direct	171	121	263	104	303	100	242	189	1,493
Disperse	243	155	371	102	371	9	27	31	1,309
Reactive	188	123	252	44	250	27	48	46	978
Pigment	194	69	262	50	74	52	42	32	775
Basic	108	68	115	50	164	6	23	11	545
Vat	48	29	61	17	74	44	84	65	405
Sulfur	23	5	14	–	20	37	96	18	213
Mordant	65	47	95	60	82	36	92	96	573
Natural	26	4	34	–	–	5	13	–	82
Solvent	174	105	232	49	134	33	58	52	837
Food	15	8	17	–	5	4	3	3	55
Total									9,475

To provide an understanding of the interrelationships that exist among the various dye classes and fiber types, a brief survey of the major fibers follows.

Textile Fibers

In this survey, commercially important textile fibers are grouped by their origin. First there are the natural fibers from plant sources, cotton and flax, and those from animal sources, wool and silk. A second group consists of those fibers that are regenerated or chemically modified natural materials, the rayon and acetate fibers. The final group consists of synthetic fibers, which include polyester, nylon, acrylics, polyolefins, and elastane.

Natural Fibers

Cotton. Cotton fibers comprise mainly cellulose, a long-chain polymer of anhydroglucose units connected by ether linkages. The polymer has primary and secondary alcohol groups uniformly distributed throughout the length of the polymer chain. These hydroxyl groups impart high water absorption characteristics to the fiber and can act as reactive sites. The morphology of the cotton fiber is a complex series of reversing spiral fibrils. The fiber in total is a convoluted collapsed tube with a high degree of twist occurring along the length of the fiber. This staple fiber occurs in nature in lengths of ½ to 2 in., depending on the variety and growing conditions. The diameter ranges from 16 to 21 μm (1 μm is 1×10^{-6} m).

Flax. Flax is also a cellulosic fiber but has a greater degree of crystallinity than cotton. The morphology of flax is quite different from that of cotton. Flax fibers have a long cylindrical shape with a hollow core. The fibers range in length from ½ to 2½ inches, with a diameter of 12–16 μm. Flax staple comprises bundles of individual fibers. Historians believe that flax was among the first fibers to be used as a textile fiber. In recent years, its commercial importance as a textile fiber has decreased significantly.

Wool. Wool fibers comprise mainly proteins: the polypeptide polymers in wool are produced from some 20 alpha-amino acids. The major chemical features of the polypeptide polymer are the amide links, which occur between the amino acids along the polymer chain, and the cystine (sulfur to sulfur) cross-links, which occur in a random spacing between the polymer chains. The polymer contains many amine, carboxylic acid, and amide groups, which contribute in part to the water-absorbent nature of the fiber.

The morphology of wool is complex. There is an outer covering over the fiber, the cortical. There are also overlapping scales having a ratchet configuration that causes shrinkage and felting. The coefficient of friction in wool fibers is vastly different between the tip and the root, depending on which way the scales point. Wool can be made washable by chemically abrading the scales or coating the fibers with another polymer.

Wool fibers are not round but are oval in cross-section. The cortical cells constitute the major component of the fiber, and are aligned along the axis of the fiber. There is a medulla section at the center region of the fiber. Each fiber has a bicomponent longitudinal crystalline arrangement. One side of the fiber contains alpha-keratin crystalline regions, and the other contains beta-keratin crystalline regions. Alpha-keratin and beta-keratin have different moisture absorption characteristics, and this difference is what gives wool fibers crimp and springiness. It is also the reason why wool fibers kink in conditions of high humidity.

Wool fibers are sheared from about 30 major sheep breeds. The length of the wool fibers varies from 1 to 14 in. and depends on the breed, the climate, and the location on the sheep's anatomy. The fibers can be very fine to very coarse, ranging from 10 to 50 μm in diameter. The longer, coarser fibers normally are used for woolen fabrics, whereas the shorter finer ones are used for worsted fabrics.

Silk. Silk, like wool, is a protein fiber, but of much simpler chemical and morphological makeup. It comprises six alpha-amino acids, and is the only continuous-filament natural fiber. Historians claim that silk was discovered in China in 2640 BC. Silk fiber is spun by the silkworm as a smooth double-strand, each part having a trilobal cross-section. This configuration helps give silk its lustrous appearance. The fiber is unwound from the cocoon the silkworm spins as it prepares its chrysalis. The filaments are smooth and have no twists in their length, which can vary from 300 to 1,800 yards. The diameter of silk is very fine, ranging from 2 to 5 μm. Because of the labor-intensiveness of sericulture and subsequent preparation of the fiber, silk remains a luxury fiber.

Regenerated Fibers

Rayon. Viscose rayon, like cotton, comprises cellulose. In the manufacturing process, wood pulp is treated with alkali and carbon disulfide to form cellulose xanthate. Subsequently, the reaction mass is forced through a spinneret and precipitated in an acid coagulation bath as it is formed into a continuous filament. The fiber has a round striated cross-section. Rayon staple is made by "breaking" the continuous strands into staple-length fibers. Viscose rayon is conventionally produced in diameters varying from 9 to 43 μm.

Acetate. Triacetate and diacetate fibers are manufactured by the chemical treatment of cellulose obtained from refined wood pulp or purified cotton lint. Most of the hydroxyl groups are acetylated (esterified) by treating the cellulose with acetic acid. This determines the chemical configuration of triacetate. Acetate or diacetate is made by the saponification of one of the acetylated groups, thus restoring a hydroxyl to each cellulosic monomer unit. Theoretically, then, diacetate has two acetylated groups in each glycoside unit. The conversion of the hydroxyl groups causes these fibers to be hydrophobic and changes the dyeing characteristics drastically from those of the normal cellulosic fibers. Triacetate fibers are spun by mixing the isolated reaction product (flake) with methylene chloride and alcohol. The spinning solution (dope) is forced through a spinneret and dry-spun into continuous filaments.

An alternate way of wet spinning is also possible. Acetate fibers are spun by mixing the isolated reaction product with acetone and water. The spinning solution is formed into filaments by evaporating the solvent and coagulating the acetate in a manner similar to that for triacetate (i.e., by the dry-spinning method).

Synthetic Fibers

Nylon. In 1939 the DuPont Company introduced the first truly synthetic textile fiber. Dr. Wallace Carothers invented nylon as a result of his basic research into polymer science. Chemically, nylon is a polyamide fiber. The two major types of nylon polymer are used in textiles: type 6,6 which is made by using hexamethylene glycol and adipic acid, and type 6, which is made by polymerizing ε-caprolactam. Nylon fibers are made by melt-spinning the molten polymer. The result is a continuous filament fiber of indeterminate length. It is spun in many deniers, with its diameter varying from 10 to 50 μm. The cross-section usually is round, trilobal, or square with hollow channels when used as carpet fiber.

Polyester. Polyester is made by the polymerization reaction of a diol and a diester. The main commercial polymer is formed by a condensation reaction using ethylene glycol and terephthalic acid. Fibers are formed by melt-spinning. Commercially introduced in 1953 by the DuPont Company as Dacron, polyester fibers have high strength, and very low moisture absorbance. The fiber is usually spun with a round cross-section. Polyester is the most-used synthetic fiber around the world.

Acrylics. The DuPont Company introduced the first commercial acrylic fiber, Orlon, in 1950. Acrylics are made from the polymerization of acrylonitrile and other comonomers to allow dyeability and to open the internal structure. The fibers are produced by either solvent-spinning (Orlon), or wet-spinning (Acrilan). In the solvent-spinning process, the polymer is dissolved in a low-boiling liquid solvent such as dimethyl formamide and extruded in a warm air chamber. In wet-spinning, the polymer is dissolved in a suitable solvent, extruded into a coagulation bath, dried, crimped, and collected. Although the acrylic fibers are extruded as continuous filaments, they subsequently are cut into staple-length fibers. Acrylics have found a niche market as a substitute for wool or in wool blends (blankets, sweaters, etc.) and in awnings and boat covers. The cross-section of the filament varies among manufacturers, Orlon having a dog-bone configuration and Acrilan having a lima-bean shape. Acrylic fibers are quick drying and wrinkle resistant.

Polyolefins. Polyolefin fibers are produced from the polymerization of ethylene or propylene gas. The catalysis research of Ziegler and Natta led to the development of these polymers to form crystalline polymers of high molecular weight. Hercules Inc. produced the first commercial fibers in 1961. The fibers made from these polymers are melt-spun. The cross-sections are round, and the fibers are smooth. They have extremely low dye affinity and moisture absorbance. Colored fiber is normally produced by mixing pigments in the melt polymer prior to extrusion.

Elastane. The DuPont Company commercialized the first manufactured elastic fiber, Lycra, in 1958. Originally categorized as a spandex fiber, the name “elastane” has become more common around the world. This specialty fiber is described as a segmented polyurethane that contains “hard” and “soft” segments; their ratio determines the amount of stretch built into the fiber. Elastane fibers are formed by dry-spinning or solvent-spinning. The continuous filaments can be coalesced multifilaments or monofilaments, depending on the manufacturer. Because most dyeings are applied from water solutions or dispersions, the effect of water absorption by the fiber is an important criterion. Table 13.3 shows the hydrophobic/hydrophilic characteristics of the important fibers. The cellulosic and natural fibers are the most hydrophilic, and polyolefin is the most hydrophobic.

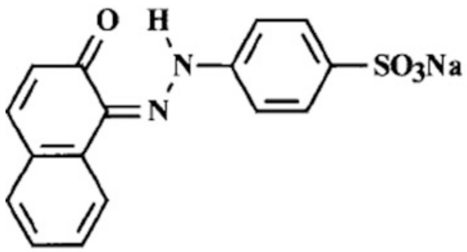
Table 13.3 Hydrophobic/hydrophilic characteristics of various fibers

Fiber	Moisture content, % (at 65% R.H. −70 °F)	Water retention, % (weight change in water)	Swelling, % (volume change in water)
Acetate	6.5	20–25	NA
Acrylic	1.0–1.5	4.5–6	2
Cotton	7	45	45
Elastane	0.3–1.5	Low	Low
Nylon	3.0–5.0	9–12	13
Polyester	0.3–0.5	3–5	0.5
Olefin	0.01–0.1	Very low	Very low
Triacetate	3.2	12–18	NA
Viscose	13	90–100	95
Wool	13–15	42	42

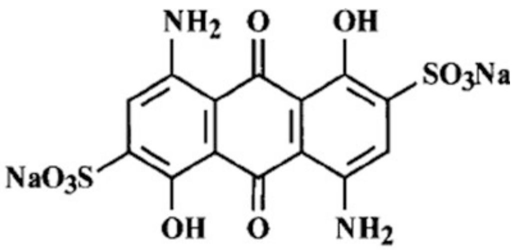
Microdenier fibers. The first commercial production of microfiber in the United States was in 1989 by the DuPont Company. Today microfibers are produced in a variety of synthetic fibers (i.e., polyester, nylon, acrylic, etc.). A microfiber is a fiber that is less than one denier per filament. Yarns made from microdenier filaments are able to give silk-like hand to fabrics.

Dye Classification

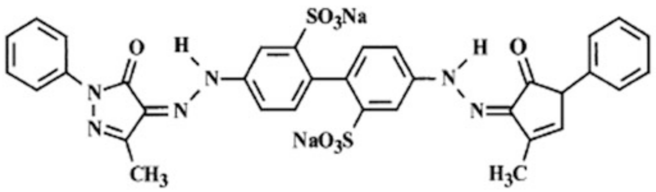
This section covers structural features that govern the classification and application of various dye classes. In this regard, the chemistry of acid, azoic, basic, direct, disperse, reactive, sulfur, and vat dyes is presented. With regard to the application of synthetic yes to textiles, it is well known that dyeing of textile fibers from an aqueous dyebath involves four steps: exhaustion, diffusion, migration, and fixation. In step 1, individual dye molecules move from the dyebath to the fiber surface and in step 2, dye molecules move from the fiber surface into the amorphous regions of the fiber. In step 3, dye molecules move from regions of high concentration to regions of low concentration



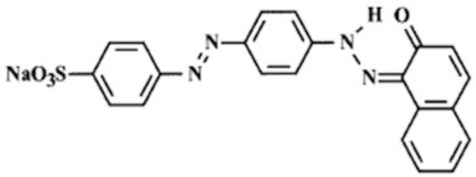
Acid Orange 7



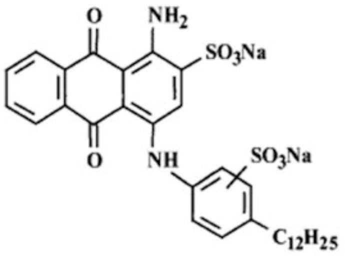
Acid Blue 45



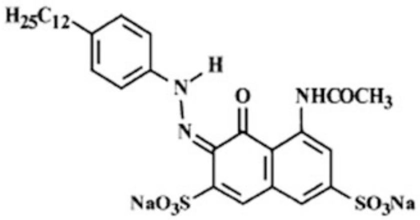
Acid Yellow 42



Acid Red 151



Acid Blue 138



Acid Red 138

(i.e., migrate) to become uniformly distributed within the polymer matrix. In step 4, dye molecules interact with groups along the polymer chain via primary or secondary valency forces. Dye–polymer interactions can involve ionic bonding (e.g., acid dyes on nylon or wool), covalent bonding (e.g., reactive dyes on cotton), mechanical entrapment (e.g., vat dyes and sulfur dyes on cotton), secondary valency forces (direct dyes on cotton), or solid–solid solution (e.g., disperse dyes on polyester).

Acid Dyes

Acid dyes derive their name from the conditions associated with their application, in that they are typically applied to textile fibers from dyebaths containing acid [9]. Most acid dyes have one or two sodium sulfonate ($-SO_3Na$) groups and, therefore, are water soluble and capable of bonding with fibers having cationic sites (cf. Fig. 13.1). They give a wide range of bright colors on textiles, especially when monoazo and anthraquinone structures are used.

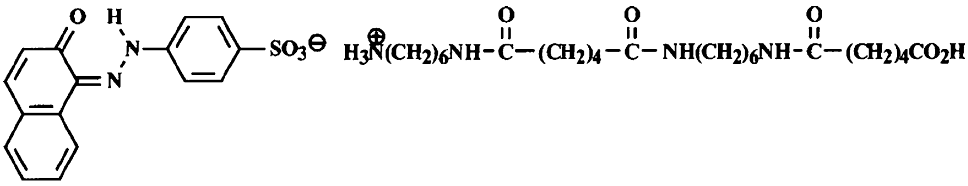
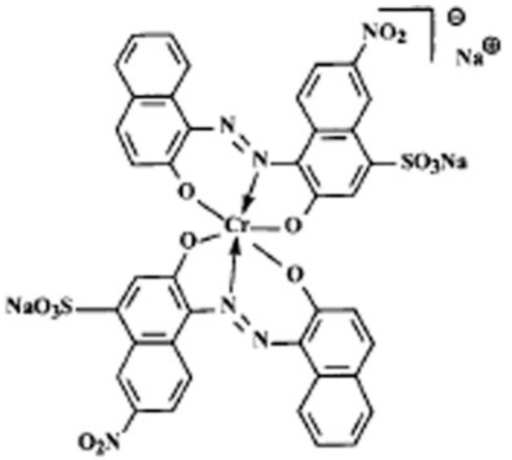
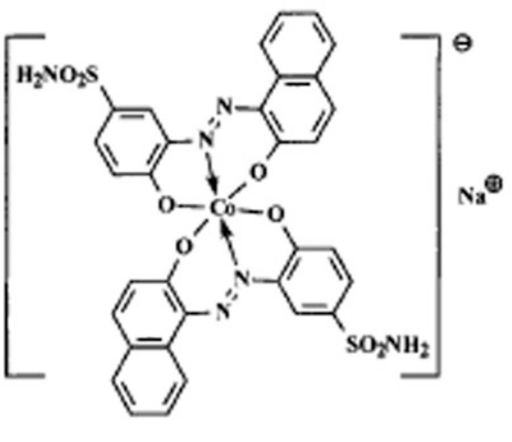


Fig. 13.1 Ionic bond formation between nylon 66 and C.I. Acid Orange 7

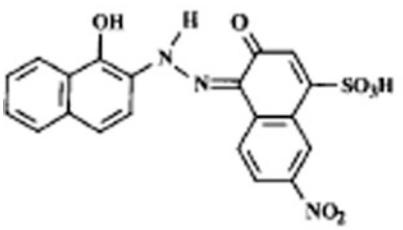
Acid dyes vary widely in molecular structure and in the level of acid required for dye application. They include relatively low molecular weight dyes such as C.I. Acid Orange 7 and C.I. Acid Blue 45, both of which are readily applied to polyamide and protein fibers and are known as *level dyeing acid dyes*. As the name suggests, these dyes are characterized by good migration and, therefore, readily produce level dyeings with time. In addition they give reasonably good lightfastness and barré coverage. The application of level dyeing acid dyes to nylon and wool utilizes weak acid and strong acid, respectively. For applications requiring good washfastness, *milling acid dyes* or *supermilling acid dyes* are employed. Both of these dye types afford relatively poor barré coverage, however. The former type dyes are applied from weakly acidic dyebaths whereas the latter are generally applied at neutral pH, with molecular size increasing as acid strength decreases. Examples of milling acid dyes are C.I. Acid Yellow 42 and C.I. Acid Red 151, and supermilling acid dyes include C.I. Acid Blue 138 and C.I. Acid Red 138.



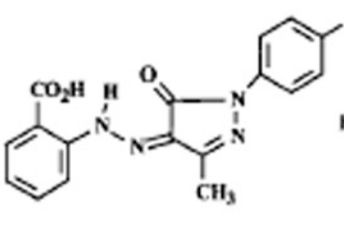
Acid Black 172



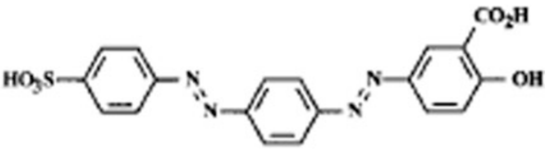
Acid Red 182



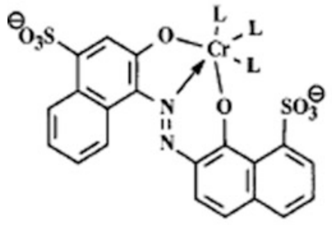
Mordant Black 11



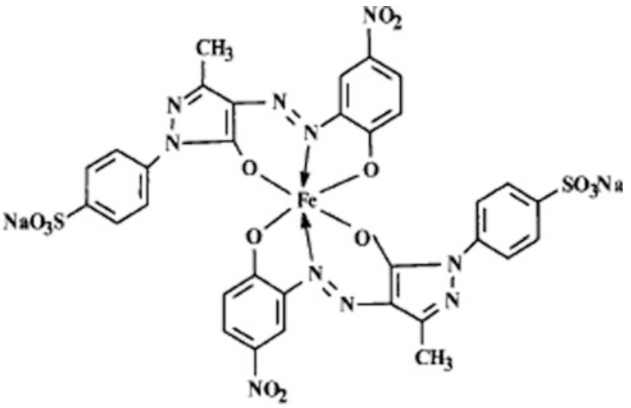
Mordant Yellow 8



Mordant Orange 6



Acid Blue 158



Acid Brown 98

Because it is well known that azo dyes derived from naphthol and pyrazolone intermediates exist predominantly in the hydrazone form, this tautomeric form is given for Acid Yellow 42, Acid Red 151, Acid Red 138, and for the appropriate dyes that follow in this chapter.

Acid dyes include metal-complexed azo structures, where the metals used are cobalt, chromium, and iron [10, 11]. Examples are 1:1 and 2:3 chromium complexes and 1:2 cobalt complexes, where the numbers employed represent the ratio of metal atoms to dye molecules. Metal-complexed dyes can be formed inside textile fibers by treating suitably dyed fibers with a solution containing metal ions [12]. In this case, the metal-free forms of these azo dyes are known as *mordant dyes* and contain mainly *ortho*, *ortho′*-bis-hydroxy or *ortho*-carboxy, *ortho′*-hydroxy groups (e.g., C.I. Mordant Black 11, Mordant Yellow 8, and Mordant Orange 6). When the metal complexes are formed prior to the dye application process, the resultant dyes are known as *premetallized acid dyes* and vary in the acid strength required in the application step [13]. The 1:1 chromium complexes (e.g., C.I. Acid Blue 158) are stable only in very strong acid, making them suitable for wool but not nylon. Neutral dyeing premetallized acid dyes contain $–SO_2NH_2$ or $–SO_2CH_3$ groups in lieu of $–SO_3Na$ groups (see Acid Black 172 vs. Acid Red 182). In this case, dye–fiber fixation occurs because the combination of trivalent metal ion (e.g., Cr^{3+}) and four attached negatively charged ligands gives the complex a net negative charge.

Metallization of azo dyes enhances lightfastness, reduces water solubility, causes a bathochromic shift in color, and dulls the shade. Iron complexes generally give brown shades (e.g., C.I. Acid Brown 98) and are most often used to dye leather.

Azoic Dyes

Azoic dyes are mainly bright orange and red monoazo dyes for cotton, with dull violet and blue colors also possible [14]. They are water insoluble and consequently give high washfastness. They are also referred to as azoic “combinations” rather than “dyes” because they do not exist as colorants until they are formed inside the pores of cotton fibers [15]. They are quite important for printing on cotton and often give good lightfastness in heavy depths. Their bleachfastness is better than direct and sulfur dyes and good crockfastness requires efficient soaping after the application step. The formation of these dyes requires two constituents: an azoic coupling component and an azoic diazo component, examples of which are shown in Figs. 13.2 and 13.3. The azoic coupling components are

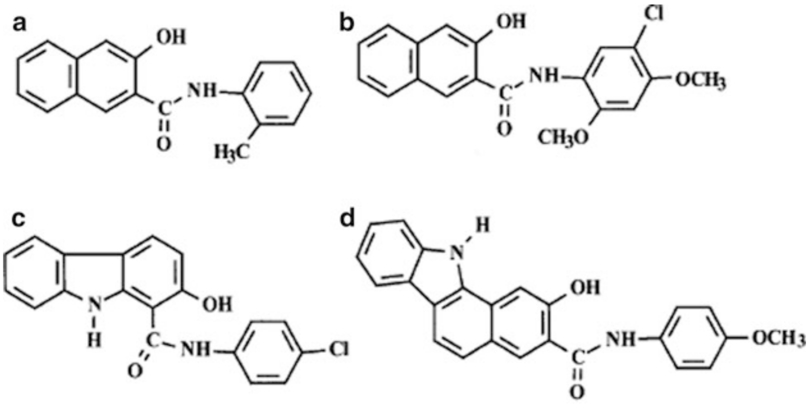


Fig. 13.2 Structures of C.I. azoic coupling components 18 (**a**), 12 (**b**), 15 (**c**), and 13 (**d**)

beta-naphthol and β-oxynaphthoic (BON) acid derivatives and the azoic diazo components are substituted anilines.

Azoic dyes are also known as naphthol dyes, because all employ a naphthol component in their formation, and they can be produced in batch or continuous processes. Because they have a limited shade range, they are best known for their ability to provide economical wetfast orange and red shades on cotton. A generic azoic dye structure is shown in Fig. 13.4.

Basic or Cationic Dyes

Basic dyes were developed to dye negatively charged acrylic fibers, forming ionic bonds in the fixation step (Fig. 13.5) [16]. They owe their name to the presence of aromatic amino (basic) groups, and in this case a cationic amino group is present. Generally, they have excellent brightness and color strength, especially among the triarylmethane types. However, their lightfastness is often low, when they are applied to fibers other than acrylics. Basic dyes include those containing a fixed cation, examples of which are C.I. Basic Blue 22, and C.I. Basic Red 18. The triarylmethane dye C.I. Basic Violet 3 has a mobile cation that produces resonance structures of comparable energy.

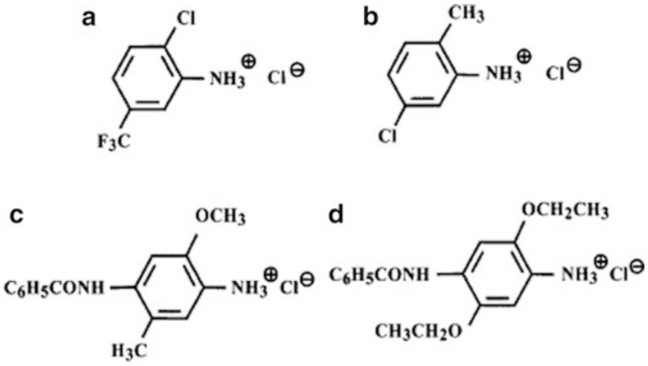


Fig. 13.3 Structures of C.I. azoic diazo components 49 (**a**), 32 (**b**), 41 (**c**), and 20 (**d**)

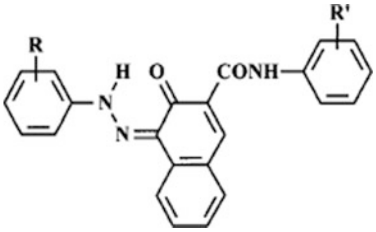


Fig. 13.4 Generic structure for azoic dyes, where R and R′ = alkyl, alkoxy, halo, and nitro groups

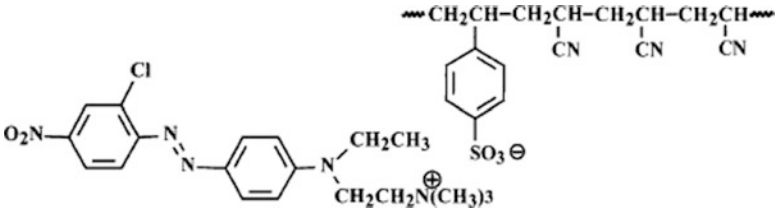
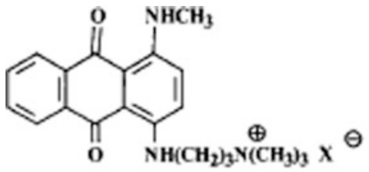
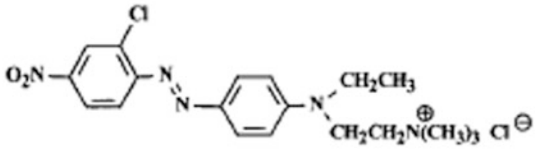


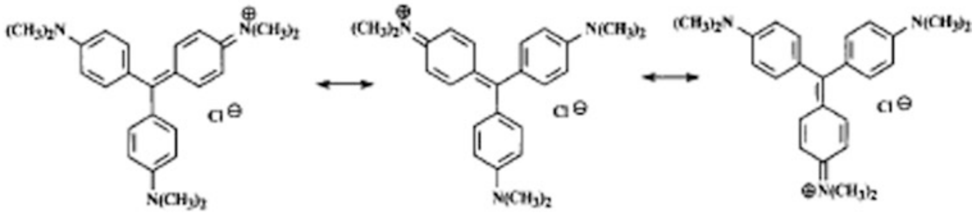
Fig. 13.5 Ionic bond formation between polyacrylonitrile and C.I. Basic Red 18



Basic Blue 22



Basic Red 18



Basic Violet 3

Basic dyes are applied from weakly acidic dyebaths ($pH = 4.5–5.5$) and often require the use of anionic or cationic retarding agents to control the rate of dye strike and give level dyeing. Suitable retarding agents either form a weak bond with dyesites along the polymer chain or interact with the dye in the dyebath. In the former case, a significant fraction of the cationic retarder employed is displaced by the dye as dyeing progresses because the dye has higher affinity for the fiber. When anionic retarding agents are used, the dye–retarder bond is broken by increasing the dyebath temperature, giving controlled release of dye molecules to facilitate leveling.

To help determine which basic dyes can be combined for shade matching, key dyebath parameters have been developed [17]. The first parameter pertains to the dyes themselves and is known as the *combinability constant* (k). This value provides a measure of how fast a basic dye will dye the fiber, and the dyes are rated on a scale of 1 (fast) to 5 (slow). The second parameter pertains to the fiber type involved and is known as the *fiber saturation value* (S_F). This value provides an indication of how much dye the fiber will hold at the saturation point. In this regard, the dye used is C.I. Basic Green 4 and typical saturation levels are 1.0–4.0% based on the weight of the fibers (*owf*) for light to deep dyeing. The third parameter is the *dye saturation factor* (f), which is a measure of the capacity of a basic dye for saturating a fiber. This factor is influenced by the molecular size and purity of the dye. In this case, the goal is to avoid placing more dye on the fiber than the number of dye sites, and the standard is C.I. Basic Green 4 ($f = 1\%$).

Direct Dyes

Direct dyes are anionic colorants that have affinity for cellulosic fibers [18]. They were the first dyes that could be used to dye cotton in the absence of a mordanting agent, giving rise to the term *direct-cotton dyes*. Like acid dyes, direct dyes contain one or more $–SO_3Na$ groups, making them water-soluble. Unlike acid dyes, they interact with cellulose (Cell–OH) chains via secondary valency forces (e.g., H-bonding and dipole–dipole interactions), as illustrated in Fig. 13.6. The combined effects of these rather weak forces and sulfonated structures cause direct dyes to have low intrinsic washfastness. Direct dye structures are based on four main chromophores: azo (e.g., C.I. Direct Red 81, C.I. Direct Yellow 28, and C.I. Direct Black 22), stilbene (e.g., C.I. Direct Yellow 12 and C.I. Direct Yellow 11), oxazine (e.g., C.I. Direct Blue 106 and C.I. Direct Blue 108), and phthalocyanine (e.g., C.I. Direct Blue 86 and C.I. Direct Blue 199). About 82% of all direct dyes have disazo or polyazo structures, with stilbene and monoazo structures occupying about 5% each and thiazole, phthalocyanine, and dioxazine structures covering the remaining few percent [19].

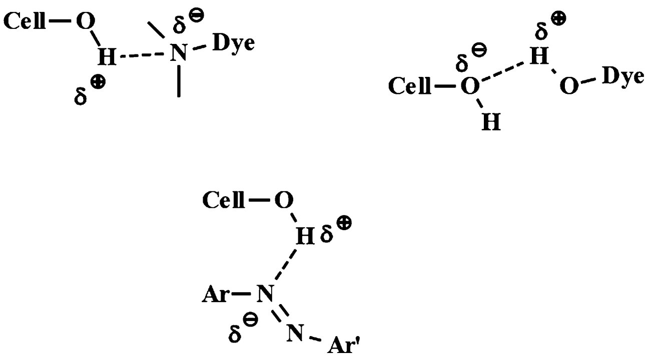
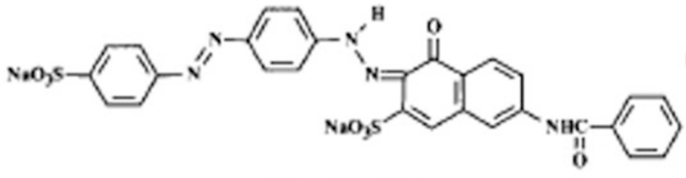
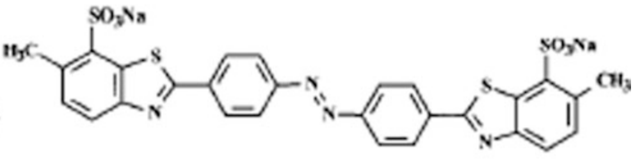


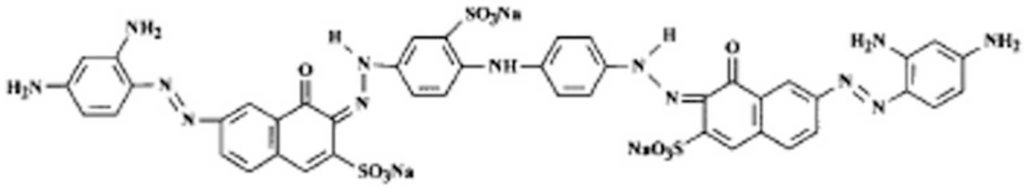
Fig. 13.6 H-bonding interactions between cellulose and polar groups in direct dye molecules



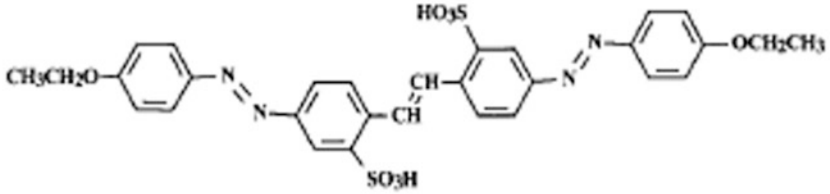
Direct Red 81



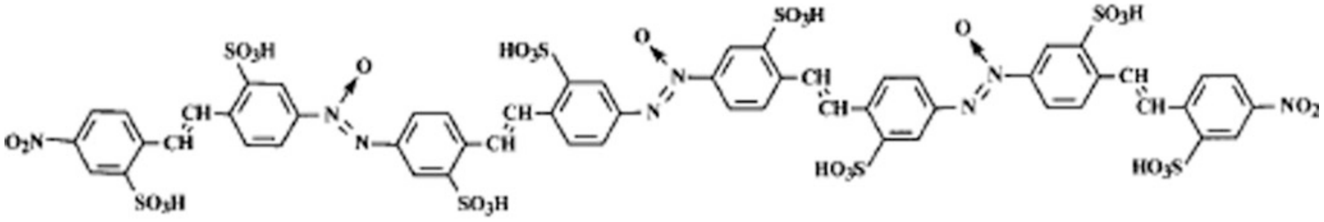
Direct Yellow 28



Direct Black 22

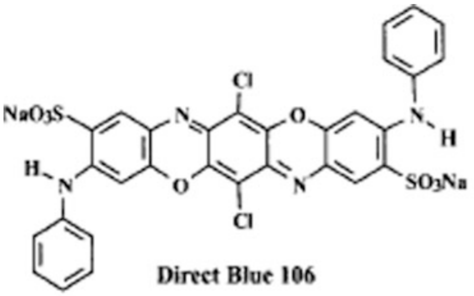


Direct Yellow 12

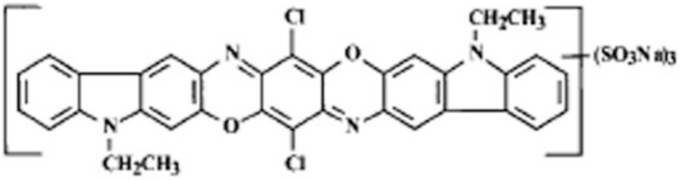


Direct Yellow 11

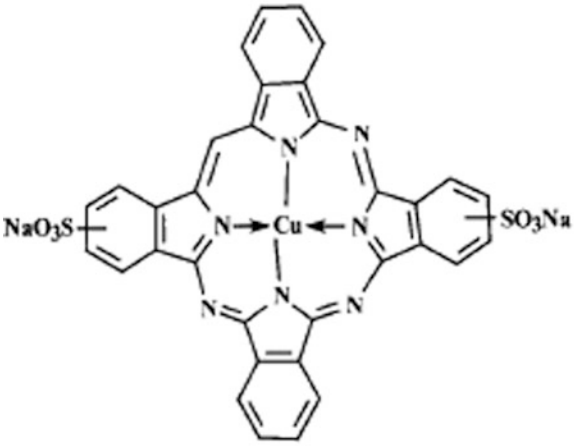
Suitably substituted direct dyes can be converted to metal complexes. In this regard, Cu is the metal of choice and examples are C.I. Direct Blue 218, C.I. Direct Red 83, and C.I. Direct Brown 95. About 5% of all azo direct dyes are metal complexes and unlike most direct dyes, these dyes have good lightfastness, as would be anticipated.



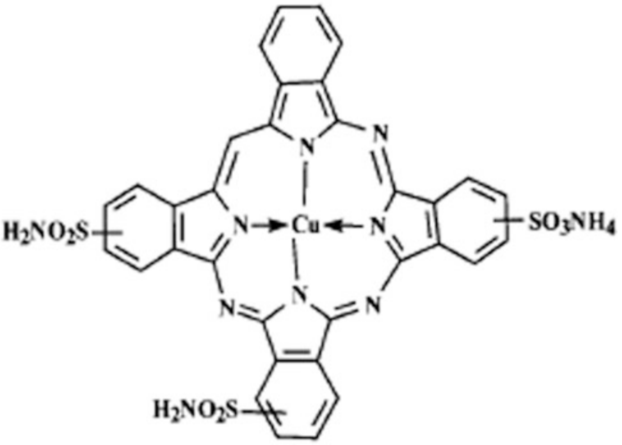
Direct Blue 106



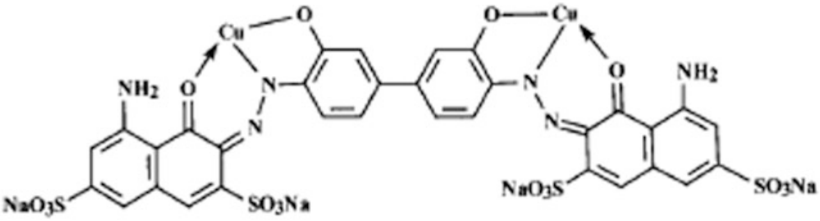
Direct Blue 108



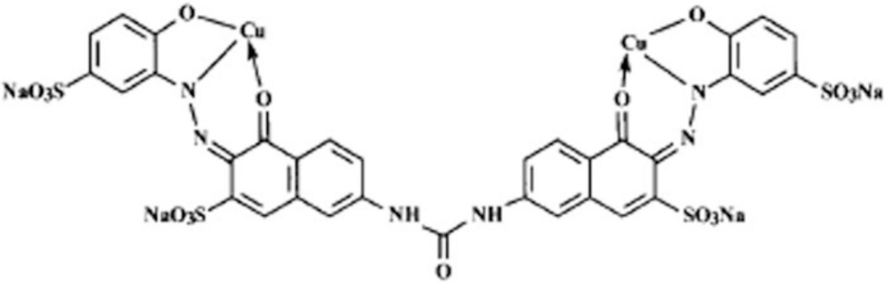
Direct Blue 86



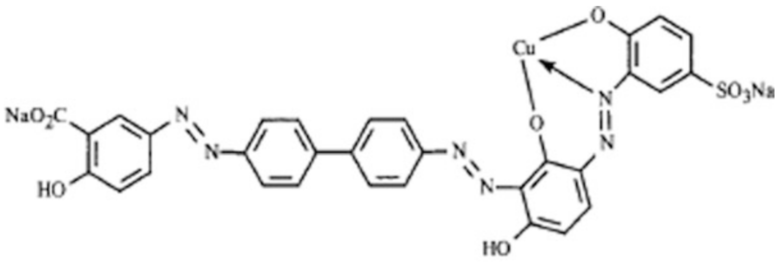
Direct Blue 199



Direct Blue 215

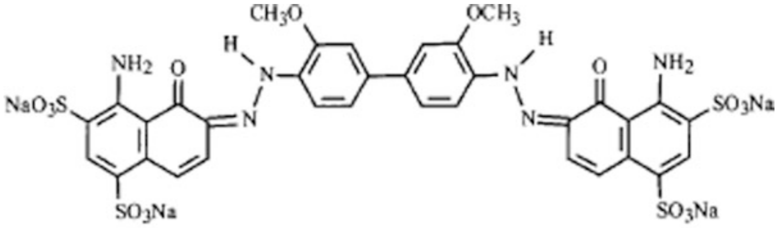


Direct Red 83



Direct Brown 95

Direct dyes are subdivided into three classes (A, B, and C), to assist the dyer in selecting appropriate combinations for color matching [20]. Class A direct dyes give good migration and leveling with time. The dyer employs 5–20% salt for their application and in this case all of the salt may be added at the beginning of the dyeing cycle. An example of this direct dye class is C.I. Direct Yellow 12. Class B direct dyes have poor migration and leveling properties and require the controlled addition of salt to afford level dyeing. They are larger than the former types and have better washfastness. An example is C.I. Direct Blue 1. Class C dyes are the largest of the direct dyes and, consequently, have the best washfastness but poorest leveling properties. Leveling requires careful control of the rate of temperature rise during the dyeing process. Some salt may be added but less than the amount used with classes A and B. An example of this dye class is C.I. Direct Black 22.



Direct Blue 1

The high temperature stability of direct dyes is an important consideration if one wishes to use these dyes as the colorant for cotton when dyeing a polyester/cotton blend at 130 °C [21]. The key to success is to choose dyes that are resistant to hydrolysis. Suitable dyes include C.I. Direct Yellow 105, C.I. Direct Orange 39, and C.I. Direct Blue 80, whereas unsuitable dyes include C.I. Direct Yellow 44, C.I. Direct Red 80, and C.I. Direct Red 83. A quick examination of the structures of the unsuitable dyes reveals that they have groups which are subject to hydrolysis. In the case of Direct Red 83, hydrolysis essentially cuts the molecule in half, eliminating fiber affinity (Fig. 13.7).

Because many direct dyes do not have good washfastness and lightfastness, their dyeing on cotton are often treated with a chemical agent, in what is commonly known as an aftertreatment process. The most widely used aftertreatment methods involve (1) cationic fixatives, (2) copper sulfate, or (3) diazotization and coupling reactions. The first and third methods are designed to enhance washfastness and are illustrated in Figs. 13.8 and 13.9. The use of cationic fixatives ties up sodium sulfonate groups, reducing the water solubility of the treated dye. Diazotization and coupling enlarges the size of the dye, making desorption more difficult, and simultaneously makes the dye less hydrophilic. This process requires the presence of at least one diazotizable primary arylamino (Ar–NH$_2$) group in the dye structure. In this two-step process, the amino group is diazotized by treatment with nitrous acid (HNO$_2$) and the resultant diazonium groups are coupled with a naphthoxide to give new azo groups. It should be pointed out that the addition of new azo groups can also affect dye color. Therefore, this process is most often used for navy and black shades, where the differences in shade variations from batch to batch are less objectionable.

Although copper sulfate aftertreatments are designed mainly to enhance lightfastness, the reduction in water solubility that accompanies Cu-complex formation can have a beneficial effect on washfastness. This treatment also dulls the fabric shade and causes a shift in dye color, so that the resultant color must be the one the dyer is seeking.

The dye used as an example in Fig. 13.10 is C.I. Direct Black 38. It is worthwhile to note that this dye is one of many that were synthesized from benzidine, an established human carcinogen [22]. Nowadays, such dyes are regarded as

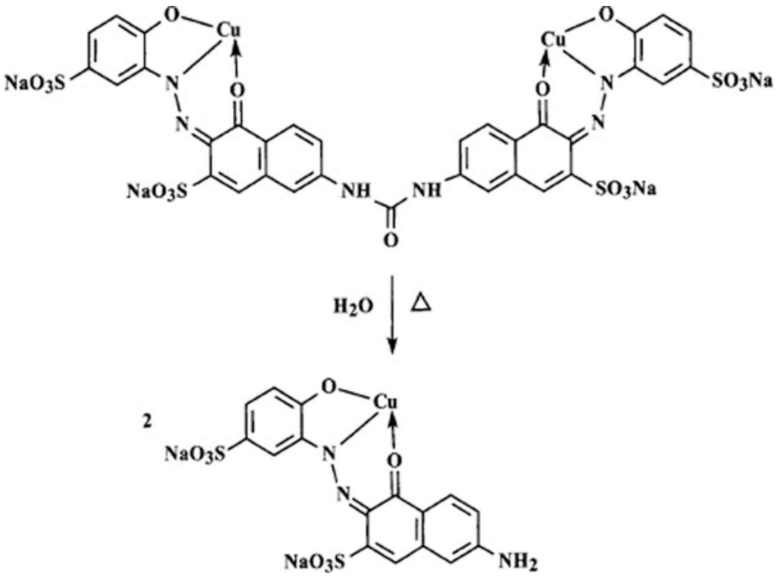


Fig. 13.7 High temperature hydrolysis of Direct Red 83

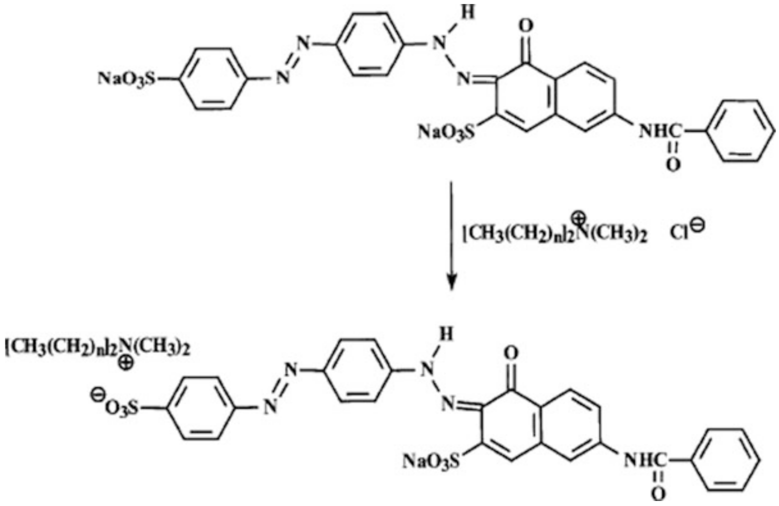


Fig. 13.8 Use of a cationic fixative to enhance direct dye washfastness

cancer-suspect agents because of their potential to generate free benzidine upon metabolic breakdown [23]. With this point in mind, regulations preventing the use of azo dyes derived from benzidine and 20 other aromatic amines in textiles have appeared [24]. This requires dye chemists to consider the genotoxicity of potential metabolites in the design of new azo dyes [25].

Disperse Dyes

Disperse dyes were invented to dye the first hydrophobic fiber developed, namely cellulose acetate, and were initially called acetate dyes [26]. The term disperse dyes is more appropriate, because these dyes are suitable for a variety of hydrophobic fibers and it is descriptive of their physical state in the dyebath. Disperse dyes have extremely low water solubility and to be applied from this medium they must be (1) dispersed in water using a surfactant (dispersing agent) and (2) milled to a very low particle size (1–3 μm). These nonionic hydrophobic dyes can be used on acetate, triacetate, polyester, nylon, acrylic, and polyolefin fibers, and their mechanism of fixation involves solid–solid solution formation.

Disperse dyes provide a wide range of bright colors on textiles and many have excellent build-up and barré coverage properties. In addition, they have good washfastness properties but their lightfastness varies with structure. They are suitable for continuous dyeing, a process that takes advantage of their sublimation properties. Disperse dye end-use applications are often based on their classification. The classification system employed is shown in Table 13.4. Low-energy disperse dyes are the easiest to exhaust under atmospheric dyeing conditions but have the lowest

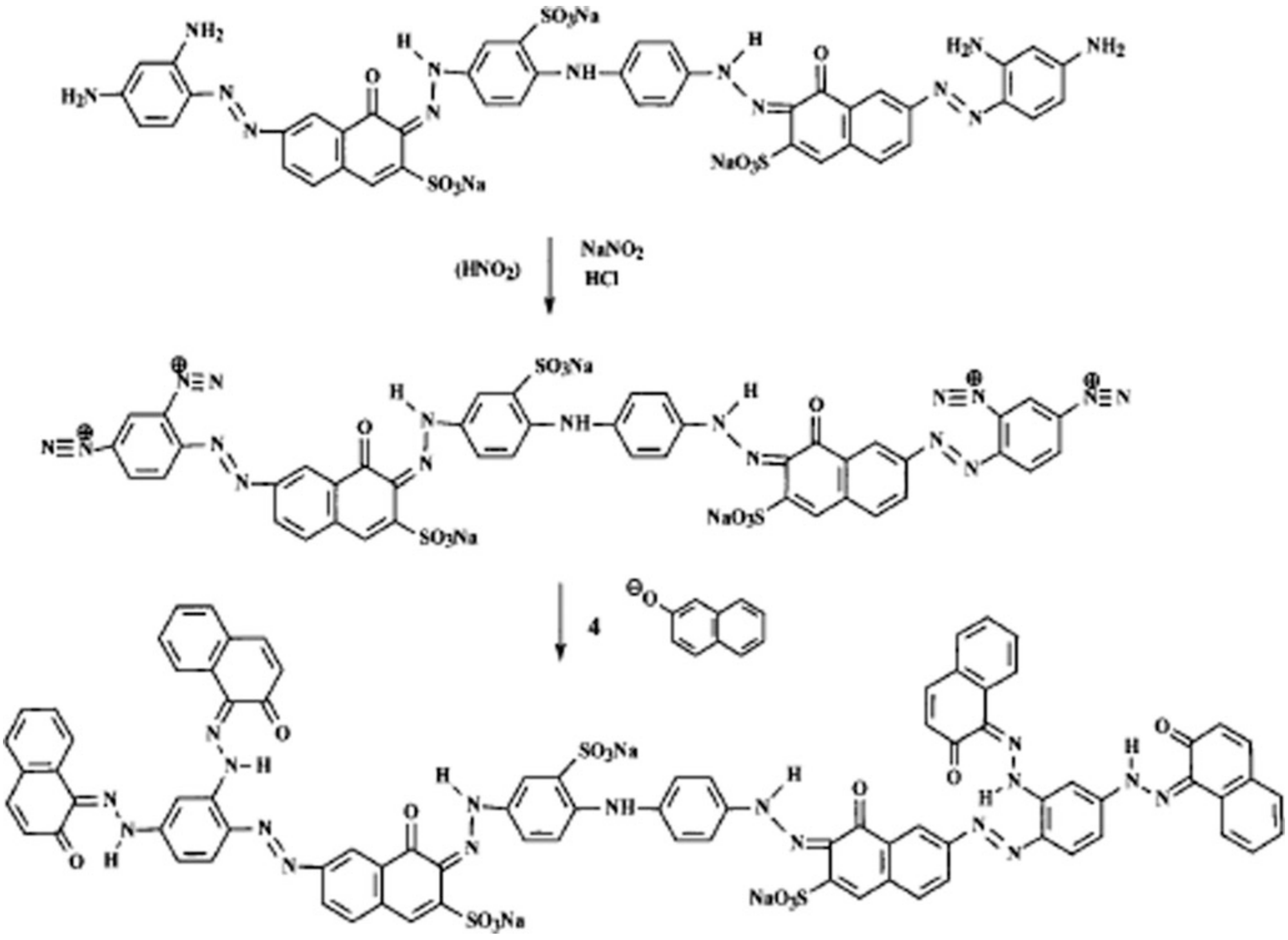


Fig. 13.9 Use of the two-step development process to enhance direct dye washfastness

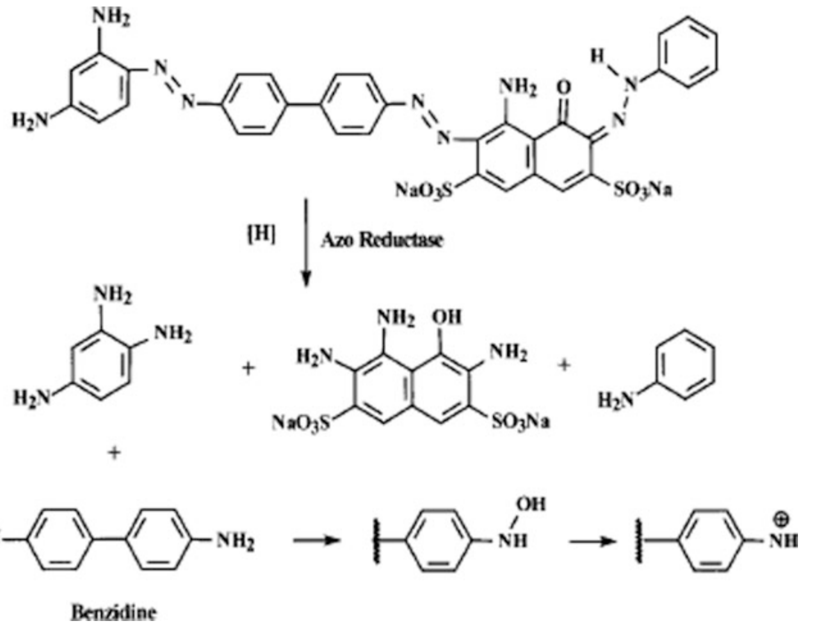
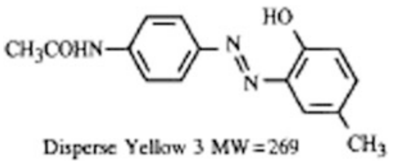
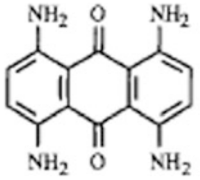
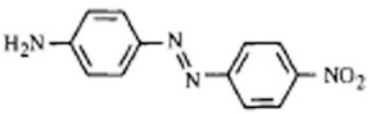
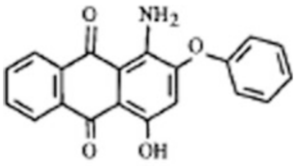
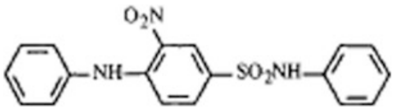
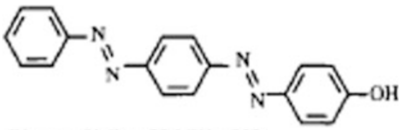
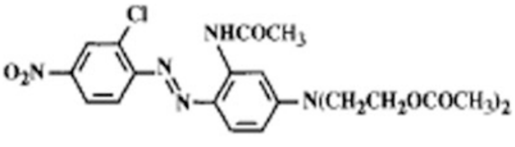
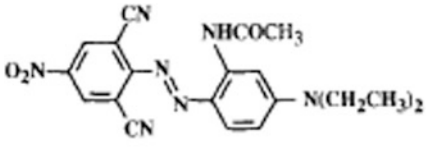


Fig. 13.10 Metabolic breakdown of a direct dye by azo reductase enzymes

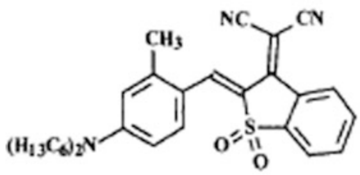
Table 13.4 Disperse dye classification and examples

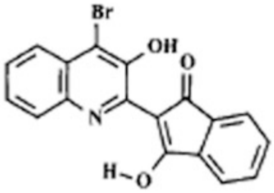
Class	Molecular mass	Sublimation temperature
Low energy	<300	<150 °C
 Disperse Yellow 3 MW=269	 Disperse Blue 1 MW = 269	 Disperse Orange 3 MW=242
Medium energy	300–400	150–210 °C
 Disperse Red 60 MW = 331	 Disperse Yellow 42 MW = 369	 Disperse Yellow 23 MW = 302
High energy	>400	>210 °C
 **Disperse Red 167:1 MW 507**	 **Disperse Blue 165 MW 405**	

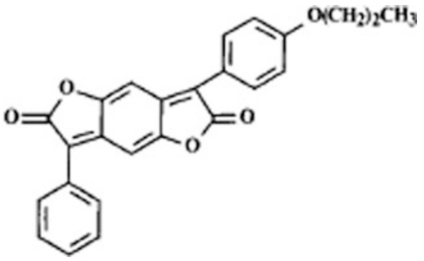
thermostability, with the latter property making them unsuitable for automotive applications. They are used to dye acetate, triacetate, and nylon fibers. On the other hand, the high energy dyes are best applied under pressure ($T = 130$ °C) and are most appropriate for polyester body cloth for automobile interiors. Medium energy dyes are also used to dye polyester and can be applied at atmospheric pressure using a carrier.

Disperse dyes vary in the type of chromophore present and include azo, anthraquinone, nitro, methane, benzodifuranone, and quinoline-based structures. Examples of the first three types are given in Table 13.4, and representative of the latter three types are C.I. Disperse Blue 354, C.I. Disperse Yellow 64, and C.I. Disperse Red 356. Most disperse dyes have azo (~59%) or anthraquinone (~32%) structures. Azo disperse dyes cover the entire color spectrum, whereas the important anthraquinone disperse dyes are mainly red, violet, and blue. The azo types offer the advantages of higher extinction coefficients ($\varepsilon_{\max} = 30{,}000-60{,}000$) and ease of synthesis, and the anthraquinones are generally brighter and have better photostability (lightfastness). The key weaknesses associated with the anthraquinone dyes are their low extinction coefficients ($\varepsilon_{\max} = 10{,}000-15{,}000$) and less environmentally friendly synthesis.

To produce disperse dyes having the brightness of the anthraquinone system and the color strength of the azo







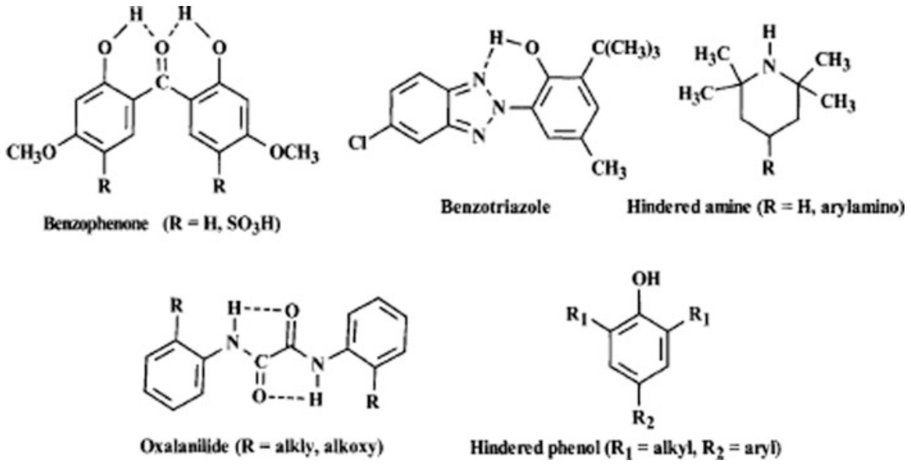
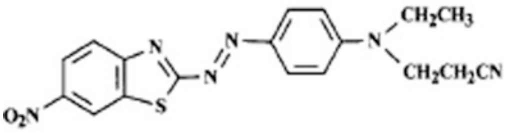


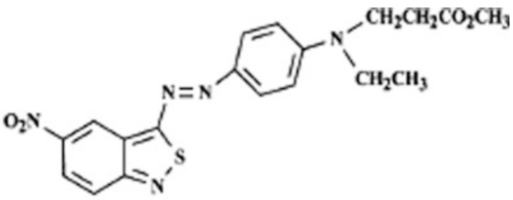
Fig. 13.11 Structures of disperse dye photostabilizers

system, azo dyes based on heteroaromatic amines were developed [27–29]. Examples are C.I. Disperse Red 145, Disperse Blue 148, Disperse Red 156, and C.I. Disperse Blue 339. These dyes employ aminated thiazoles, benzothiazoles, benzisothiazoles, and thiadiazoles in their synthesis. They have $\varepsilon_{max} = 50{,}000-80{,}000$, good brightness, and good washfastness. Another key feature of disperse dyes with heteroaromatic systems is their less complex structures. Compare, for example, the fewer number of substituent groups in the diazo component (left side of the azo bond) of Disperse Red 167 vs. Disperse Red 156 and Disperse Blue 165 vs. Disperse Blue 102. However, these dyes are more expensive than disperse dyes derived from benzeneamines, owing to their low reaction yields, and have lower lightfastness than the anthraquinone dyes.

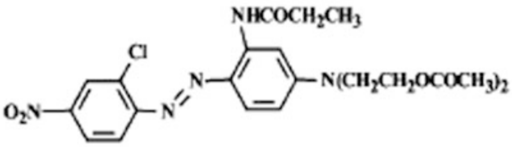
The use of disperse dyes in applications requiring high lightfastness involves the coapplication of photostabilizers. These agents enhance dye stability by quenching the excited states of disperse dyes, probably via energy transfer, or by preferential absorption (screening) of UV radiation. They are also known as UV absorbers and exhaust from the dyebath like disperse dyes. They encompass benzophenone, benzotriazole, oxalanilide, and hindered amine/phenol structures (see Fig. 13.11) [30].



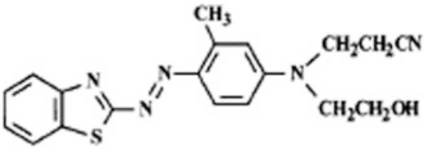
Disperse Red 145



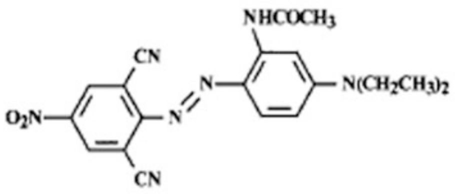
Disperse Blue 148



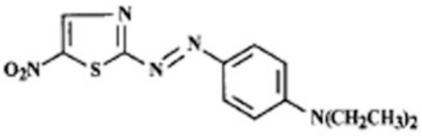
Disperse Red 167



Disperse Red 156



Disperse Blue 165



Disperse Blue 102

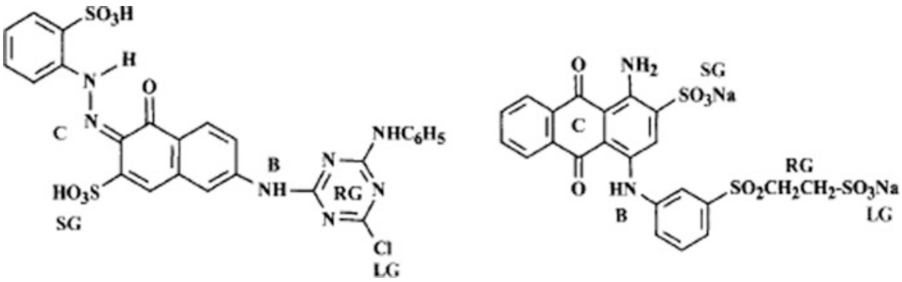


Fig. 13.12 Structures showing the basic parts of two fiber-reactive dyes

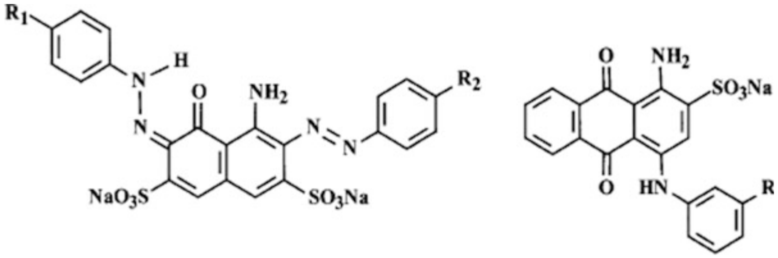


Fig. 13.13 Comparison of some acid dye and fiber-reactive dye structures. Acid Black 1: R_1 = H; R_2 = NO_2; Acid Blue 25: R = H; Reactive Black 5: $R_1 = R_2 = SO_2(CH_2)_2OSO_3Na$; Reactive Blue 19: $R_1 = SO_2(CH_2)_2\ OSO_3Na$

Reactive Dyes

Reactive dyes are used mainly as colorants for cotton, although they are also suitable for nylon and wool [31]. They are water soluble, due to the presence of one or more $–SO_3Na$ groups, and undergo fixation to polymer chains via covalent bond formation. Reactive dyes have very high washfastness and are used for leisure wear and other applications requiring stability to repeating laundering. Each dye is composed of five basic parts:

$$SG - C - B - RG - LG$$

In this regard, SG = water solubilizing group ($-SO_3Na$), C = chromogen (e.g., azo, anthraquinone), B = bridging or linking group (e.g., –NH–), RG = reactive group (e.g., chlorotriazine, vinylsulfone), and LG = leaving group (e.g., –Cl, –F, $–SO_4H$). These parts are illustrated for the structure in Fig. 13.12. This structure also shows that reactive dye structures can be quite small, much smaller in fact than those characterizing direct dyes. As a consequence, reactive dyes have significantly lower inherent affinity for cotton and can require high levels of salt (200–300 g/L) in their dyebaths to promote exhaustion [32].

In addition to giving high washfastness on cotton, reactive dyes usually give bright shades. The latter property arises from the fact that reactive dyes are often acid dye structures linked to reactive groups, as shown in Fig. 13.13. Reactive dyes have moderate-to-good lightfastness and fair-to-poor chlorine fastness.

Although the most commonly used reactive systems involve the halotriazine and sulfatoethyl sulfone (vinyl sulfone) groups, halogenated pyrimidines, phthalazines, and quinoxalines are also available (Fig. 13.14). For all of these systems, alkali is used to facilitate dye–fiber fixation, and fixation occurs either by nucleophilic substitution or addition (Figs. 13.15 and 13.16).

The requirement for alkali in the application of reactive dyes to cotton leads to an undesirable side reaction, namely hydrolysis of the reactive groups before dye–fiber fixation can occur (Fig. 13.17). Because the hydrolyzed dye cannot react with the fiber, this leads to wasted dye and the need to treat the residual color in the wastewater prior to dyehouse discharges. To improve percentage fixation, dyes with two or more reactive groups were developed (Fig. 13.18). This makes it possible for dye–fiber fixation to occur even when one reactive group undergoes hydrolysis [33].

Sulfur Dyes

Sulfur dyes are water-insoluble dyes that are applied to cotton [34]. They are used primarily for their economy and high washfastness, are easy to apply, and give mainly dull shades. Yellow, red, brown, olive, and blue colors can be produced; however, sulfur dyes are most important for their ability to delivery washfast black shades on cotton. In this regard, C.I. Sulfur Black 1 is the main dye used commercially. Sulfur dyes have acceptable lightfastness but poor bleachfastness.

Fig. 13.14 Examples of reactive groups found in fiber-reactive dye chemistry, where
1 = dichlorotriazine,
2 = monochlorotriazine,
3 = trichloropyrimidine,
4 = sulfatoethyl sulfone,
5 = dichlorophthalazine,
6 = monofluorotriazine,
7 = dichloroquinoxaline,
8,9 = difluorochloropyrimidine

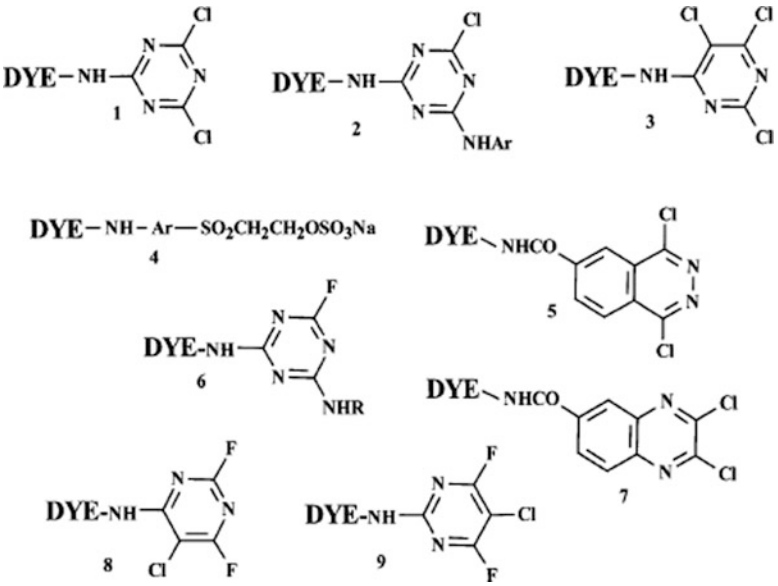


Fig. 13.15 Reactive dye fixation to cellulose via nucleophilic substitution

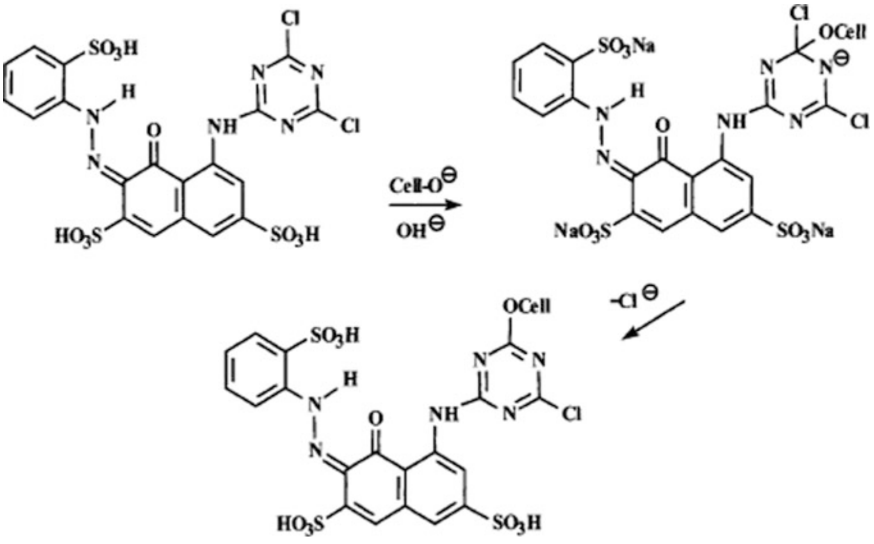
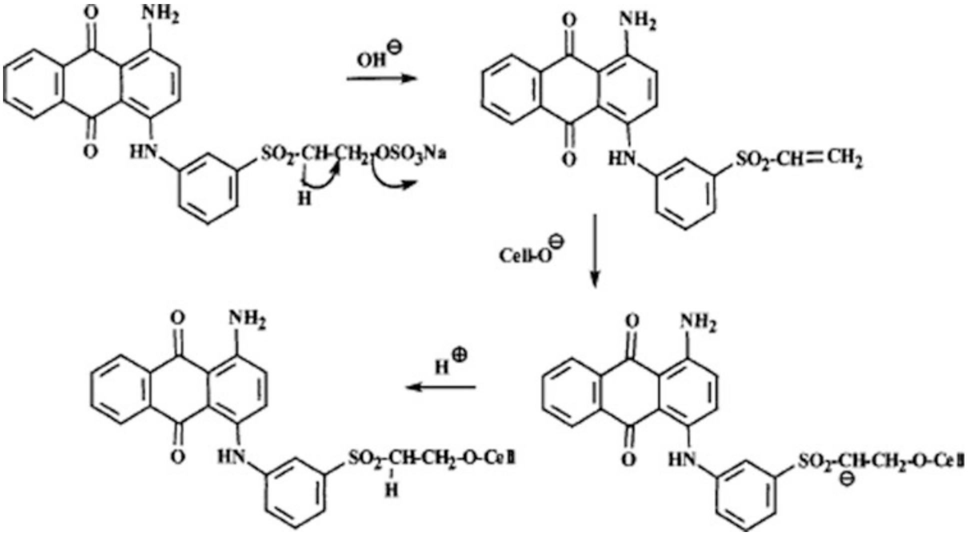


Fig. 13.16 Reactive dye fixation to cellulose via nucleophilic addition



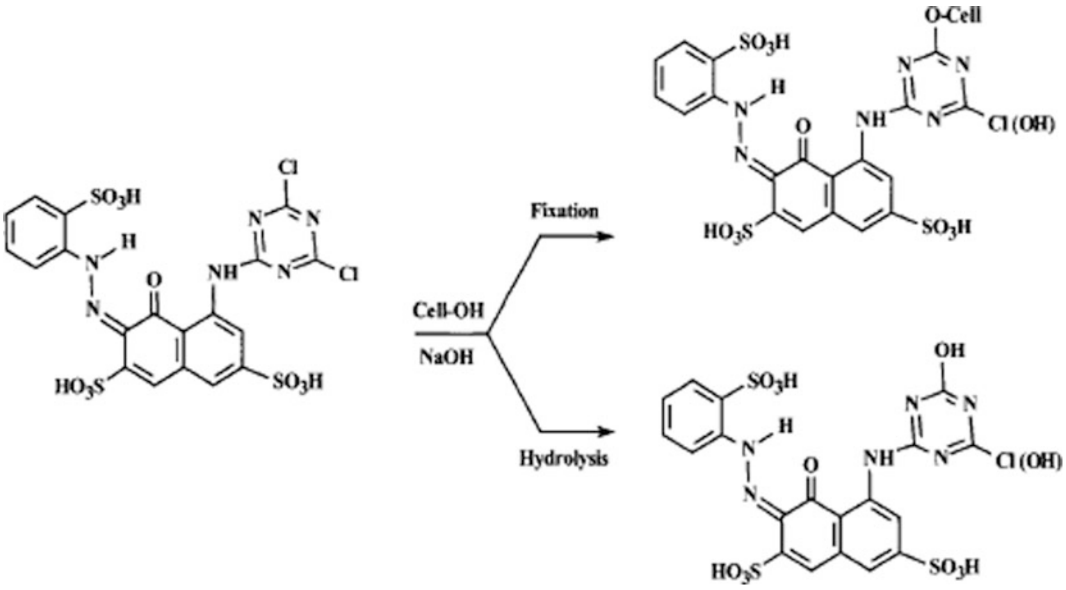


Fig. 13.17 Competing reactions when reactive dyes are applied to cotton

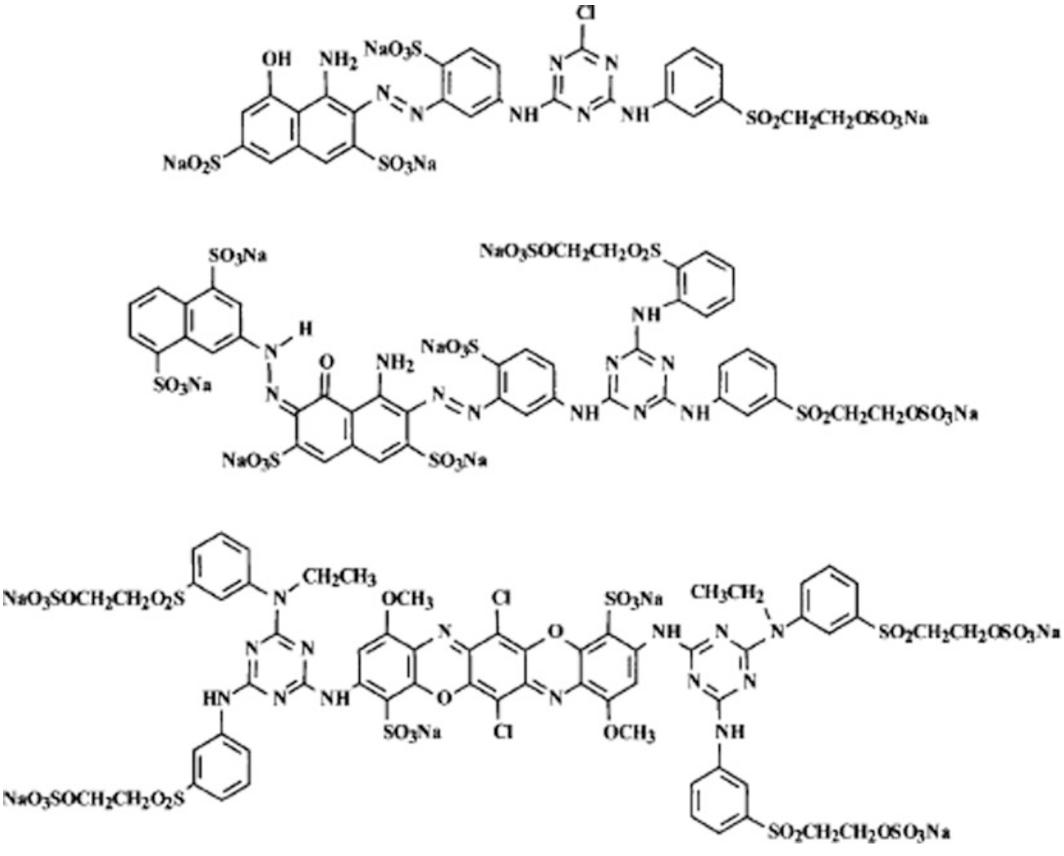


Fig. 13.18 Examples of bireactive and polyreactive dyes for cotton

Due to extremely low solubility, the precise structures of most sulfur dyes remain unknown. Much of what we know about sulfur dye structures arises from the characterization of certain degradation products or reaction precursors [35]. Based on such work, it has been possible to determine that structures of the type shown in Fig. 13.19 are covered in this dye class. A key common feature of sulfur dyes is the presence of sulfide ($-S_n-$) bonds, and it is this feature that makes dye application from an aqueous medium possible.

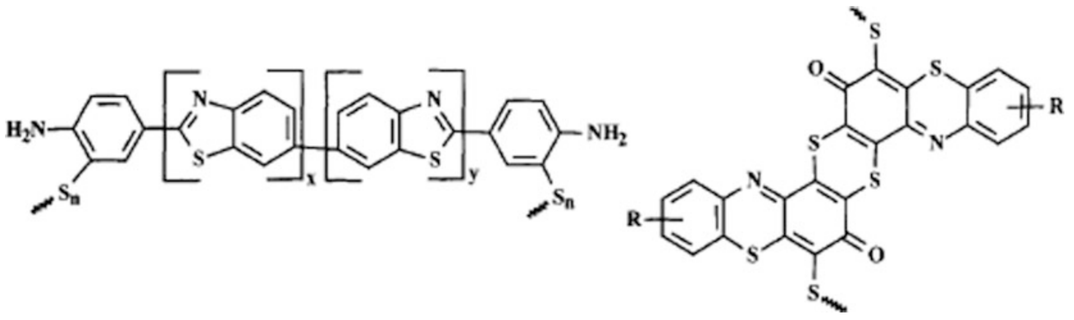


Fig. 13.19 Examples of sulfur yellow (*left*) and red (*right*) dye structures

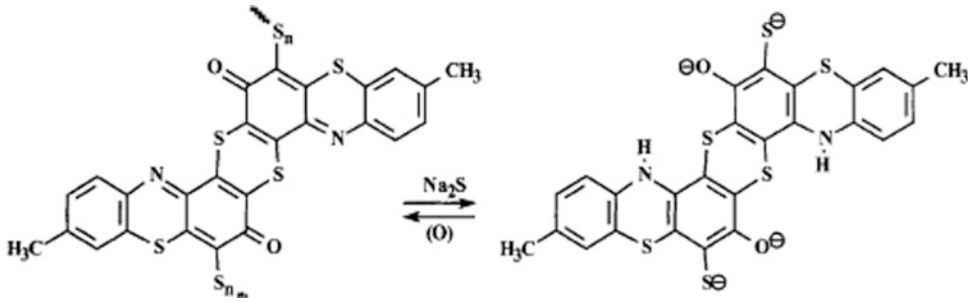


Fig. 13.20 Steps involved in the application of sulfur dyes to cotton

The reaction of sulfur dyes with sodium sulfide (Na_2S) at pH >10 effects the reduction of the sulfide bonds, giving their water soluble (leuco) forms. The reduced forms behave like direct dyes, in that they exhaust onto cotton in the presence of salt. Once applied, the reduced dyes are reoxidized to their water-insoluble forms, giving dyeing with good washfastness. This chemistry is illustrated in Fig. 13.20. Although the oxygen in air can be used for the oxidation step, an agent such as hydrogen peroxide is used because it works faster. Sulfur dyes have also been marketed in their prereduced form (Dye-S^-Na^+), as ready-to-use C.I. Leuco Sulfur dyes. Dye exhaustion in the presence of salt is followed by oxidation. Similarly, water-soluble sulfur dyes containing thiosulfate groups are sold as C.I. Solubilized Sulfur dyes. They are known as “Bunte salts” [36], have better leveling properties than the C.I. Sulfur dyes, and are attractive for package dyeing. The C.I. Solubilized sulfur dyes are applied with Na_2S and the chemistry associated with their two-step application is summarized in Fig. 13.21, along with a representative dye structure.

$$\textbf{Step 1:}\quad \text{Dye-S-SO}_3^{\ominus} + \text{S}_2^{\ominus} \longrightarrow \text{Dye-S}^{\ominus}$$

$$\textbf{Step 2:}\quad \text{Dye-S}^{\ominus} + \text{Dye-S-SO}_3^{\ominus} \longrightarrow \text{Dye-S-S-Dye}$$

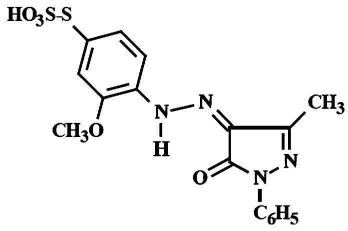


Fig. 13.21 Two-step chemistry employed in the application of C.I. Solubilized Sulfur dyes (*top*) and a representative dye structure (*bottom*)

Vat Dyes

Like sulfur dyes, vat dyes are water-insoluble colorants for cotton that must be reduced to their soluble “leuco” forms to be applied from an aqueous dyebath [37, 38]. Their name originates from their early application from wooden vessels known as vats. The term “vatting” is used to refer to the application of these dyes via chemical reduction followed by oxidation. Vat dyes are easier to reoxidize than sulfur dyes and the oxygen in air is often the agent used. As would be anticipated, most vat dyes display high washfastness. As a class, they have the best lightfastness and bleach fastness among the dyes families suitable for cotton. Some cause catalytic fading or phototendering on cotton [39].

Vat dyes have mainly anthraquinone (82%) or indigoid/thioindigoid (9%) structures, with the former having much better fastness properties. The anthraquinone vat dyes exhibit a bathochromic color shift (λ_{max} of higher wavelength) upon reduction to their leuco forms, whereas the

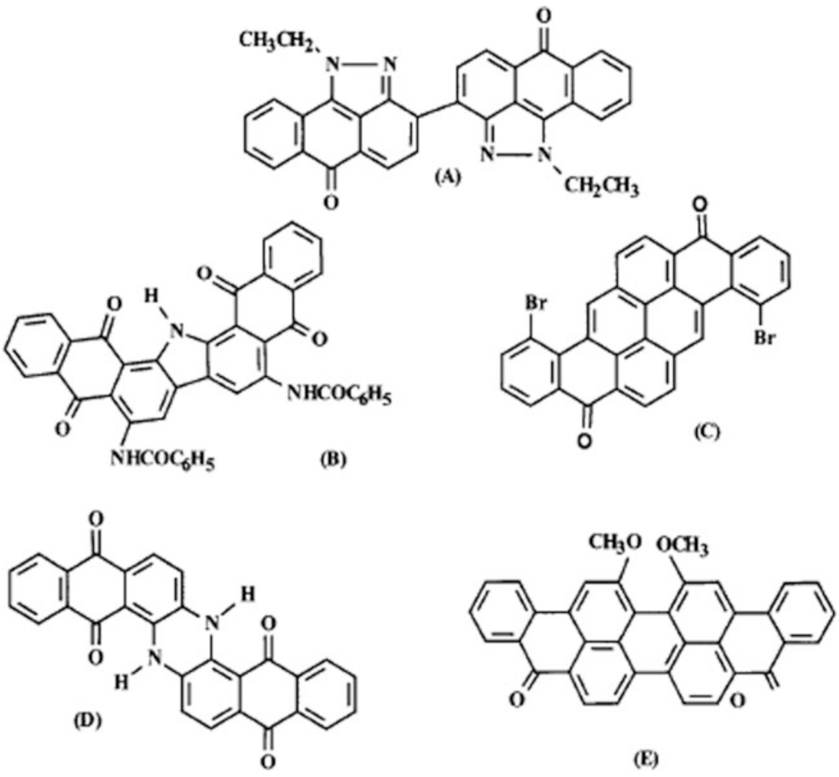


Fig. 13.22 Representative anthraquinone vat dye structures: CI Vat Red 13 (**a**), Vat Black 27 (**b**), Vat Orange 2 (**c**), Vat Blue 4 (**d**), and Vat Green 1 (**e**)

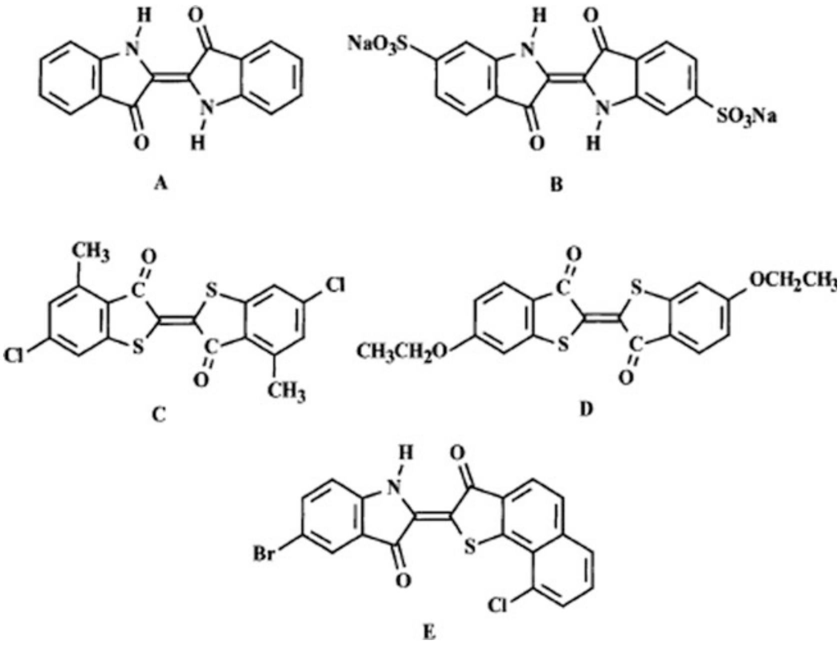


Fig. 13.23 Representative indigoid and thioindigoid vat dye structures: CI Vat Blue 1 (**a**), Food Blue 1 (**b**), Vat Red 1 (**c**), Vat Orange 5 (**d**), and Vat Black 1 (**e**)

indigoids exhibit a hypsochromic shift. Examples of the two structural types are shown in Figs. 13.13 and 13.22. Anthraquinone vat dyes having a single anthraquinone unit exist; however, those with the best fastness properties seem to have the equivalent of two anthraquinone units.

No doubt the best-known and biggest volume vat dye is C.I. Vat Blue 1, indigo, the denim blue dye. Closely related structures are the thioindigoids (4%), which have a sulfur atom in lieu of the –NH– group (Fig. 13.23). The thioindigoids are used mainly as colorants for printing and give orange and red hues. A few dyes having the features of both indigoid types are also known (e.g., Ciba Violet A).

The chemistry associated with the vatting process is illustrated in Fig. 13.24. For the reduction step, a mixture

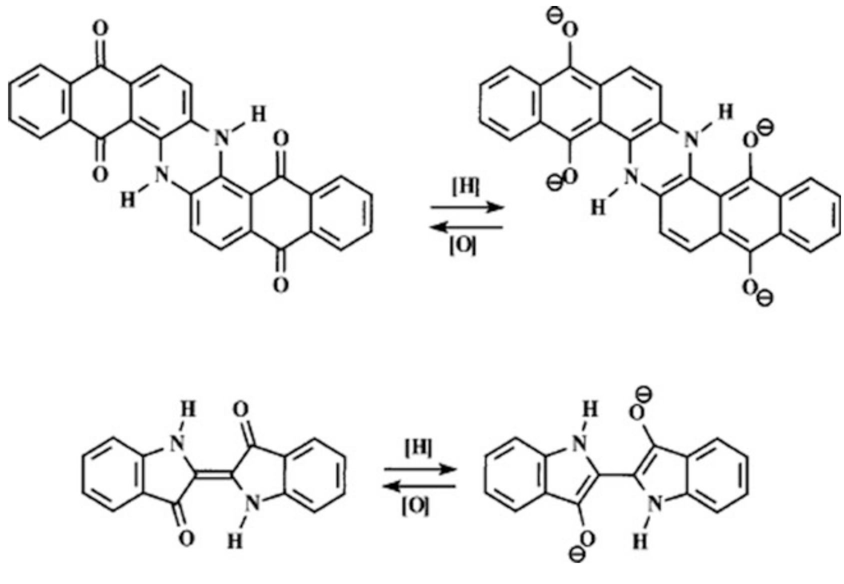
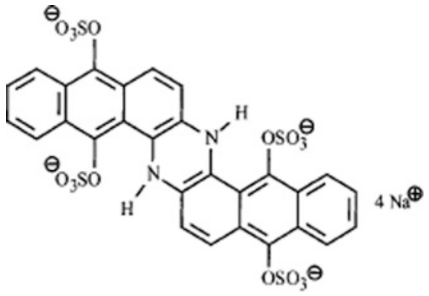


Fig. 13.24 Chemistry involved in the vatting of the CI Vat dyes Blue 4 (*top*) and Blue 1 (*bottom*)

of sodium hydroxide (caustic) and sodium hydrosulfite (hydro, $Na_2S_2O_4$) is used. Depending upon the amount of caustic and hydro employed one or both of the anthraquinone rings may undergo reduction.

Vat dyes are also available in prereduced forms (3%), an example of which is the leuco sulfuric acid ester C.I. Solubilized Vat Blue 4. These water-soluble forms have affinity for cellulose and exhaust like direct dyes. They are oxidized to the insoluble form using hydrogen peroxide.



C.I. Solubilized Vat Blue 4

Vat dyes are brighter than direct and sulfur dyes but less so than reactive dyes. They are the colorants of choice when dye bleachfastness on cotton is important. They span the entire color spectrum and can be applied to cotton using a variety of methods. With regard to the latter point, they can be further classified based on the temperatures involved in their application. Accordingly, there are hot (50–60 °C), warm (40–50 °C), and cold (25–30 °C) dyeing vat dyes. The hot dyeing types are large planar leuco forms having high affinity and no salt is required for their application. The cold types are small molecules with low affinity and require repeated application to get good build-up. Indigo falls into the cold dyeing category.

The Application of Dyes

The process of dyeing may be carried out in batches or on a continuous basis. The fiber may be dyed as stock, yarn, or fabric. However, no matter how the dyeing is done, the process is always fundamentally the same: dye must be transferred from a bath—usually aqueous—to the fiber itself. The basic operations of dyeing include: (1) preparation of the fiber, (2) preparation of the dyebath, (3) application of the dye, and (4) finishing. There are many variations of these operations, depending on the kind of dye. The dyeing process is complicated by the fact that single dyes seldom are used. The matching of a specified shade may require from two to a dozen dyes.

Fiber Preparation

Fiber preparation ordinarily involves scouring to remove foreign materials and ensure even access to dye liquor. Some natural fibers are contaminated with fatty materials and dirt, and synthetic fibers may have been treated with spinning lubricants or sizing that must be removed. Some fibers also may require bleaching before they are ready for use.

Dyebath Preparation

Preparation of the dyebath may involve simply dissolving the dye in water, or it may be necessary to carry out more involved operations such as reducing the vat dyes. Wetting

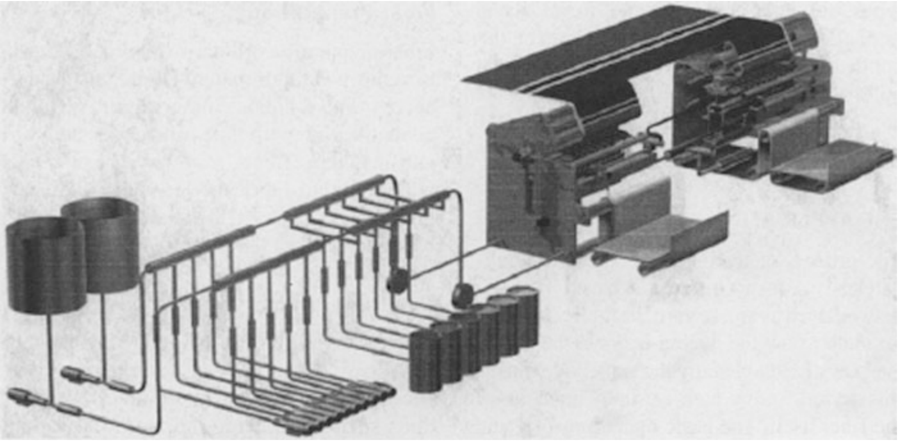


Fig. 13.25 Fluidyer carpet range dye applicator (Courtesy of Zima Corporation)

agents, salts, "carriers," retarders, and other dyeing assistants also may be added. Carriers are swelling agents that improve the dyeing rate of very hydrophobic fibers such as the polyesters. Examples are *o*-phenylphenol and biphenyl. Retarders are colorless substances that compete with dyes for dye sites or form a complex with the dye in the bath and act to slow the dyeing rate. Their use is necessary when too-rapid dyeing tends to cause unevenness in the dyeings.

Finishing

The finishing steps for many dyes, such as the direct dyes, are very simple: the dyed material merely is rinsed and dried. Vat-dyed materials, on the other hand, must be rinsed to remove the reducing agent, oxidized, rinsed again, and soaped before the final rinsing and drying steps are carried out. Generally, the finishing steps must fix the color (if fixation has not occurred during application) and remove any loose dye from the surface of the colored substrate. Residual dyeing assistants such as carriers also must be removed.

The types of textile structures that lend themselves to continuous dyeing methods are woven and tufted carpets. Continuous dyeing is designed for long runs of similar product; it is a high-output method of dye application.

The first volume-yardage continuous process was the continuous pad-steam process for vat dyes on cotton. The vat dye dispersion was padded onto the cloth and dried; this was followed by passage through a reducing bath, steaming for 30 s, passage through an oxidizing bath and, finally, washing. When it was discovered that disperse dyes could be thermosoled into polyesters by treatment with dry heat for 60 s and 400 °F, this procedure was readily adapted to continuous processing. The advent of large volumes of dyed polyester-cotton-blend fabrics in the late 1960s made it possible to combine these two processes into one thermosol pad-steam system.

Tufted nylon carpet grew to be the number-one floor covering in the United States in recent decades. Continuous open-width ranges were developed but not without a great deal of ingenuity to deliver the precise loading of liquid to the tufted surface. This was accomplished by a dye applicator that flooded the dye solution onto the carpet surface. The advancing technology in continuous, metered dyeing systems has created a need for dyes in liquid form, both dispersions and solution. The dyes used in carpet dyeing, for the most part, are supplied by the dye manufacturers as liquids. See Fig. 13.25.

Dyeing Methods/Batch

Virtually all types of textile structures are dyed by batch (or exhaust) methods of dyeing, such as stock, yarn, circular knits, warp knits, woven fabrics, garments, carpets, and so on. Batch methods include beck dyeing, jig dyeing, pad-batch, beam dyeing, and others. These methods are dictated primarily by the physical structure of the textile product and the type of fiber(s) it contains. Each of these batch methods employs a different type of machine. As an example, a circular knit fabric comprises cotton could be dyed in a beck, whereas the same structure comprises polyester most likely would be dyed in a high-pressure jet machine, and a garment constructed from the circular knit cotton likely would be dyed in a garment machine.

Stock dyeing often is carried out in large heated kettles made of stainless steel or other corrosion-resistant metal. These kettles can be sealed and used for dyeing at temperatures somewhat above the boiling point of water at atmospheric pressure.

Yarns are dyed in package machines. In this arrangement the yarn is wound onto perforated dye tubes and placed on spindles that are fit into a closed kettle. The dye solution is heated and pumped through the spindle and yarn package. A cycle of inside-outside flow usually is used to provide level

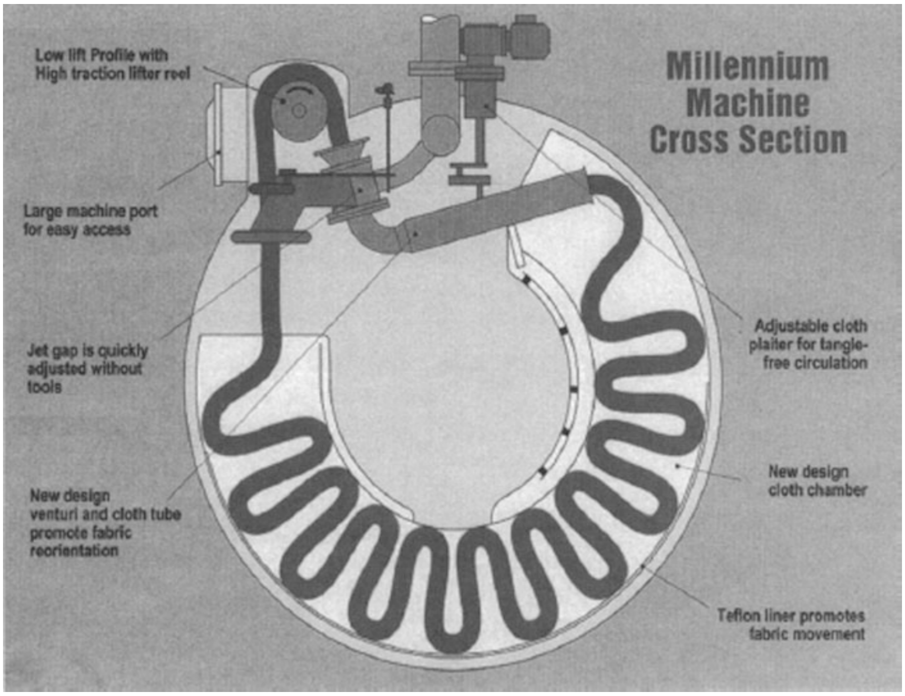


Fig. 13.26 Millennium jet dyeing machine (Courtesy of Gaston County Dyeing Machine Co.)

dyeing by equal exposure of the dye to yarns. Although the basis of package dyeing has not changed, a number of refinements have been introduced in recent years. Precision winding of the yarn has improved quality by giving a more uniform package density. Horizontal machines and valving between chambers to allow reconfiguration of the dye machine to control the size of the dyeing have changed the way package dyehouses are built. Robotization has been widely utilized to load and unload machines. Also lower-ratio dyebaths with higher flow rates have improved the energy efficiency of the newer machines.

Fabrics are dyed in machines that move them through the dye liquor either under tension (jig) or relaxed (beck). Fabrics also can be dyed in full width by winding them on a perforated beam through which hot dye liquor is pumped. This is the principle of the beam dyeing machine.

The pressure-jet dyeing machine is unique in that it has no moving parts. The cloth, in rope form, is introduced into a unidirectional liquid stream enclosed in a pipe. Liquor is pumped through a specially designed xanthen jet imparting a driving force that moves the fabric. The two fabric ends are sewn together to form a continuous loop.

The first jet machine was introduced in 1965. There are two major types of jet dyeing machines: the vertical kier and the elongated horizontal kier (see Fig. 13.26). In general, the kier uses small water volumes, whereas the elongated types use larger-volume ratios in dyeing. The kier types normally are used for more substantial fabrics, and the elongated types are suited for fine or delicate fabric styles. Important features in today's machines are improved corrosion-resistant alloys and the ability to operate at higher efficiencies with minimum energy consumption. The control systems have been refined; there is simultaneous loading and unloading. Larger-capacity machines also are being built; a jet dye machine has been developed for carpet dyeing.

Printing

Printing is a special kind of localized dyeing that produces patterns. Four kinds of printing have long been recognized: (1) direct, (2) dyed, (3) discharge, and (4) resist. In direct printing, a thickened paste of the dye is printed on the fabric to produce a pattern. The fabric then is steamed to fix the dye and is finished by washing and drying. Dyed printing requires that the pattern be printed on the fabric with a mordant. The entire piece then is placed in a dyebath containing a mordant dye, but only the mordanted areas are dyeable. Washing then clears the dye from the unmordanted areas, leaving the pattern in color.

In discharge printing, the cloth is dyed all over and then printed with a substance that can destroy the dye by oxidation or reduction, leaving the pattern in white. When a reducing agent such as sodium hydrosulfite is used to destroy the dye, the paste may contain a reduced vat dye. Finishing the goods by oxidation and soaping then produces the pattern in color. In resist printing, certain colorless substances are printed on the fabric. The whole piece then is dyed, but the dye is repelled from the printed areas, thus producing a colored ground with the pattern in white.

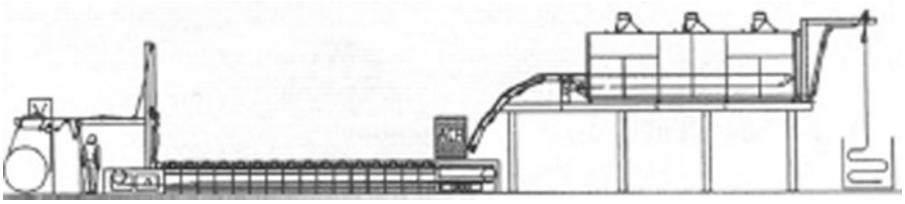


Fig. 13.27 Rotascreen V rotary screen print machine (Courtesy of Zimmer Machinery Co.)

Printing is most often done with rotary screens etched in the design to be printed. Printing paste is fed constantly to the center of the rotating screen from a nearby supply, and a squeegee pushes the colored paste through the holes in the screen, leaving the dye paste only in the intended areas, a separate screen is required for each color in the pattern. See Fig. 13.27.

An important recent advance in the pattern-coloring of textiles is ink-jet or digital printing. Milliken's Millitron and Zimmer's ChromoJet have been successfully used for carpet and upholstery markets for over two decades. Finer-resolution machines began to emerge in the late 1980s when Stork introduced a prototype machine. During the past 5 years a number of manufacturers have introduced digital ink-jet printers that use either CYMK (Cyan, Yellow, Magenta, and Black) to make a composite color or true-color machines that use mixed pigment systems. The current machines are very successful at furnishing one of a kind and for use in rapid prototyping.

Pigment Dyeing and Printing

Pigment dyeing and printing are processes that compete with the more conventional means of dyeing and printing described above. These processes use water-insoluble dyes or pigments that are bound to the surfaces of fabrics with resins. A paste or an emulsion, containing pigment and resin or a resin-former, is applied to the fabric. The goods then are dried and cured by heat to produce the finished dyeing or print. During the heating or curing, fabric, resin, and pigment become firmly bonded together. This method of color application is economical and produces good results. It should be noted that the pigment is confined to the surface of the fabric and can be selected without regard for fiber affinity.

Nontextile Uses of Dyes

Colorants for nontextile use have been developed mainly for use in hair dyeing, photography, biomedical application, and electronics and reprographics [34–41]. Dye application areas involving the latter areas include ink-jet printing, thermal or pressure dye transfer, laser printing, liquid crystal displays, optical data storage, and nonlinear optics. In several nontextile applications, dyes are not used for their ability to deliver color. Instead, they are used because of their potential electrical properties, such as photoconduction and electrostatic charging of toners, and in some cases they are used because they absorb IR radiation, which induces heating effects. The latter property is important in optical data recording.

Liquid Crystal Dyes

Dyes for liquid crystalline media typically have (1) nonionic structures, (2) high purity, (3) solubility and compatibility with the medium, (4) a transition dipole that is parallel with the alignment axis of the molecular structure, and (5) good alignment with the liquid crystal molecule [42]. Examples include the disazo and anthraquinone dyes in Fig. 13.28.

Ink-Jet Dyes

Dyes used in this area must have the following properties: (1) very good water solubility; (2) low toxicity; (3) good stability to UV light, heat, and moisture; (4) quick fixation to paper following application (deposition); and (5) good color strength [43]. To achieve high resistance to removal by water (wetfastness), ink-jet dyes often contain fewer sulfonate groups and one or more carboxylate groups. This change in structural features allows the dyes to have good solubility in alkaline ink formulations but high wetfastness following deposition. This change in solubility behavior is known as *differential solubility* [44]. Structures in Fig. 13.29 illustrate the type changes made to the early ink-jet dye C.I. Food Black 2 to enhance wetfastness.

New water-soluble yellow dyes for ink-jet printing are similar to the initially used dye C.I. Direct Yellow 86, except that they are smaller (Fig. 13.30). The size change is designed to provide the solubility needed for high throughput ink cartridge systems, without clogging the ink-jet nozzles.

Most of the new water-soluble magenta dyes are based on H-acid. Examples include dyes that contain a fluorocyanophenyl group (Fig. 13.31). To improve the lightfastness of magenta dyes, gamma acid can be used as the coupling component. For very bright magenta prints, dyes based on xanthene structures can be used. Examples include C.I. Acid Red 52,

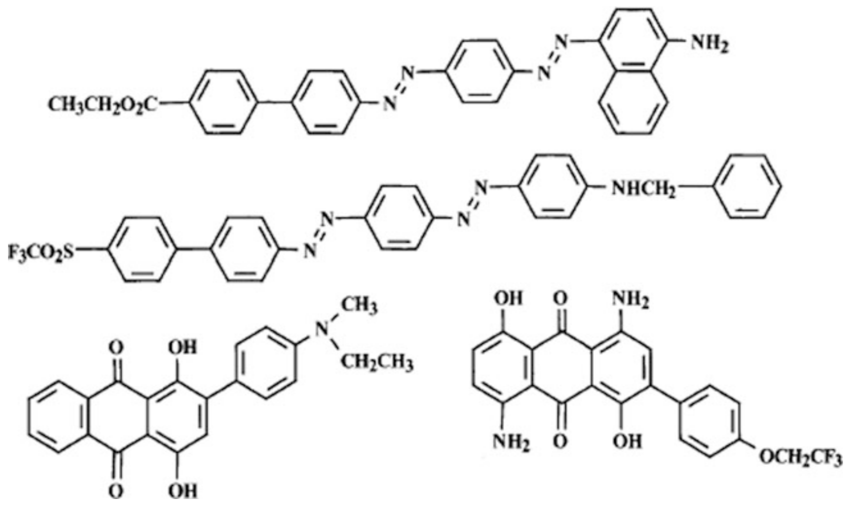


Fig. 13.28 Azo and anthraquinone dyes suitable for liquid crystalline media

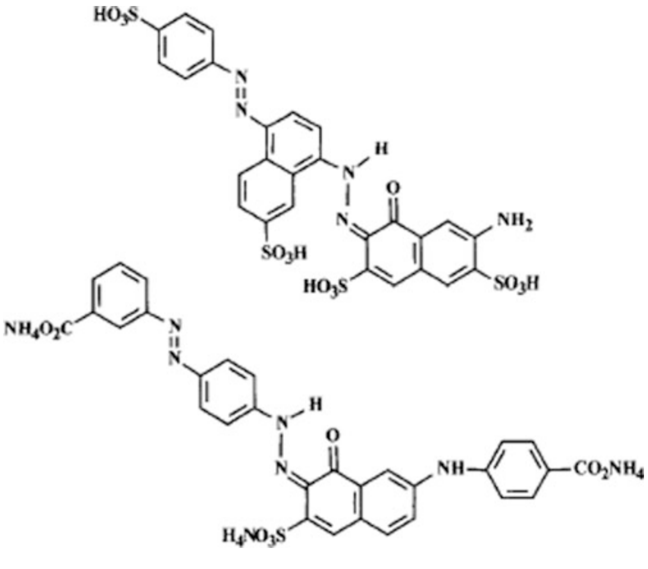
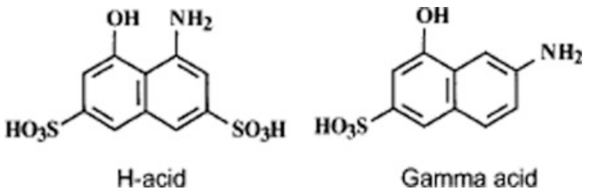


Fig. 13.29 Food Black 2 (*top*) and its wetfast analog (*bottom*)



which has low photostability, and a carboxylated analog, which has better photostability and wetfastness (Fig. 13.32).

The water-soluble cyan dyes continue to be based on the copper phthalocyanine (CuPc) system. In this regard, C.I. Direct Blue 199 has proved effective, due to its good color strength and photostability. Carboxylated analogs of this type of dye have been developed to enhance wetfastness on paper (Fig. 13.33).

Thermal and Pressure-Sensitive Printing

In direct thermal printing, a color former (colorless) and a developer (acidic) are brought into contact in the presence of heat, to produce color on paper [45, 46]. The most important color in thermal printing is black and the majority of the color formers are fluorans (Fig. 13.34). The most important application of direct thermal printing is in facsimile machines.

In pressure-sensitive printing technology the color former is dissolved in a solvent and encapsulated [47]. The use of pressure (pen, typewriter key) ruptures microcapsules containing the color former, which generates color upon contacting a developer. Black prints are usually obtained either from fluorans or from color former mixtures. Compounds of the type shown in Fig. 13.35 can be used in two- and three-component mixtures.

Organic Photoconductors and Toners

Photoconductors and toners are used in photocopiers and laser printers to produce images [48]. Organic photoconductors consist of a charge-generating layer and a charge-transporting layer. The former comprises pigments and the latter comprise electron-rich organic compounds that are usually colorless. Suitable organic pigments for charge generation include azo pigments, tetracarboxydiimides, polycyclic quinones, phthalocyanines, perylenes, and squarylium compounds (e.g., Fig. 13.36).

Colorants are used in toners to provide color and control the electrostatic charge on toner particles. Diarylides and monoarylides have been used as the yellow pigments in colored toners. The magenta pigments are often quinacridones and the cyan pigments are CuPcs.

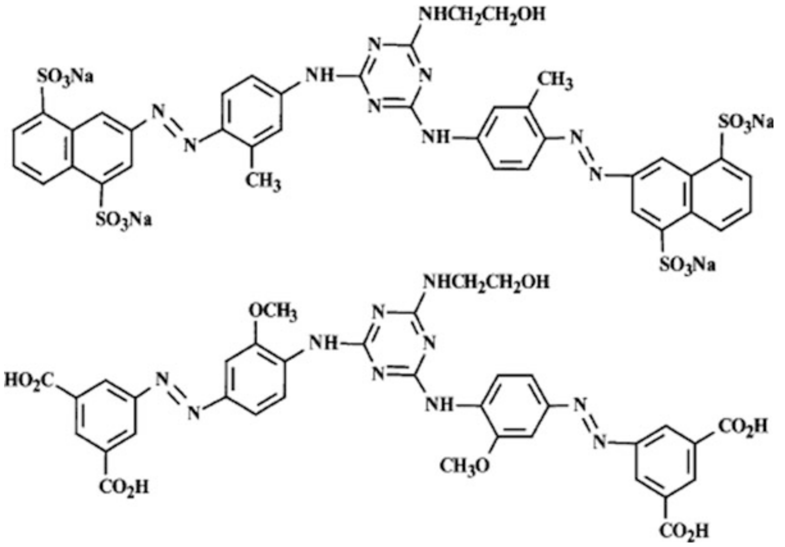


Fig. 13.30 Examples of yellow ink-jet dyes (upper dye is Direct Yellow 86)

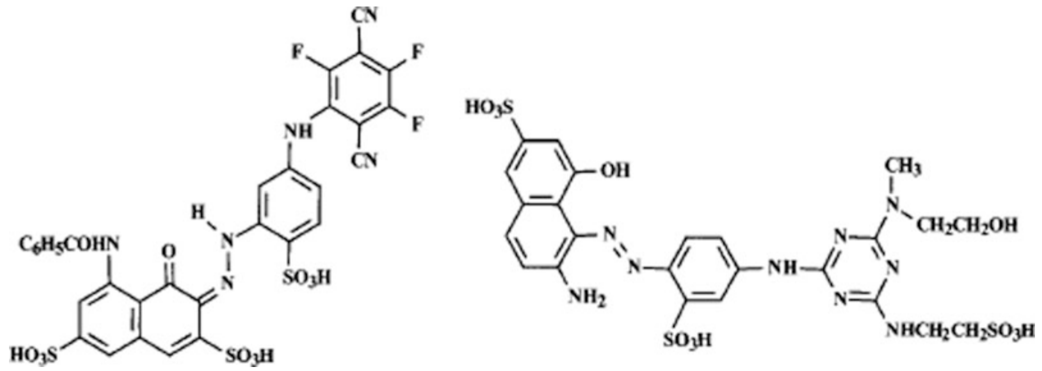


Fig. 13.31 Examples of magenta ink-jet dyes based on H-acid (*left*) and gamma acid (*right*)

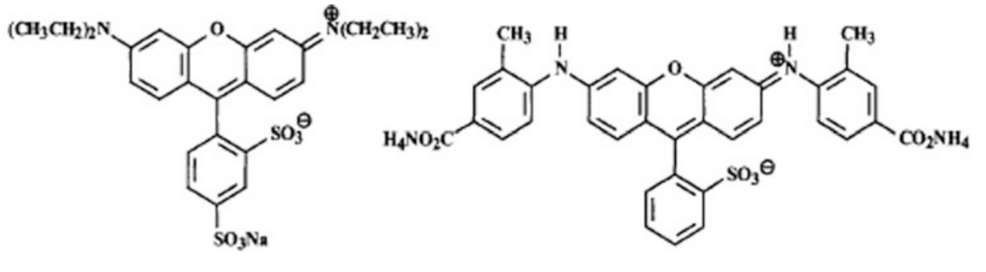


Fig. 13.32 Acid Red 52 (*left*) and a more lightfast analog (*right*)

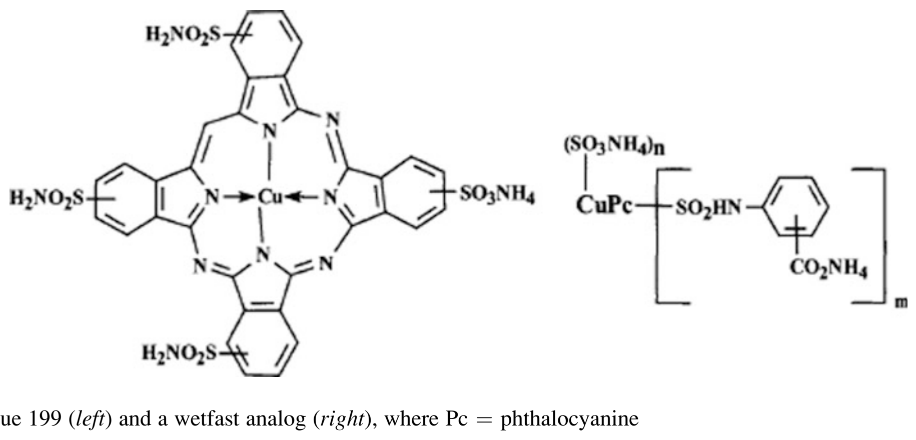


Fig. 13.33 Direct Blue 199 (*left*) and a wetfast analog (*right*), where Pc = phthalocyanine

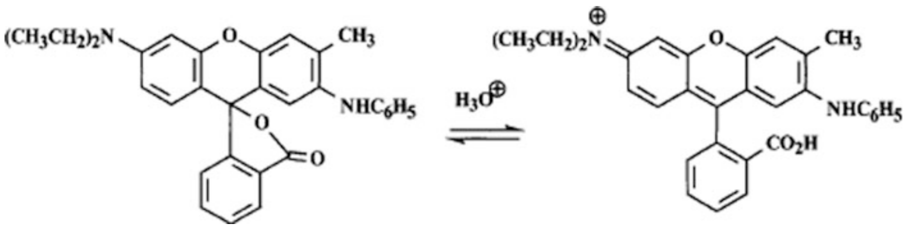


Fig. 13.34 Acid-induced conversion of a color former to its colored form

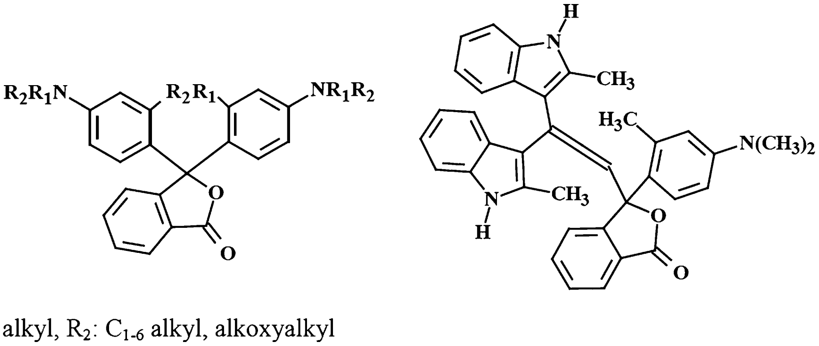


Fig. 13.35 Color former combinations suitable for pressure-sensitive printing

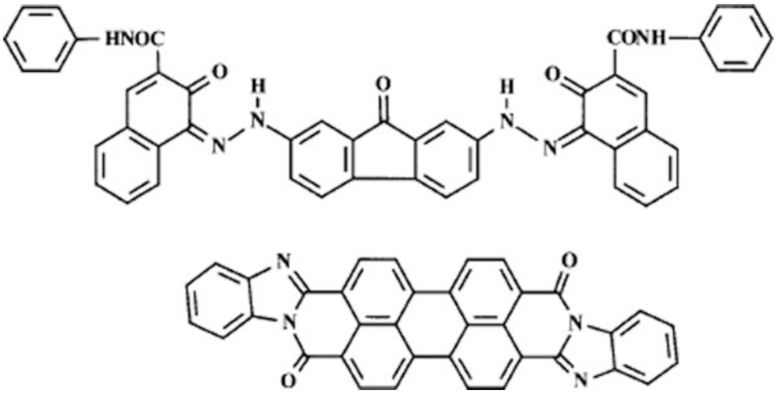


Fig. 13.36 Azo (*top*) and perylene (*bottom*) pigments for charge generation

Infrared Absorbing Dyes

Infrared dyes include indoleninecyanines and azulenium compounds, both of which are used in optical recording materials [49]. Other examples are metal (Mn, Fe, Co, Cd, Al, Cu, Pd)-complexed phthalocyanines, quinones, quinonoids, and imminium and diiminium compounds (Fig. 13.37).

Laser Dyes

Lasers in which dyes comprise the active medium have become one of the most widely used types [50]. The key virtue of these systems is their ability to cover virtually the entire fluorescence spectral region. Accordingly, the most commonly used dyes are highly fluorescent and include coumarin, rhodamine, oxazine, and *syn*-bimane structures (Fig. 13.38). Dye lasers are employed in liquid form, which allows them to dispel excessive heat by recirculating the dye solution. Good photostability and efficient laser action under flashlamp excitation are important properties.

Biomedical Dyes

Dyes can be used clinically in bioanalysis and medical diagnostics and in the treatment of certain diseases (cf. Fig. 13.39). For instance, near infrared (NIR) and IR-absorbing dyes can be used in automated DNA sequencing (dye A), fluorescent dyes can be used in cancer detection (dye B), and certain azo and heterocyclic dyes can be used in virus (dye C), cell (dye D), and bacteria (dye E) detection. An in-depth summary of dyes in this area has been published recently [51].

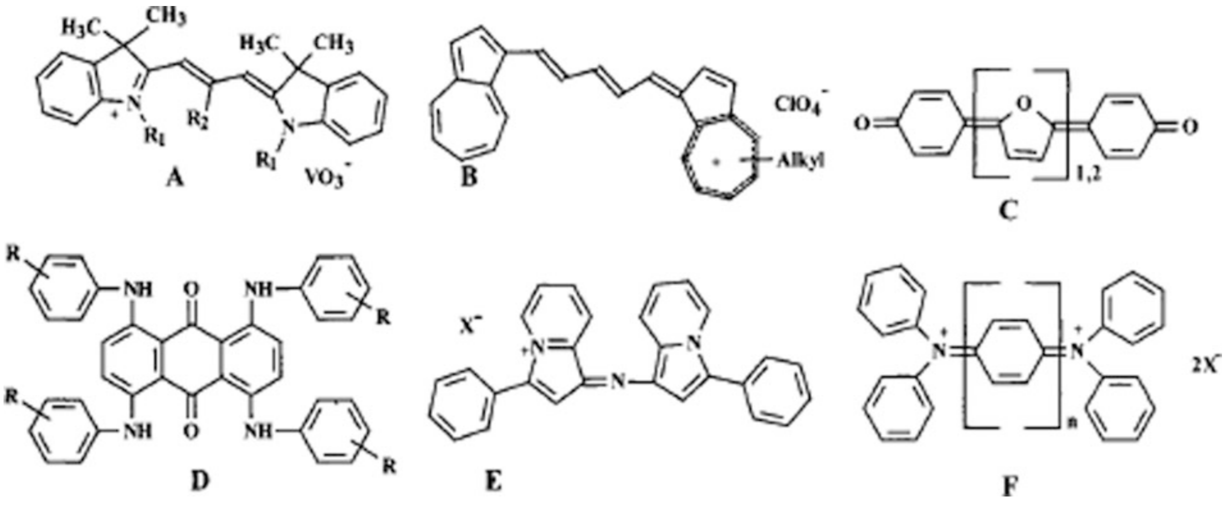


Fig. 13.37 Structures of indoleninecyanine (**a**), azulenium (**b**), quinonoid (**c**), anthraquinone (**d**), imminium (**e**), and diiminium (**f**) IR-absorbing dyes, where R = alkyl or alkoxy

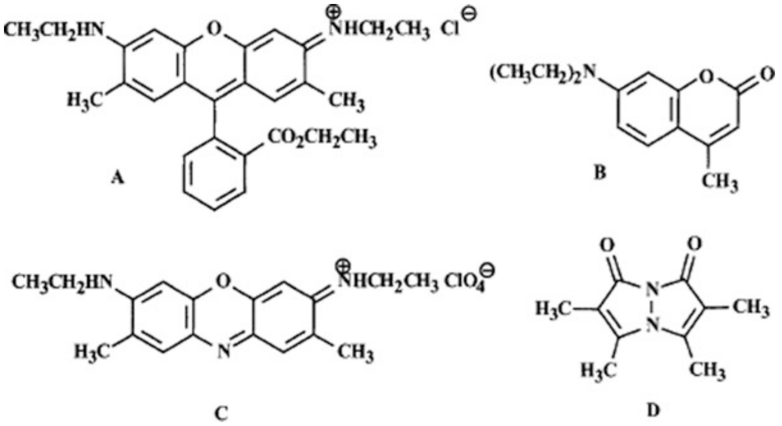


Fig. 13.38 Rhodamine (**a**), coumarin (**b**), oxazine (**c**), and *syn*-bimane (**d**) laser dye structures

Hair Dyes

About 80% of the dyes used in hair coloring are known as oxidation hair dyes [52, 53]. The remaining 20% of the available hair dyes are mainly synthetic dyes that have affinity for protein substrates. Oxidation dyes are produced directly on hair by oxidizing aromatic diamines (e.g., *para*-phenylenediamine or 2,5-diaminotoluene) with a suitable oxidizing agent. In this regard, the diamines have been referred to as "primary intermediates" and the oxidizing agents (e.g., hydrogen peroxide) as "developers." Other suitable primary intermediates are aminodiphenylamines, amino-methylphenols, and *para*-aminophenol.

When used alone, the primary intermediates give a quite limited shade range following oxidation on hair. To enhance the range of available hair colors, the primary intermediates are oxidized in the presence of suitable "couplers." Whereas most couplers do not produce colors when exposed to developers alone, they give a wide array of hair shades in combination with primary intermediates. Suitable couplers include 3-aminophenol, resorcinol, and α-naphthol.

The chemistry associated with the oxidation of primary intermediates is now reasonably well known. For *para*-phenylenediamine and *para*-aminophenol, this involves the process outlined in Fig. 13.40. It can be seen that dye formation is a two-step process involving oxidation and self-coupling.

C.I. Basic dyes such as Yellow 57, Red 76, Blue 99, Brown 16, and Brown 17 have been used in color refreshener shampoos and conditioners. Similarly, C.I. Acid dyes such

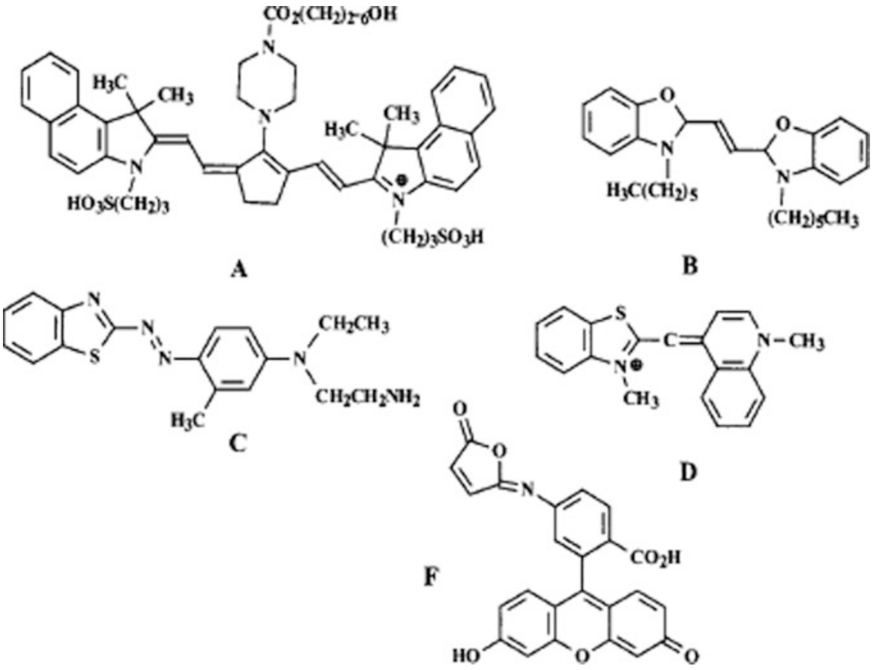


Fig. 13.39 Examples of experimental dyes used in biomedical applications

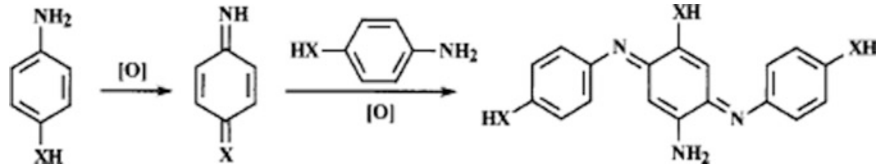


Fig. 13.40 Oxidation of primary intermediates, where X = O, NH

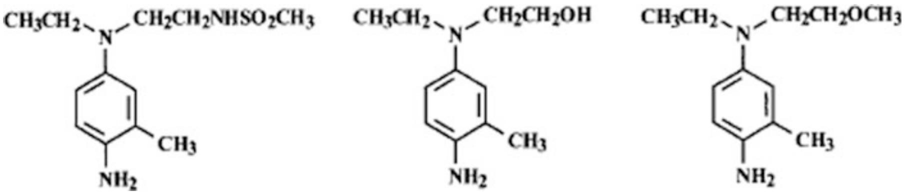


Fig. 13.41 Structures of some developers used in color photography

as Yellow 3, Orange 7, Red 33, Violet 43, and Blue 9 have been used in shampoos, in this case to deliver highlighting effects [54].

Photographic Dyes

Color photography is still one of the most important and interesting nontextile uses for synthetic dyes. The chemistry employed is comparable to that described above for oxidation hair dyes, in that an oxidizable substrate (e.g., phenylenediamine) is combined with a coupler to produce the target colorant. In this case the diamine is referred to as the "developer," and it is oxidized by silver halide in the photographic film. The oxidized developer then reacts with the coupler to form the dye. This process produces a negative dye image consisting of yellow, magenta, and cyan dyes in proportion to the amount of red, blue, and green light absorbed by the film [55].

Some widely used developers are shown in Fig. 13.41. They can be used to produce the yellow, magenta, and cyan dyes shown in Fig. 13.42. These dye structures demonstrate that acetoacetanilide, pyrazolone, and indoaniline intermediates are useful for producing yellow, magenta, and cyan colors, respectively.

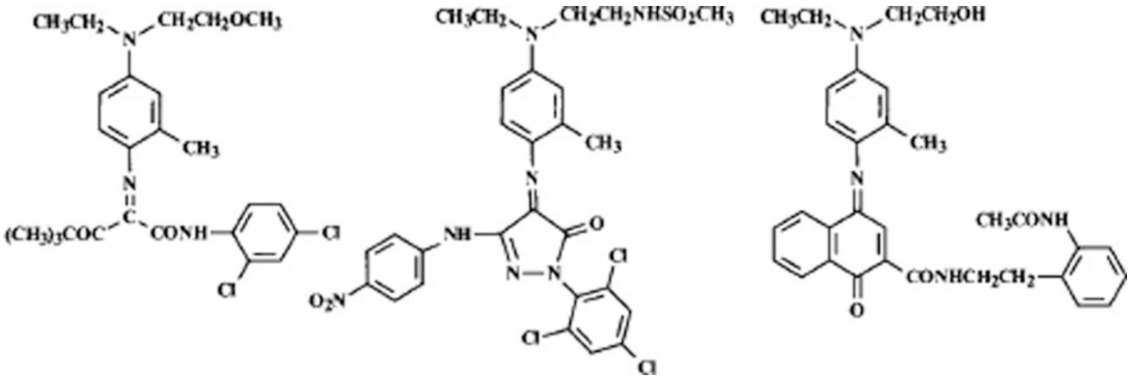


Fig. 13.42 Structures of yellow (*left*), magenta (*center*), and cyan (*right*) photographic dyes

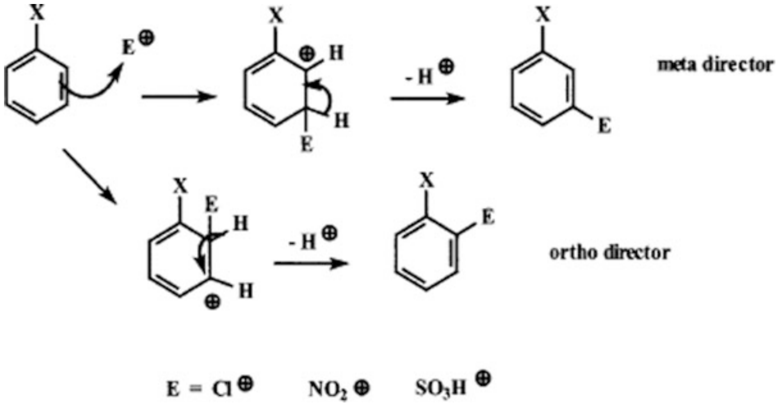


Fig. 13.43 Electrophilic attack of an aromatic ring containing deactivating and activating groups

Dye Intermediates

The raw materials used to synthesize organic dyes are commonly referred to as dye intermediates. Largely, they are derivatives of aromatic compounds obtained from coal tar mixtures. The majority of these derivatives are benzene, naphthalene, and anthracene-based compounds. This section provides an overview of the chemical reactions used to prepare the key intermediates employed in dye synthesis. In this regard, emphasis is placed on halogenated, aminated, hydroxylated, sulfonated, and alkylated derivatives of benzene, naphthalene, and anthraquinone.

Most dye intermediates are prepared by reactions involving electrophilic or nucleophilic substitution processes. The electrophilic processes include nitration, sulfonation, and halogenation reactions, and the nucleophilic processes include hydroxylation and amination reactions. Electrophilic substitution reactions are of the form shown in Fig. 13.43. In this regard, the incoming electrophile (electron-seeking species) reacts with the more electron-rich positions. When the aromatic ring contains ring-activating groups (e.g., hydroxy, alkoxy, amino, alkyl), the incoming group will attack *ortho*/*para* positions. If ring-deactivating groups (e.g., nitro, sulfonic acid, carboxylic) are present, the positions *meta* to the deactivating groups will be attacked.

Other key dye intermediates are prepared by oxidation and reduction processes. Examples of each of these processes are covered in the sections that follow.

Nitration

For dye intermediates, this process involves the introduction of one or more nitro (NO_2) groups into aromatic ring systems. Nitro groups serve as chromophores (color bearers, precursors for amino groups, and as auxochromes (color aiders). Because they are *meta*-directing groups they are also useful in the strategic placement of another incoming group.

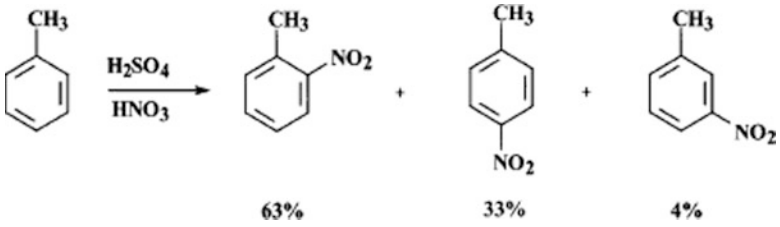


Fig. 13.44 Nitration of toluene using mixed acid

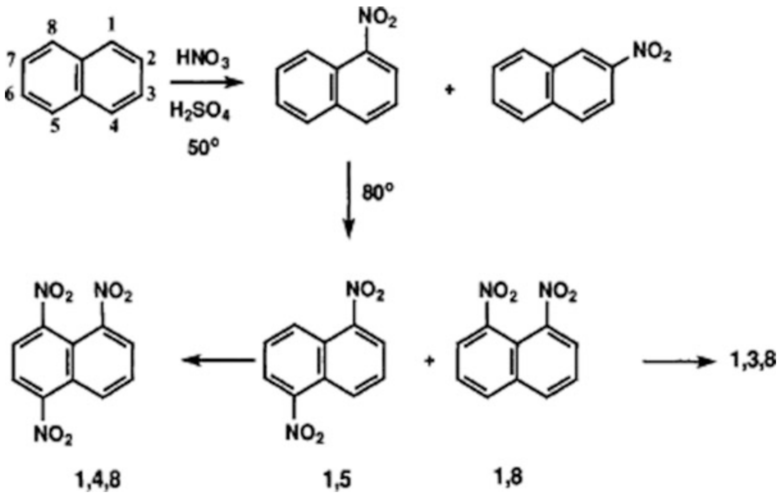


Fig. 13.45 Nitration of naphthalene using mixed acid

Nitric acid (HNO_3) is the chemical agent commonly used in nitration reactions. Depending upon the degree of ring activation, HNO_3 may be used in combination with other acids. In fact, nitrations are often conducted by using a mixture of HNO_3 and sulfuric acid (H_2SO_4). This combination is known as "nitrating mixture" or "mixed acid," and it is especially effective when deactivated ring systems are to be nitrated. Dilute HNO_3 or a HNO_3/acetic acid (CH_3CO_2H) mixture can be used for nitrating very reactive ring systems. When the former is used there is also the potential for ring oxidation to occur rather than the desired nitration, depending upon the actual compound undergoing nitration. Examples of nitration reactions are shown in Figs. 13.44, 13.45, and 13.46. The nitration of toluene (Fig. 13.44) is selected because it illustrates what can happen when monosubstituted benzenes having a ring-activating group are used. In this case, the principal products reflect a statistical mixture of *ortho* and *para* isomers, with only a small amount of the *meta* isomer obtained. Nitration is conducted near 20 °C and the products are separated by distillation.

Nitration of naphthalene gives mostly the 1-nitro isomer (~90%), initially. Introduction of a second nitro group takes place in the opposite ring because the existing nitro group reduces the reactivity of the ring to which it is attached. Although, nitro groups are *meta* directors, in this case they can also direct the incoming second (or third) nitro group to a *peri* position. In the naphthalene ring system, the *peri* positions are those that are 1,8 and 4,5 to each other (Fig. 13.45).

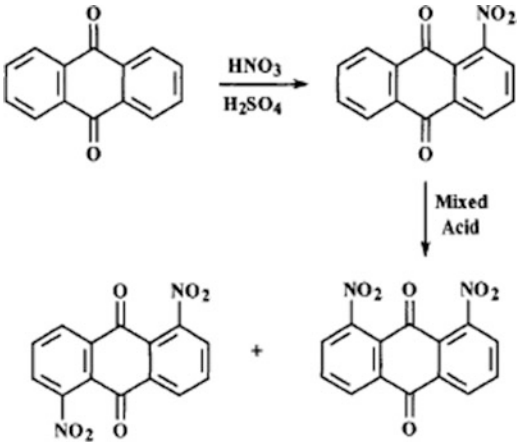
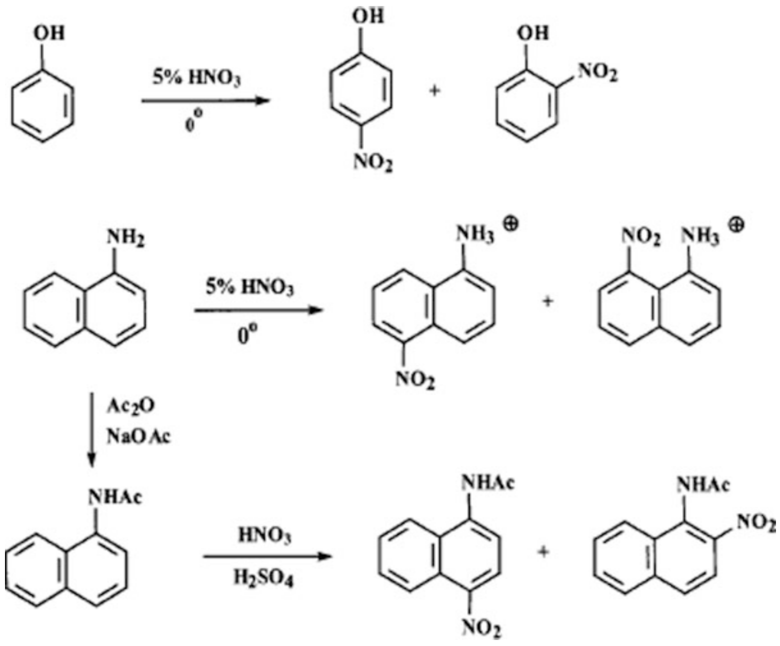


Fig. 13.46 Nitration of anthraquinone using mixed acid

The nitration of anthraquinone at 50 °C gives, initially, the 1-nitro isomer, and if nitration continues at 80–90 °C, the 1,5 and 1,8 isomers are obtained. Further nitration is impractical and serves to point out that the anthraquinone ring is appreciably less reactive than the naphthalene system. This will be more evident as the chemistry reported in this section continues to unfold.

Fig. 13.47 Nitration of phenol and 1-naphthylamine



The nitration of phenols and amines must be conducted with care, as these systems are subject to ring oxidation if the temperature gets too high. For instance, the nitration of phenol itself is conducted near 0 °C using 5% HNO_3. This gives a mixture of *ortho* and *para* isomers that can be separated by steam distillation (Fig. 13.47). Aromatic amines are often protected by *N*-acetylation prior to nitration. This reduces both the potential for ring oxidation and the amount of *meta* isomer that forms when the amino group undergoes protonation. The protonated amino group ($-NH_3^+$) is a *meta* director, unlike the free amino (NH_2) and the acetylated amino (NHAc) groups. This chemistry is illustrated in Fig. 13.47 for 1-naphthylamine. Following nitration, the acetyl group can be removed by hydrolysis.

Reduction

The most important reduction reactions are those leading to aromatic amines that are suitable for azo dye formation. Although this usually involves the reduction of a nitro group to an amino (NH_2) group, the reduction of azo groups to amino groups is also an important process. Agents that are commonly used to effect chemical reductions include: Fe + HCl or H_2SO_4; Na_2S; NaSH; Zn + NaOH; H_2 + transition metal catalysts; and $Na_2S_2O_4$. Examples of these reductions are given in Figs. 13.48, 13.49, 13.50, and 13.51. While the reduction of nitrobenzene can be conducted in a number of ways, a key commercial process involves the method in Fig. 13.48, where high-temperature hydrogenation is used.

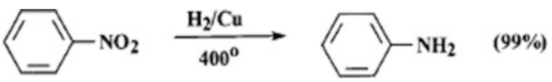


Fig. 13.48 Commercial process for the reduction of nitrobenzene

The reduction of azo compounds using sodium hydrosulfite ($Na_2S_2O_4$) and NaOH is an important reaction, as it provides an indirect method for the amination of phenols and naphthols (Fig. 13.49). The reduction of nitro groups in anthraquinone compounds works best when a mild reducing agent (e.g., sodium hydrosulfide, NaSH) is used. In this way one avoids reducing the quinoid system.

An example of an important reduction reaction involving Fe + H^+ is shown in Fig. 13.50. In this case the sequential use of nitration and reduction is illustrated.

It must also be pointed out that the medium employed in the reduction process can play a major role in the outcome of the reaction. A good example is the reduction of nitrobenzene in the presence of acid or alkali. One should expect the reduction to follow the course shown in Fig. 13.48 under normal conditions, however, in acidic media the product obtained is mainly *para*-aminophenol. In fact, this has long been the key step in the commercial route to acetaminophen [55], which is obtained by *N*-acetylation of the reduction product. When the reduction is conducted in the presence of alkali and Zn, the nitro compound is converted to a hydrazo compound via azoxy and azo intermediates. The hydrazo compound is important because it can be treated with acid to form diaminobiphenyls known as benzidines. These reactions are shown in Fig. 13.51. Because benzidine (4,4′-diaminobiphenyl) itself is known to be a human carcinogen, its use as a dye intermediate is substantially curtailed in the western world.

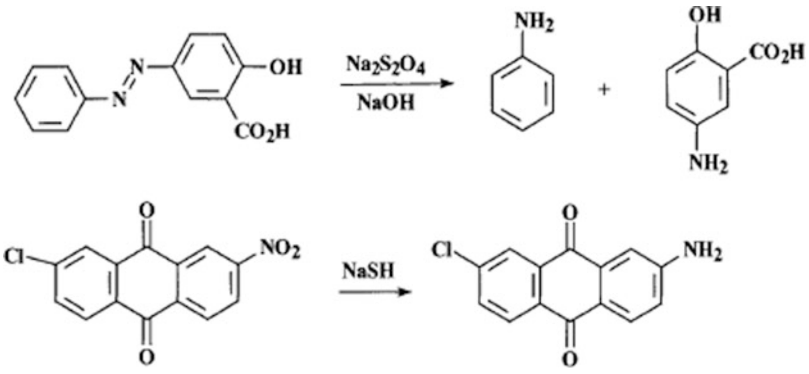


Fig. 13.49 Reduction of azo (*upper*) and nitroanthraquinone (*lower*) compounds

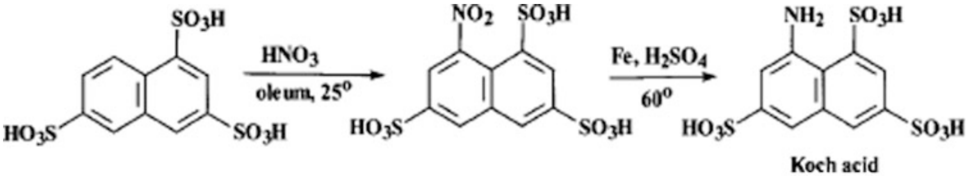


Fig. 13.50 Formation of Koch acid via nitration and reduction steps

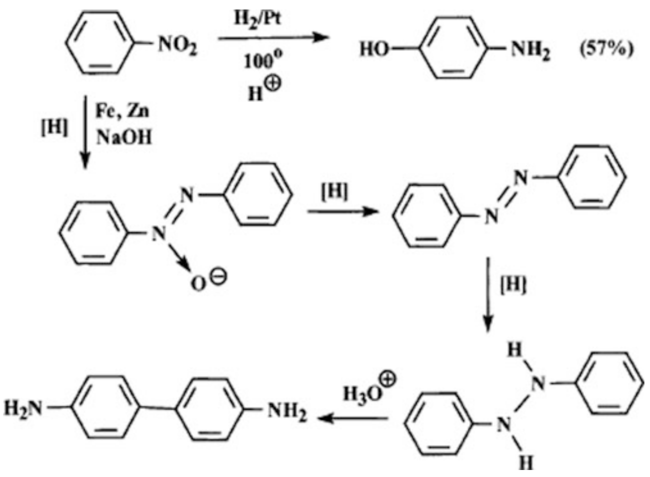


Fig. 13.51 Reduction of nitrobenzene in acidic and alkaline media

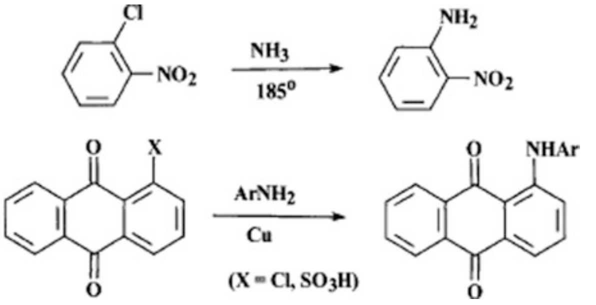


Fig. 13.52 Amination reactions involving benzene and anthraquinone compounds

Amination

In as much as the previous section covers the reduction of nitro and azo compounds as a method for introducing amino groups, the focus of this section is direct aminations involving replacement reactions and examples of indirect amination. In the former case, amination via the replacement of activated halogens using an alkyl or arylamine is widely used. The examples given in Fig. 13.52 show that halogens positioned *ortho* to a nitro group or in an α-position on the anthraquinone ring can replaced by amino groups. The former reaction also works well when the groups are *para*. However, the reaction is difficult and usually impractical when electron-donating rather than electron-attracting groups are situated *ortho* and/or *para* to the halogen. In the case of the anthraquinone system, α-sulfonic acid and α-nitro groups can also be replaced.

An important amination reaction involves hydroxy-substituted naphthalenes (Fig. 13.53). In a process known as the Bucherer reaction, naphthols are heated under pressure with a mixture of ammonia and sodium bisulfite. As the second and third examples indicate, the reaction works with aromatic amines and is selective. Note that the β-hydroxy group reacts preferentially when an α-hydroxy group is also present, and that two hydroxy groups in the same compound can be replaced.

An alternative route to the synthesis of aminoanthraquinones is the two-step sequence shown in Fig. 13.54. In this case, amination occurs via the condensation of *para*-toluenesulfonamide with chloroanthraquinone followed by hydrolysis of the sulfonamide bond. This method provides a way to introduce an –NH_2 group without the use of ammonia gas and the associated high temperatures and pressures.

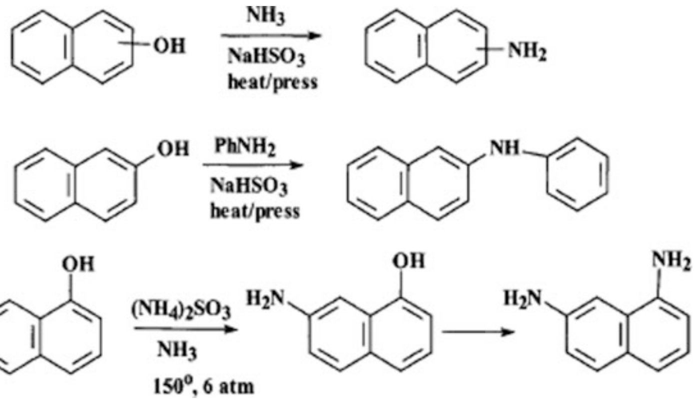


Fig. 13.53 Amination of naphthalene compounds via the Bucherer reaction

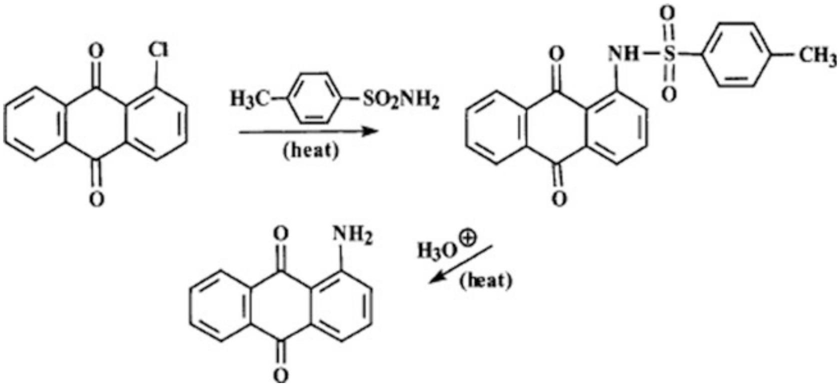


Fig. 13.54 An indirect amination of the anthraquinone ring system

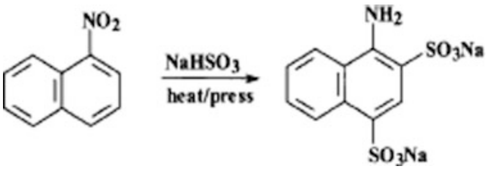


Fig. 13.55 A one-step amination and sulfonation of naphthalene

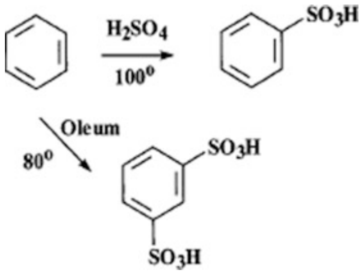


Fig. 13.56 The mono and disulfonation of benzene

Another interesting reaction is shown in Fig. 13.55. In this example, amination and sulfonation occur when α-nitronaphthalene is reduced by heating it under pressure with $NaHSO_3$.

Sulfonation

The introduction of one or more sulfonic acid groups (sulfonation) into dye intermediates is often conducted to confer water solubility, to provide fiber affinity, and to direct other incoming groups in the steps that follow sulfonation. In most cases this process employs sulfuric acid but in difficult cases, for example, deactivated ring systems, oleum (an SO_3/H_2SO_4 mixture) is used. This chemistry is illustrated in Fig. 13.56 for benzene. Here we see that benzene can be sulfonated using sulfuric acid and that the introduction of a second sulfonic acid group requires oleum. When a more reactive system is sulfonated, less stringent conditions are required. For example, naphthalene (Fig. 13.57) is readily sulfonated up to four times without using oleum. It is important to note that it is not possible to have sulfonic acid groups that are *ortho*, *para*, or *peri* to each other in the naphthalene system.

By contrast, the sulfonation of anthraquinone requires oleum and no more than two sulfonic acid groups can be introduced. In this system, sulfonation in the α-position requires the use of HgO as a catalyst. Examples of the possible products are shown in the scheme in Fig. 13.58.

Fig. 13.57 Examples of products obtained from the sulfonation of naphthalene

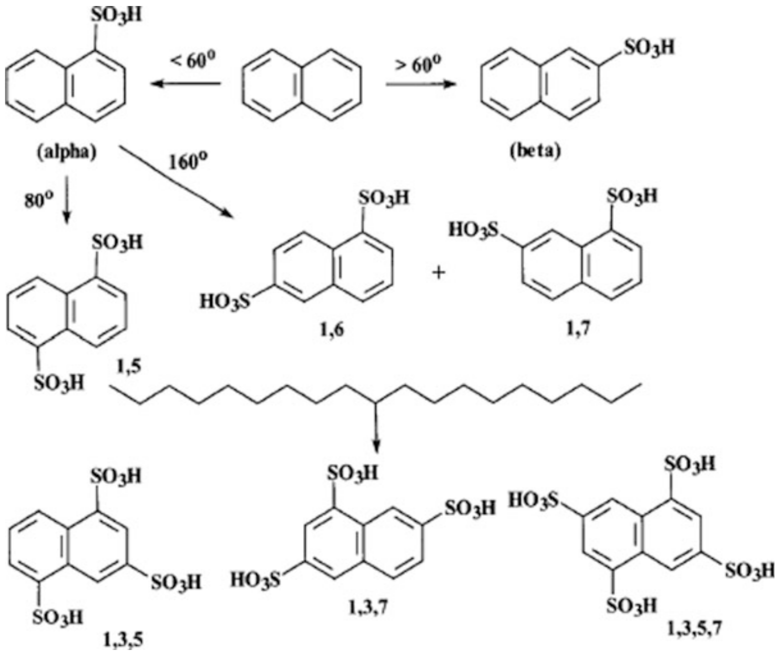
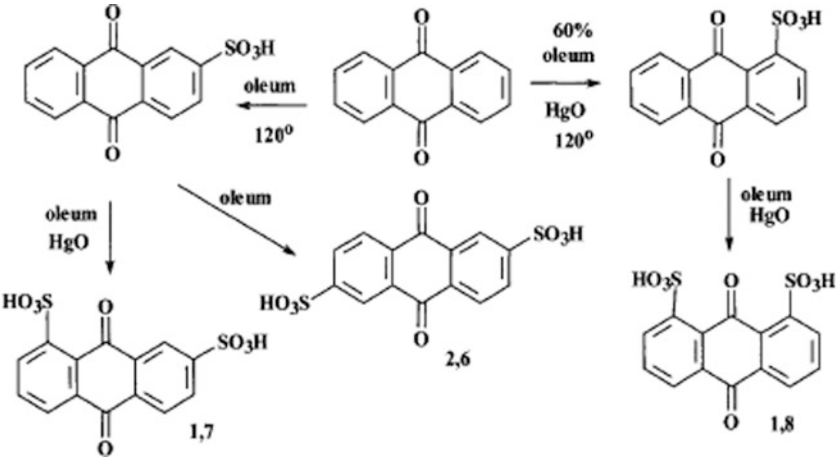


Fig. 13.58 Typical products produced from the sulfonation of anthraquinone



The sulfonation of β-naphthol produces several important dye intermediates, the nature of which depends upon the conditions employed (Fig. 13.59). At low temperatures, sulfonation occurs in the α-position to giveoxy-Tobias acid. Under ambient conditions Crocein acid is produced and at elevated temperatures three other products are obtained, including two that are disulfonated.

The sulfonation of aromatic amines such as aniline can give a mixture of products that must be separated prior to dye synthesis. When a single product is sought, the "baking" reaction is often employed (Fig. 13.60). In this process, the sulfate salt of aniline is prepared, dried, and then "baked" in an oven under vacuum. The product in this case is the important dye intermediate, sulfanilic acid. Similarly, naphthylamine sulfonic acid can be produced, and if the *para*-position is occupied, sulfonation of an *ortho*-position occurs (Fig. 13.61).

Halogenation

For dye intermediates, halogenation most often involves the incorporation of chloro groups. As pointed out earlier, halogens are important as leaving groups in the amination process, but they can also be used to enhance brightness and influence color. Later, we show that halogens are important as leaving groups in reactive dye chemistry, and in this regard chloro and fluoro groups are used.

Figures 13.62, 13.63, 13.64, and 13.65 provide examples of chlorination reactions. In the first example, the commonly

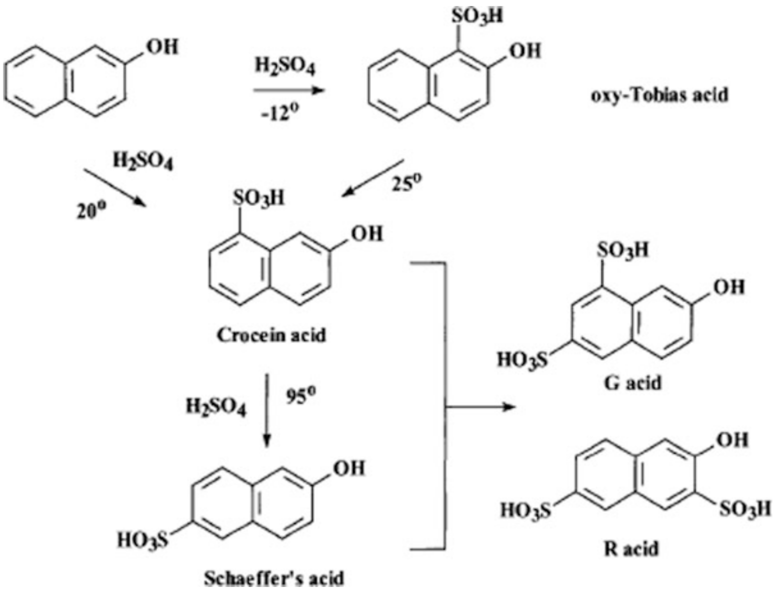


Fig. 13.59 Dye intermediates prepared from the sulfonation of β-naphthol

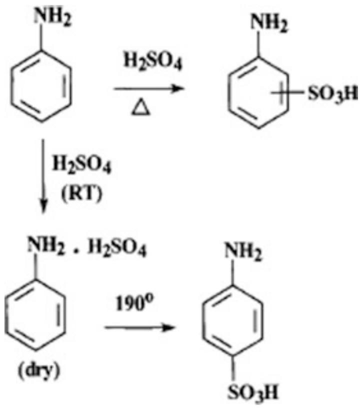


Fig. 13.60 Direct sulfonation of aniline (*upper*) vs. the baking reaction (*lower*)

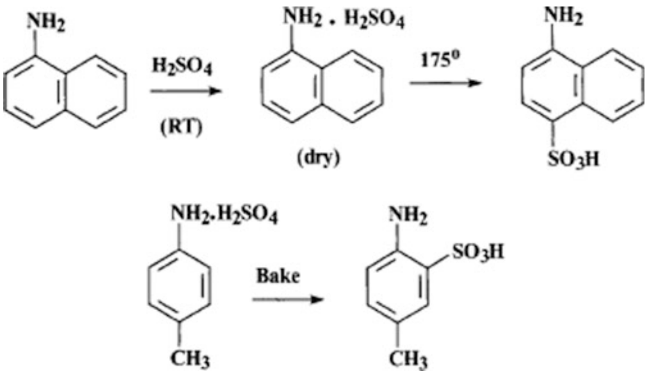


Fig. 13.61 Additional examples of the baking reaction

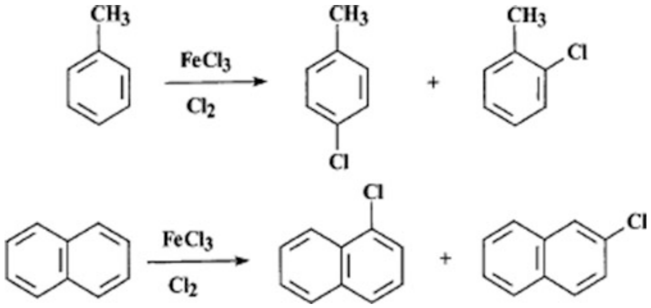


Fig. 13.62 Halogenation of benzene and naphthalene rings

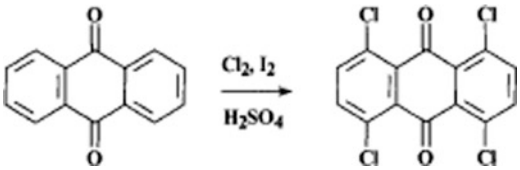


Fig. 13.63 Synthesis of the dye intermediate 1,4,5,8-tetrachloroanthraquinone

used agent $FeCl_3/Cl_2$ is employed for the chlorination of benzene and naphthalene rings. This method is not practical for the chlorination of anthraquinone. In this case the most important reaction is the tetrachlorination process shown in Fig. 13.63.

Because the chlorination of phenols and aromatic amines can be difficult to control, chlorination of these systems usually employs agents that will give a single chloro group when this is the desired outcome. In this regard, NaOCl and

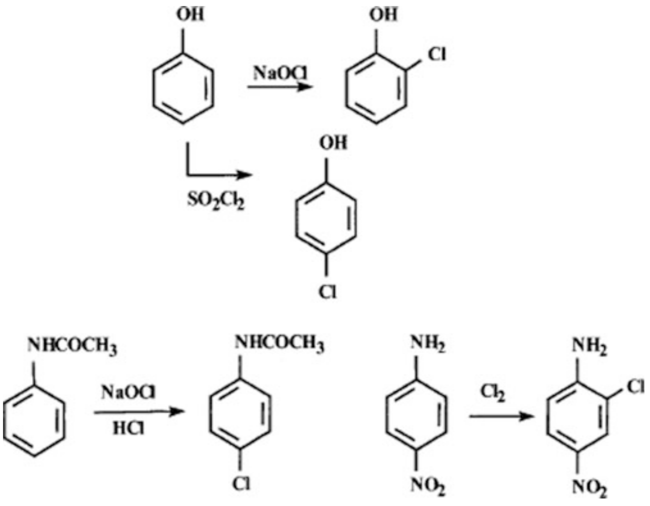


Fig. 13.64 Chlorination of phenol and anilines

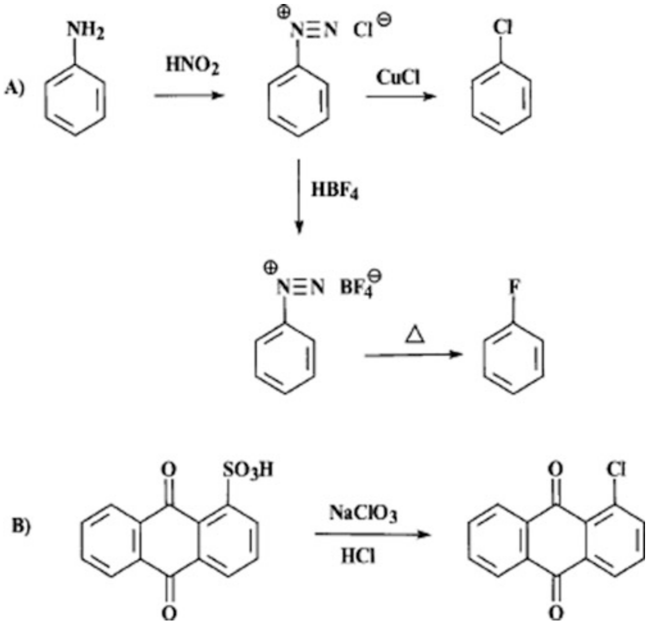


Fig. 13.65 Three important indirect halogenation reactions

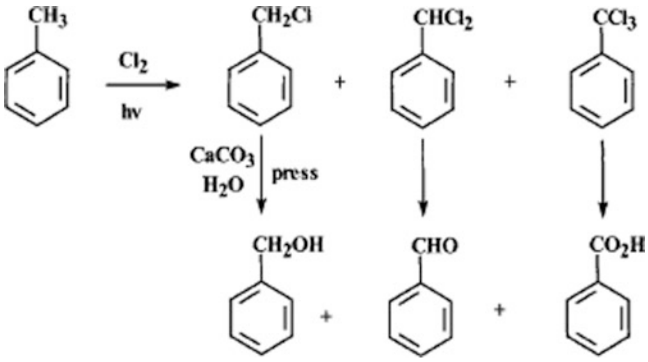


Fig. 13.66 Free radical chlorination of toluene and hydrolysis of the products

SO_2Cl_2 are quite useful chlorinating agents (see Fig. 13.64). In cases involving amines, often such compounds are protected by acetylation prior to chlorination. If the reactivity of the ring has been reduced by the presence of a deactivating group (e.g., $–NO_2$), acetylation may not be needed.

Halogens are also introduced via indirect methods, three examples of which are shown in Fig. 13.65. In the first case (sequence "A"), aniline is diazotized and the resultant diazonium compound is heated with cuprous chloride to give chlorobenzene, in a process known as the Sandmeyer reaction [56]. Alternatively, the diazonium compound can be converted to the tetrafluoroborate salt, which in turn is heated to give fluorobenzene. In sequence "B," anthraquinone-2-sulfonic acid is converted to the corresponding chloro compound by treatment with $NaClO_3$/HCl. All three reactions can be used to prepare a wide array of halogenated aromatics.

Chlorination is also an important step in the synthesis of oxygenated aromatic compounds. In this case, chlorination takes place at alkyl groups attached to the rings and is conducted in the absence of iron. The use of UV light speeds up this reaction, which is illustrated for toluene in Fig. 13.66. This free radical chlorination of toluene gives a mixture of benzyl chloride, benzal chloride, and benzotrichloride, which in turn can be hydrolyzed to benzyl alcohol, benzaldehyde, and benzoic acid.

Hydroxylation

The introduction of hydroxy groups is important in dye chemistry because it opens the door to azo dye formation, using phenols and naphthols, and provides an important auxochrome. Hydroxylation methods include alkali fusion, replacement of labile groups, and the reverse Bucherer reaction. In the alkali fusion reaction, naphthalene sulfonic acids are reacted with molten NaOH, KOH, or combinations of the two, as illustrated in Fig. 13.67. When disulfonated naphthalenes are used, the reaction can be stopped at the mono-hydroxylation stage if this is the desired outcome. The second example shows that the α-sulfonic acid group reacts faster.

When sulfonated anthraquinones are used, hydroxylation is conducted with $Ca(OH)_2$ to avoid over oxidation that occurs when hot NaOH is used. Example reactions are shown in Fig. 13.68.

The short sequence in Fig. 13.69 shows that aqueous alkali can also be used in hydroxylation reactions. In both cases, however, elevated temperatures are required.

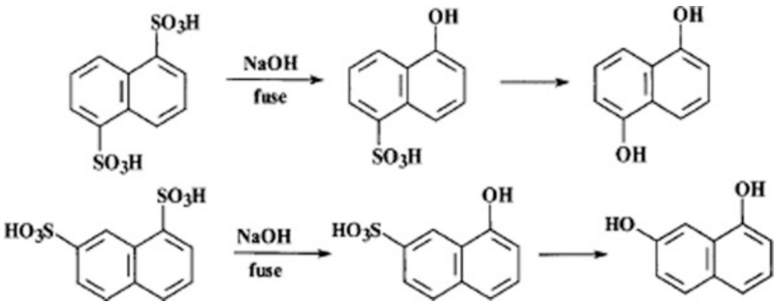


Fig. 13.67 Sodium hydroxide fusion as a hydroxylation process

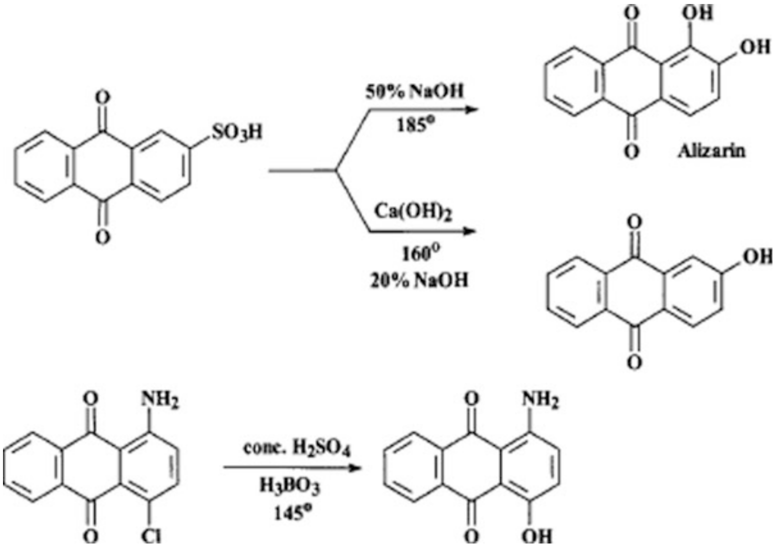


Fig. 13.68 Hydroxylation of the anthraquinone system

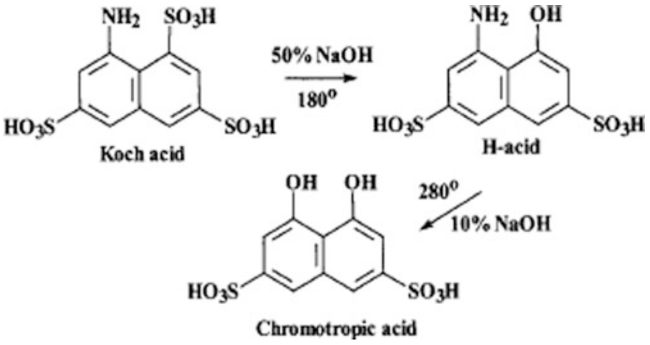


Fig. 13.69 Synthesis of two key dye intermediates using hydroxylation steps

Other important hydroxylation reactions are shown in Fig. 13.70. Here it can be seen that the Bucherer reaction is reversible, that the fusion reaction works for sulfonated benzene compounds, and that diazonium compounds undergo hydrolysis to produce phenols/naphthols.

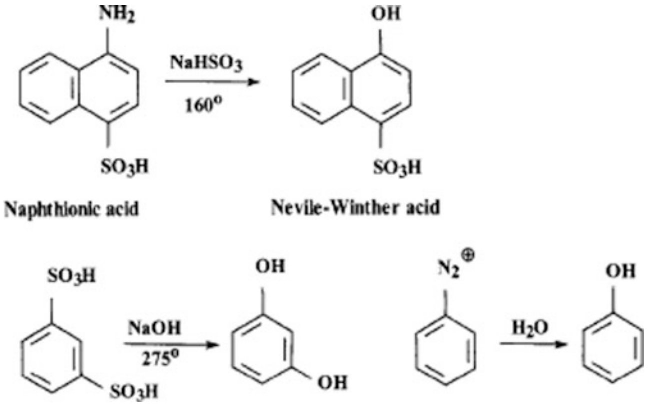


Fig. 13.70 Other examples of hydroxylation reactions

Oxidation

Although the oxidation of aromatic methyl groups can be conducted via the two-step sequence shown in Fig. 13.66, a convenient alternative process involves potassium dichromate.

In this case, the ring system involved must be stable to the conditions of the reactions. Another important oxidation reaction involves the conversion of naphthalene to phthalic anhydride, which can be accomplished using hot $KMnO_4$ or V_2O_5. These two reactions are illustrated in Fig. 13.71. Later we show that the oxygen in air can be used as the oxidant for certain organic dyes.

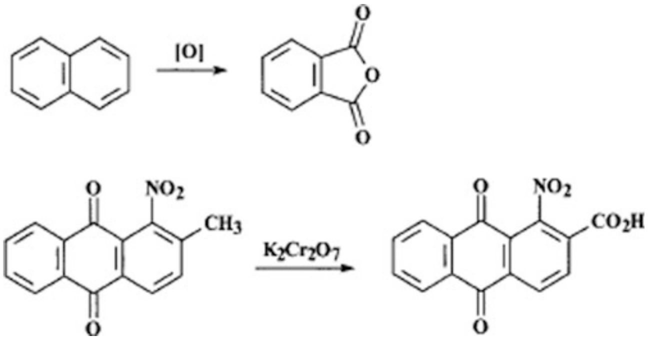


Fig. 13.71 Oxidation of naphthalene (*top*) and a methylated anthraquinone (*bottom*)

Other Important Reactions

Diazotization. The conversion of a primary aromatic amine to a diazonium compound is known as diazotization. Although this process is covered in more detail in our discussion of azo dye synthesis, it is worthwhile to point out that the diazonium group ($-N_2^+$) is used to produce a wide range of intermediates. As indicated in Fig. 13.72 diazotization is often achieved through the action of nitrous acid (HNO_2) and the resultant diazonium group can be replaced by various groups or reduced to give arylhydrazines.

Carboxylation. The introduction of carboxyl groups into the structures of phenols and naphthols produces some important dye intermediates, including salicylic acid and BON acid. This process is conducted under pressure at elevated temperatures using the sodium salts of phenols/naphthols and in the case of β-naphthol, the carboxyl group enters the 3-position (cf. Fig. 13.73). The free acid ($-CO_2H$) group is produced by acid treatment in the final step.

Fig. 13.72 Diazotization of aniline and its conversion to other types of intermediates

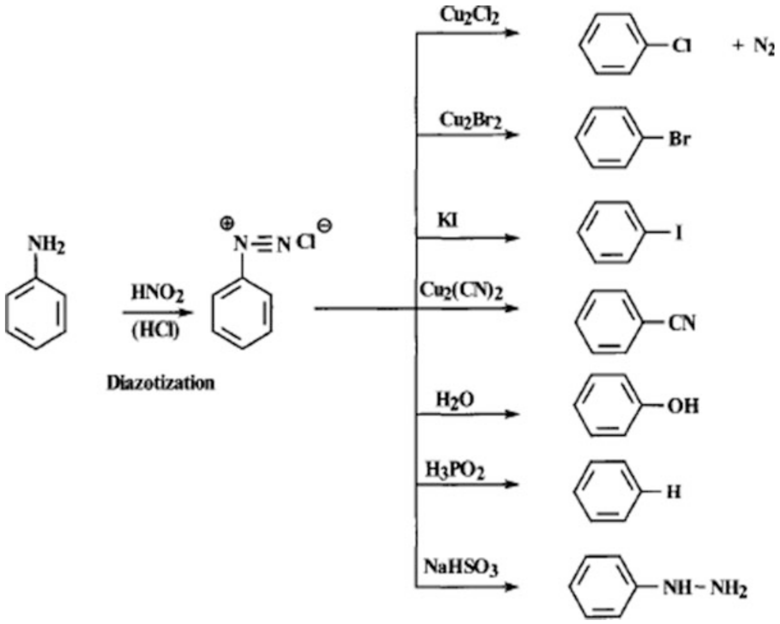
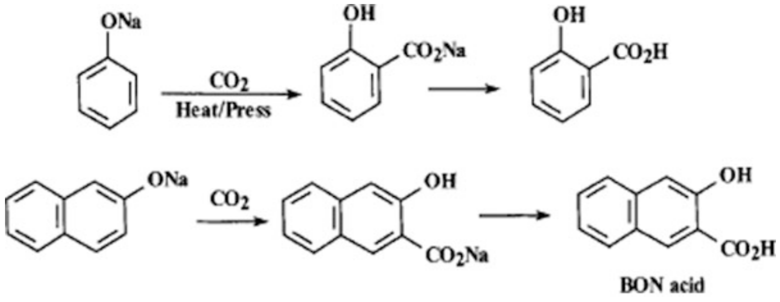


Fig. 13.73 Carboxylation of phenol and β-naphthol



Dye Manufacture

In this section, we summarize the principal methods of synthesis for different dye classes. Emphasis is placed on dyes presently in commerce and the industrial methods suitable for making them. Before doing so, we review the important principles that set dyes apart from other classes of organic compounds.

Unlike other organic compounds dyes possess color because they (1) absorb light in the visible spectrum (400–700 nm), (2) have at least one chromophore (color bearing group), (3) have a conjugated system (system of alternating double and single bonds), and (4) exhibit resonance (a stabilizing force in organic compounds). Table 13.5 shows the relationships between wavelength of visible light and color absorbed/seen and the other three factors are illustrated in Fig. 13.74, 13.75, and 13.76.

Table 13.5 Wavelength of light vs. color

Wavelength absorbed	Color absorbed	Color seen
400–435	Violet	Yellow-green
435–480	Blue	Yellow
480–490	Green-blue	Orange
490–500	Blue-green	Red
500–560	Green	Purple
560–580	Yellow-green	Violet
580–595	Yellow	Blue
595–605	Orange	Green-blue
605–700	Red	Blue-green

Concerning the various factors responsible for color in organic compounds, it is worthwhile to point out that the chromophore must be part of a conjugated system. This is illustrated through the examples in Fig. 13.77. When the azo group is connected to methyl groups the resultant compound is colorless. When it is attached to aromatic rings, the compound possesses color. Similarly, the structures in Fig. 13.75 illustrate the importance of having an extended conjugated system. In this case, doubling the length of the conjugated system for Vitamin A to give β-carotene causes the λ_{max} value to shift from 325 nm to 466 and 497 nm.

Nitro Dyes

As the name suggests, this very small class of organic dyes has at least one nitro group as the chromophore. Nitro dyes invariably are yellow or orange and are important for their economical cost and good lightfastness. Examples include the dyes shown in Fig. 13.78—C.I. Acid Orange 3 (a), C.I. Disperse Yellow 42 (b), C.I. Acid Yellow 1 (c), and C.I. Disperse Yellow 70 (d). A key disadvantage of nitro dyes is their low color strength ($\varepsilon_{max} = 5{,}000$–$7{,}000$). Improvements in color strength have been achieved by incorporating an azo group, as illustrated in dye D.

Representative syntheses are shown in Figs. 13.79 and 13.80. In the first example, C.I. Disperse Yellow 42 is prepared by condensing two molecules of aniline with one molecule of 4-chloro-3-nitrobenzenesulfonyl chloride, using ethanol as the solvent. In the second example, C.I. Acid Orange is prepared in a 3-step synthesis, starting from 2-chloro-5-nitrobenzenesulfonic acid.

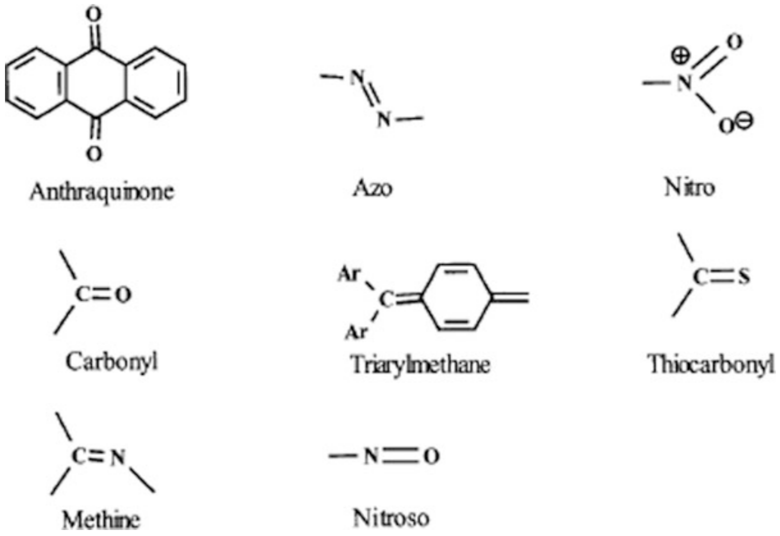


Fig. 13.74 Chromophores commonly found in organic dyes

Azo Dyes

Azo dyes are by far the largest family of organic dyes. They play a prominent role in acid, direct, reactive, azoic, and disperse dye structures, as shown previously, and include structures that cover the full color spectrum. Generally, the synthesis of azo dyes involves a process known as diazo coupling. In this process, a diazotized aromatic amine is coupled to a phenol, naphthol, aromatic amine, or a compound that has an active methylene group, as illustrated in the two-step synthesis in Fig. 13.81. Step 1 is the conversion of aniline to benzenediazonium chloride, a process known as diazotization, and step 2 is the reaction of the diazo compound with phenol to produce the corresponding azo dye, a process known as diazo coupling.

Diazotizations are normally conducted in an aqueous medium containing nitrous acid, generated in situ from HCl + $NaNO_2$, and a primary aromatic amine. When weakly basic or heteroaromatic amines are used in azo dye synthesis, H_2SO_4 is often used as the reaction medium, forming $H(NO)SO_4$ (nitrosylsulfuric acid) as the diazotizing agent [57]. The stoichiometry associated with this reaction is given in Fig. 13.82, and although only 2 moles of acid per mole amine are required, in practice 2.2–2.5 moles are used. Diazotizations are most often conducted at 0–10 °C because the resultant diazo compounds are usually unstable at higher temperatures.

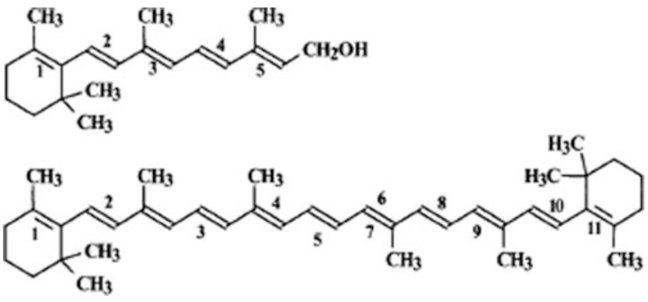


Fig. 13.75 Comparison of the conjugated systems in Vitamin A (*top*) and β-carotene (*bottom*)

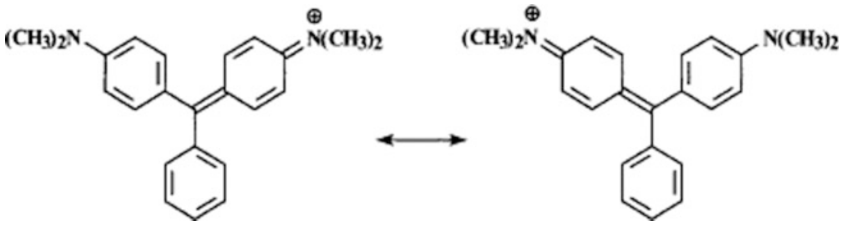
Examples of aromatic amines that can be diazotized are shown in Fig. 13.83. This extremely abbreviated list is designed to show that a wide variety of amines can be used, including hydrophobic, weakly basic, hydrophilic, and heterocyclic compounds. *ortho*-Diamines are not typically used because of their propensity to undergo triazole formation (Fig. 13.84).

Examples of compounds that can be used as coupling components in azo dye synthesis are shown in Figs. 13.85, 13.86, and 13.87. The first group comprises phenols and naphthols, the second group comprises amines that couple, and the third contains couplers that have an active methylene group (see Fig. 13.87). Compounds in the first and third groups require ionization using alkali, to give sufficient reactivity for diazo coupling, and the pH employed is usually 8–9. Because aromatic amines are appreciably more reactive, they couple at pH 5–6. Arrows have been used to indicate the coupling positions for the various couplers. Compounds such as 1-naphthol or 1-naphthylamine give a mixture of monoazo dyes by coupling in the 2-position or the 4-position. When couplers containing –OH and $–NH_2$ groups are employed (see Fig. 13.85), coupling may occur twice, giving disazo dyes. In such cases, coupling is first conducted in acid, *ortho* to the $–NH_2$ group, and then in alkali. This is important because the introduction of the first azo group decreases the reactivity of the coupler. The ability to ionize the –OH group provides sufficient ring activation for the second coupling. In the case of gamma acid, one has the lone option of coupling under acidic or alkaline conditions.

When primary amines are used as couplers, coupling can occur on the ring or at the amino group itself unless the amino group is blocked. One good way to block this group is by converting it to the *N*-sulfomethyl group, as illustrated in Fig. 13.86. The products formed are also known as omega salts [58]. The blocking group can be removed following the coupling step, by treating the resultant azo dye with an alkaline solution.

There are also important examples of phenolic compounds that do not couple (see Fig. 13.88). In these examples, the required coupling positions are blocked, the ring is too deactivated, or the compounds undergo oxidation in the presence of the diazo compound.

Fig. 13.76 Resonance structures for Malachite Green (C.I. Basic Green 4)



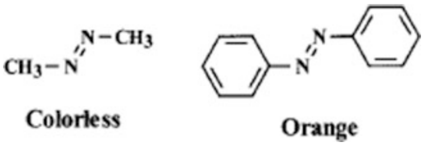


Fig. 13.77 Impact of having a chromophore apart from (*left*) or part of (*right*) a conjugated system

The synthesis of azo dyes can be illustrated using the following letter designations:

- A = Diazotizable amine
- D = Tetrazotizable diamine
- E = Coupler that couples once
- M = 1° Amine that couples once and is diazotized and coupled again

Fig. 13.78 Examples of nitro dye structures

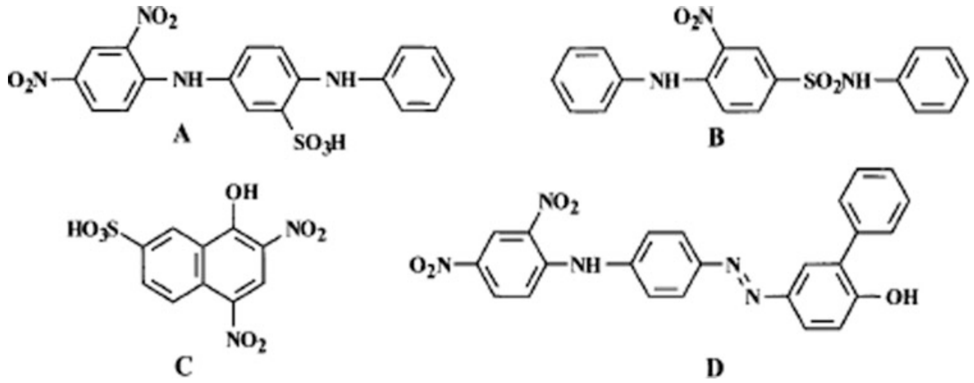


Fig. 13.79 Synthesis of Disperse Yellow 42

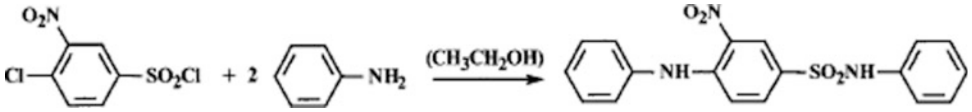


Fig. 13.80 Synthesis of Acid Orange 3

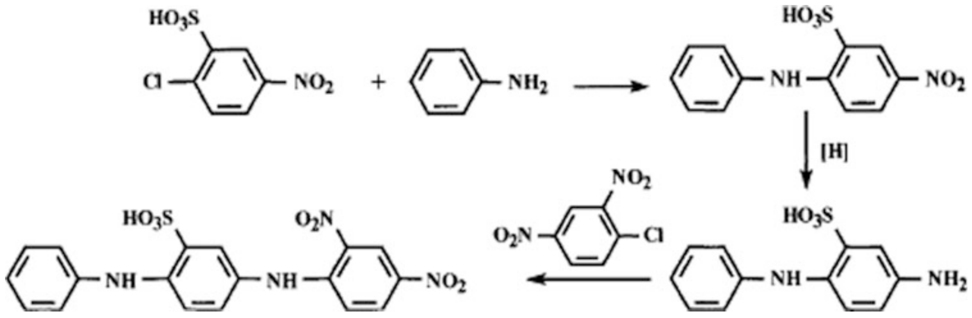


Fig. 13.81 Two-step synthesis of an azo dye from aniline and phenol

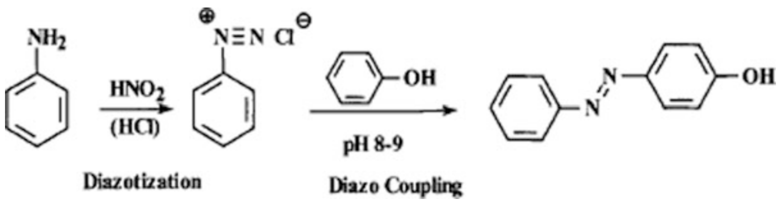


Fig. 13.82 Summary of the diazotization process

$$\mathrm{Ar{-}NH_2 + 2\ HCl + NaNO_2 \longrightarrow Ar{-}\overset{\oplus}{N}{\equiv}N\ Cl^{\ominus} + 2\ H_2O + NaCl}$$

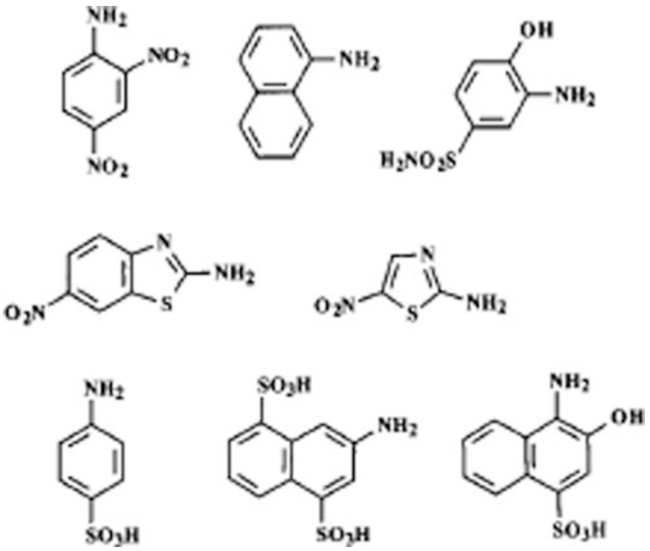


Fig. 13.83 Representative aromatic amines used in azo dye synthesis

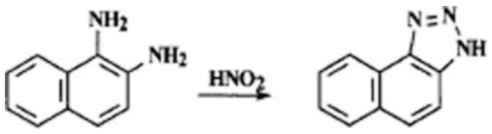
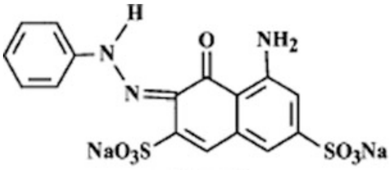


Fig. 13.84 Triazole formation from the diazotization of an *ortho*-diamine

- Z = Coupler that couples twice
- Z·X·Z = Binuclear coupler that couples twice



Acid Red 33

These designations are used to provide an indication of how a given dye has been assembled, and will be used in describing the azo dye syntheses covered in the subsections that follow.

Monoazo Dyes

Azo dyes of this type are manufactured predominantly by the reaction between a diazotized amine ("A") and a type "E" or "Z" coupler. The synthesis can be as simple as coupling diazotized aniline to H-acid, in an A → E process, to produce C.I. Acid Red 33. An example of a reactive dye that is manufactured via an A → E process is C.I. Reactive Red 1. In this case, the target dye is manufactured as shown in Fig. 13.89, which shows that the reactive group can be

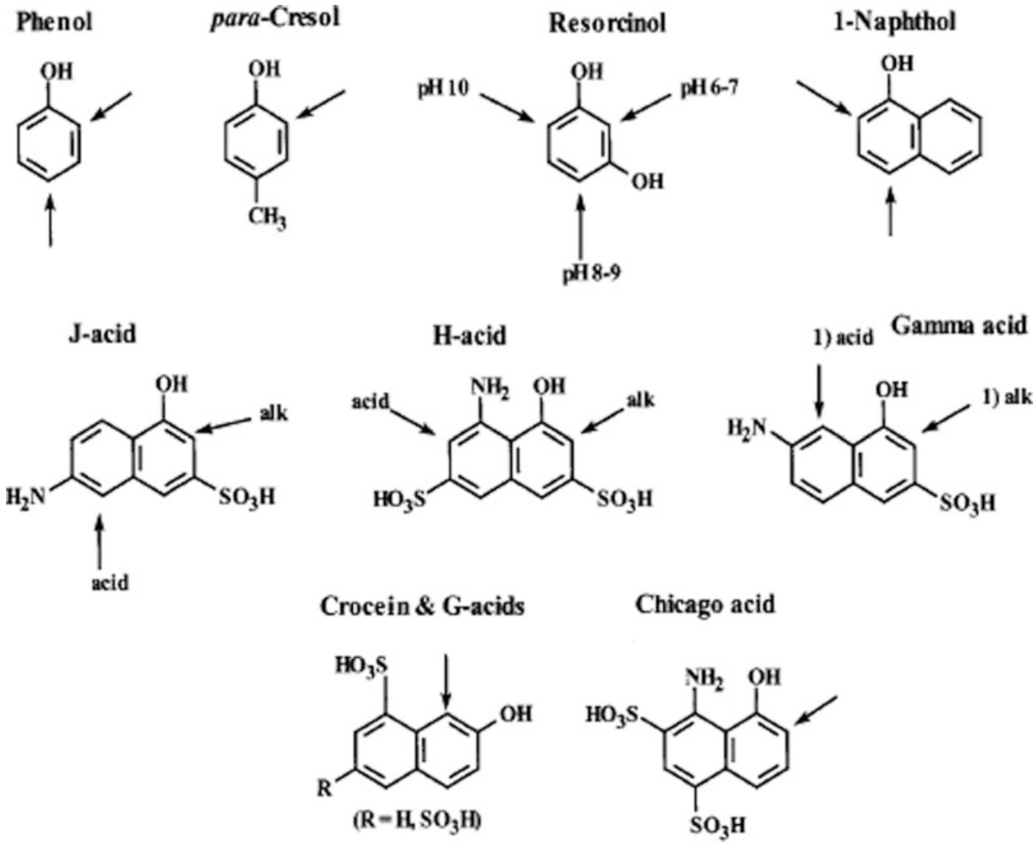


Fig. 13.85 Examples of phenol and naphthol-based couplers

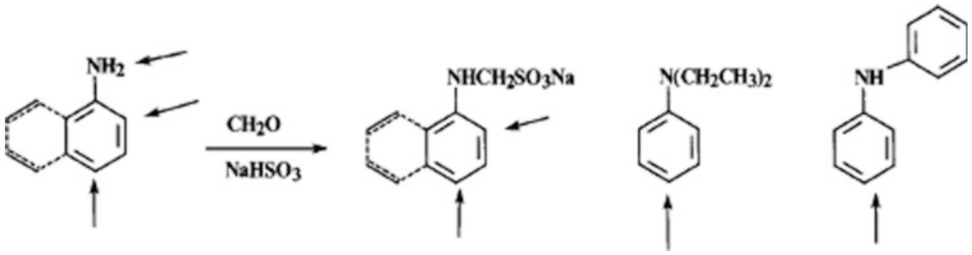


Fig. 13.86 Examples of aromatic amine couplers

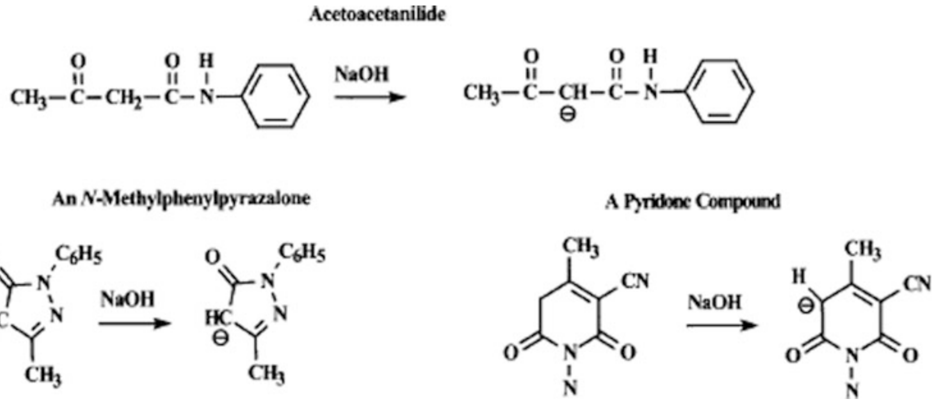


Fig. 13.87 Examples of couplers having an active methylene ($-CH_2-$) group

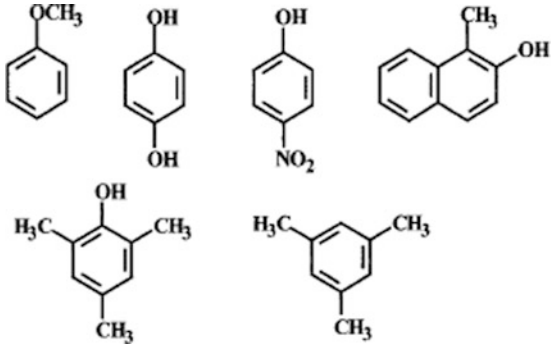


Fig. 13.88 Examples of compounds that do not couple

introduced prior to (sequence 1) or after (sequence 2) the coupling step. Similarly, monoazo bireactive dyes are made by this process (Fig. 13.90). This illustrates that a quite complex arylamine can be used as the diazo compound.

Other examples of monoazo dyes that are synthesized via an A → E process are shown in Fig. 13.91, further illustrating the wide range of structural types that can be manufactured this way.

Disazo Dyes

There are four often-used methods for synthesizing dyes containing two azo linkages, each of which requires two diazo coupling reactions. A nontraditional "disazo" dye involves 1:2 metal complex formation.

Type $A^1 \rightarrow Z \rightarrow A^2$ *synthesis*. Dyes of this type include those shown in Fig. 13.92 (C.I. Acid Black 1 (**7**), C.I. Mordant Brown 1 (**8**), and C.I. Acid Black 17 (**9**), C.I. Direct Orange 18 (**10**)), in which couplers such as H-acid, resorcinol, and *meta*-phenylenediamine are coupled twice. Although A^1 and A^2 are different in the present examples, they need not be different. As pointed out above, coupling *ortho* to the amino group of H-acid is usually conducted first, under weakly acidic conditions, followed by coupling with diazotized aniline under alkaline conditions. This is also true for the structurally similar dye **9**, which is prepared from S-acid. In the case of dye **10**, however, coupling with aniline under slightly acidic conditions is the second step. For dye **8**, both couplings are conducted under acidic conditions, with 2-amino-4-nitrophenol introduced first.

Type $E^1 \rightarrow D \rightarrow E^1$ *synthesis*. Dyes of this type require the conversion of an aryldiamine to a tetrazonium compound (one that has two diazonium groups) in a process know as tetrazotization. See Fig. 13.93 which involves environmentally friendly alternatives to benzidine. Following tetrazotization, one tetrazonium molecule reacts with two coupler molecules to produce the target dye, examples of which are provided in Fig. 13.94 (**11**: C.I. Direct Red 28 (Congo Red), **12**: C.I. Direct Yellow 12, and **13**: C.I. Acid Yellow 42). Disazo dyes prepared this way include dye **14**

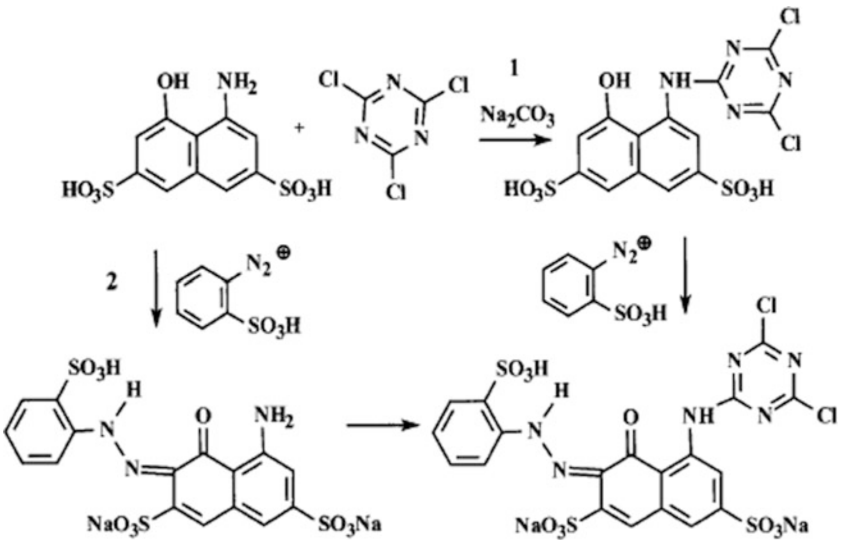


Fig. 13.89 Two approaches to the synthesis of Reactive Red 1

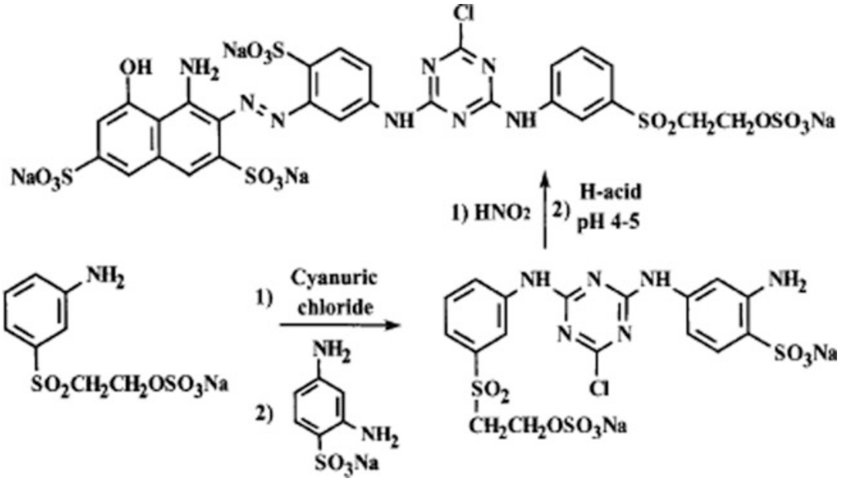


Fig. 13.90 Synthesis of a monoazo dye containing two different reactive groups

(C.I. Direct Blue 15), which is converted to the important bis-copper complex, C.I. Direct Blue 218 (see Fig. 13.95).

Type A → M → E synthesis. This is one of the largest groups of disazo dyes, as they include acid, disperse, direct, and reactive dye structures. A representative synthesis is shown in Fig. 13.96. The second diazotization and coupling steps can be conducted inside certain textile fibers. For instance, disperse black dyes are produced in the presence of cellulose acetate by conducting the chemistry shown in Fig. 13.97 after dyeing cellulose acetate with the monoazo dye.

Examples of dyes made via an A → M → E synthesis are shown in Fig. 13.98. Although most azo disperse dyes are based on monoazo structures, disazo structures such as **15** (C.I. Disperse Orange 13) and **16** (C.I. Disperse Orange 29) are manufactured. An important direct dye of this type is **17** (C.I. Direct Red 81), a reactive dye is **18** (C.I. Reactive Blue 40), and acid dyes include **19** (C.I. Acid Red 151) and **20** (C.I. Acid Blue 116).

Type $A^1 \rightarrow Z \cdot X \cdot Z \leftarrow A^1$ synthesis. Disazo dyes of this type are produced from coupling twice to dye intermediates such as those shown in Fig. 13.99, and are largely direct dyes for cotton. A representative synthesis is shown in Fig. 13.100, for C.I. Direct Red 83. In this case the target dye is prepared by metallization after the coupling step.

Disazo dyes such as C.I. Direct Yellow 44 are prepared according to the sequence shown in Fig. 13.101. In this example, a pair of monoazo dyes is reacted with phosgene.

1:2 Metal complexes. Although somewhat different from the previous examples and methods, dyes containing two azo groups can also be synthesized by forming 1:2 metal complexes of suitably substituted monoazo dyes.

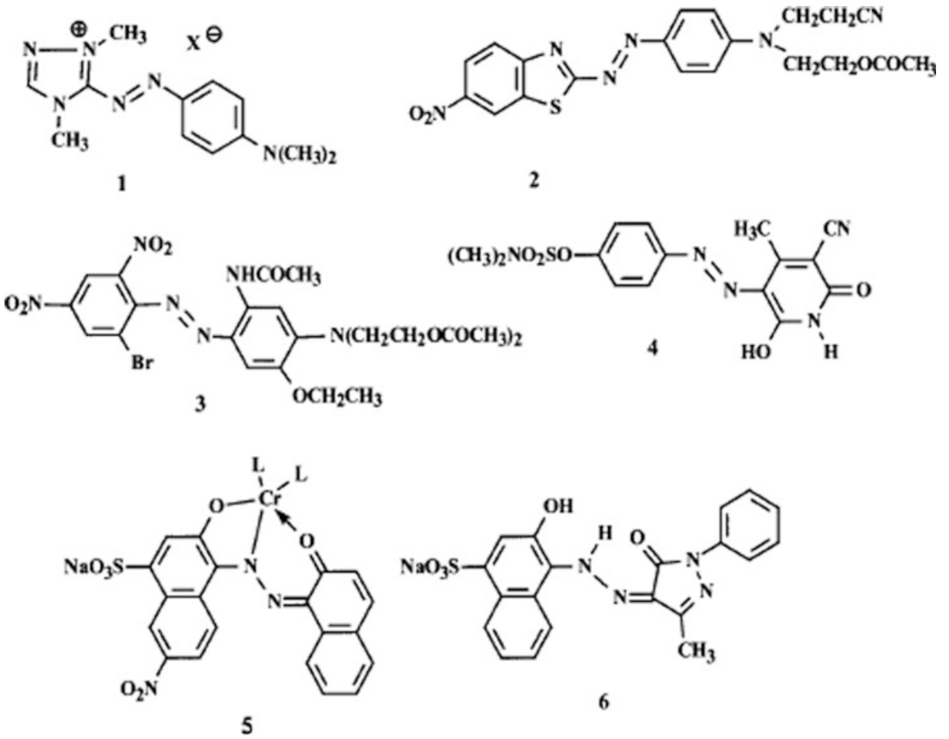


Fig. 13.91 Monoazo cationic (**1**), disperse (**2**–**4**), 1:1 chromium complexed (**5**), and mordant (**6**) dyes prepared via an A → E process

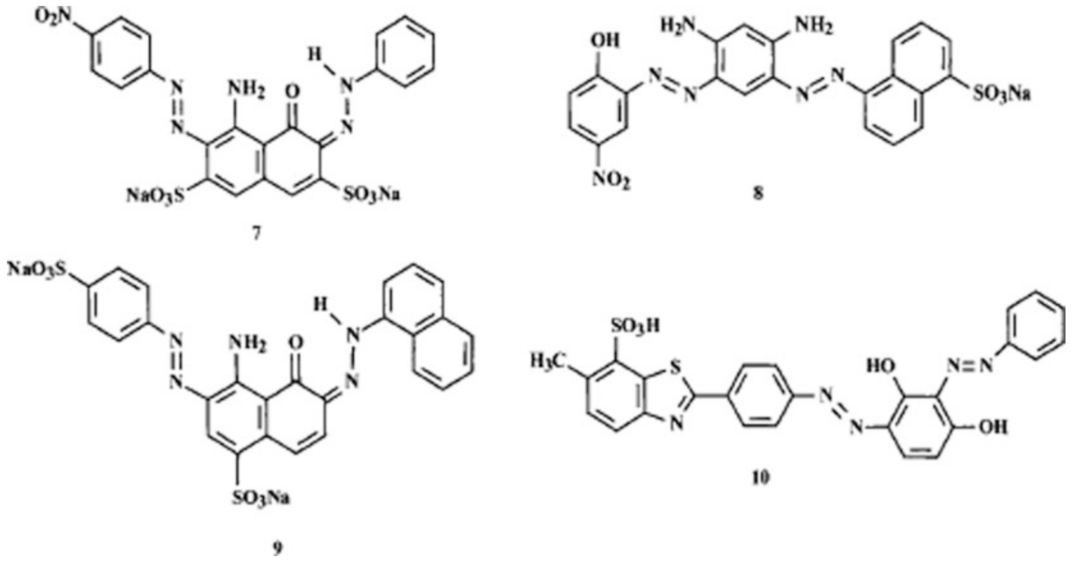


Fig. 13.92 Disazo dyes prepared by coupling twice to H-acid (**7**), *meta*-phenylenediamine (**8**), S-acid (**9**), and resorcinol (**10**)

The resultant dyes are mostly acid dyes for protein and polyamide substrates and the metals employed are Cr, Co, and Fe. Examples shown in Fig. 13.102 are for C.I. Acid Black 172 (**21**) and C.I. Acid Yellow 151 (**22**). In these examples, the corresponding monoazo dye is treated with one-half the molar amount of $Cr_2(SO_4)_3$ or $CoCl_3$, respectively.

Polyazo Dyes

In this section, we cover the synthesis of dyes containing three or more azo linkages. In this regard, methods for producing trisazo dyes (those having three azo linkages) include $E \leftarrow D \rightarrow Z \leftarrow A$ and $A \rightarrow M^1 \rightarrow M^2 \rightarrow E$ syntheses.

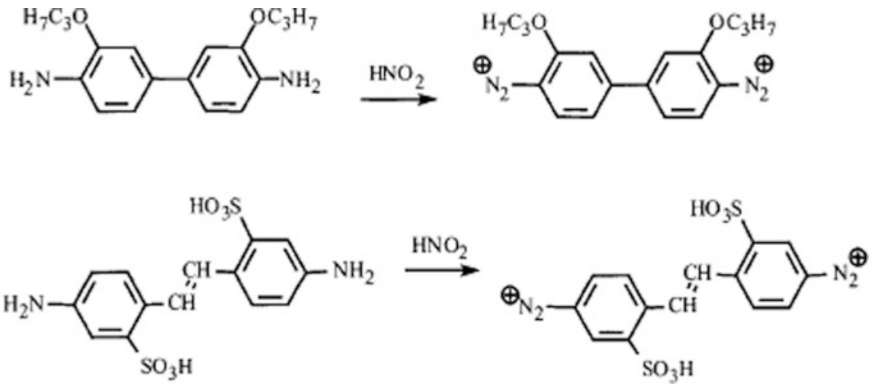
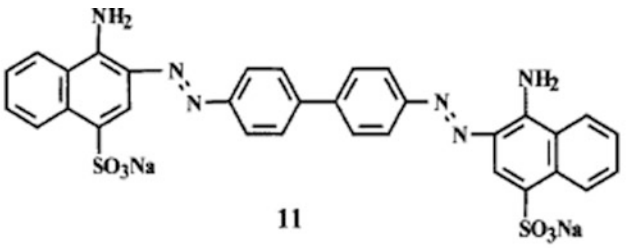
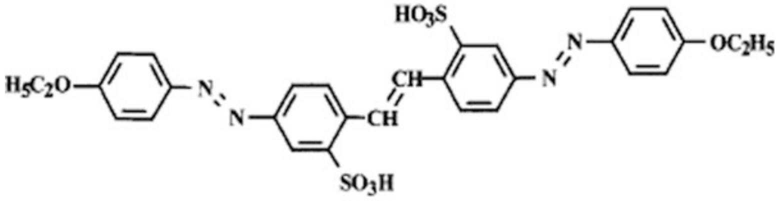


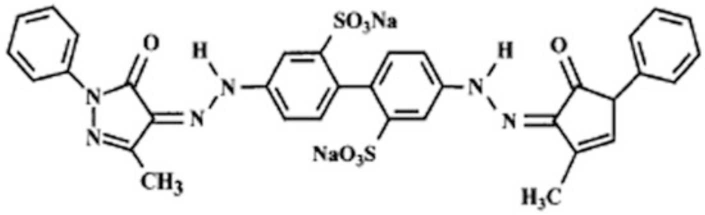
Fig. 13.93 Tetrazotization of a di-*n*-propoxybenzidine (*top*) and a diaminostilbene disulfonic acid (*bottom*)



11



12



13

Fig. 13.94 Examples of disazo dyes (**11**–**13**) prepared via a type $E^1 \leftarrow D \rightarrow E^1$ synthesis

Examples are shown in Fig. 13.103 for C.I. Acid Black 234 (**23**) and C.I. Direct Blue 71 (**24**). In the synthesis of dye **23**, an unsymmetrical dye can be made from diamine **25** because the end of the tetrazonium compound (cf. **26**) that is *para* to the $–SO_2$ moiety is more reactive than the one that is *para* to the –NH moiety (Fig. 13.104).

Dyes containing four azo linkages are direct dyes for cotton and can be prepared in several ways, including via $A \rightarrow M \rightarrow Z \leftarrow D \rightarrow E$, $A^1 \rightarrow Z^1 \leftarrow D \rightarrow Z^2 \leftarrow A^2$, $E^1 \leftarrow M^1 \leftarrow D \rightarrow M^2 \rightarrow E^2$, $E^1 \leftarrow D^1 \rightarrow Z \leftarrow D^2 \rightarrow E^2$, and $E^1 \leftarrow D \rightarrow M^1 \rightarrow M^2 \rightarrow E^2$ sequences. Examples of the second and third methods are shown in Fig. 13.105. Note that both are symmetrical molecules, the first of which (C.I. Direct Brown 44) employs *meta*-phenylenediamine as a type “Z” coupler and a type “D” diazo component. In the second example (C.I. Direct Black 22), gamma acid is twice used as the “M” moiety, and the dye is synthesized by (1) coupling tetrazotized benzidine disulfonic acid to two molecules of gamma acid, (2) diazotizing the amino groups on the gamma acid moieties, and (3) coupling to two molecules of *meta*-phenylenediamine.

Fig. 13.95 Formation of Direct Blue 218

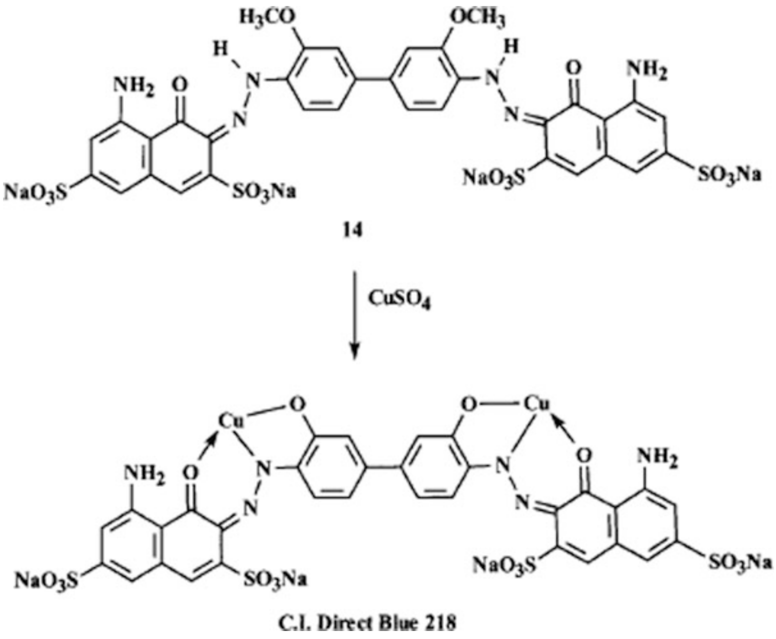
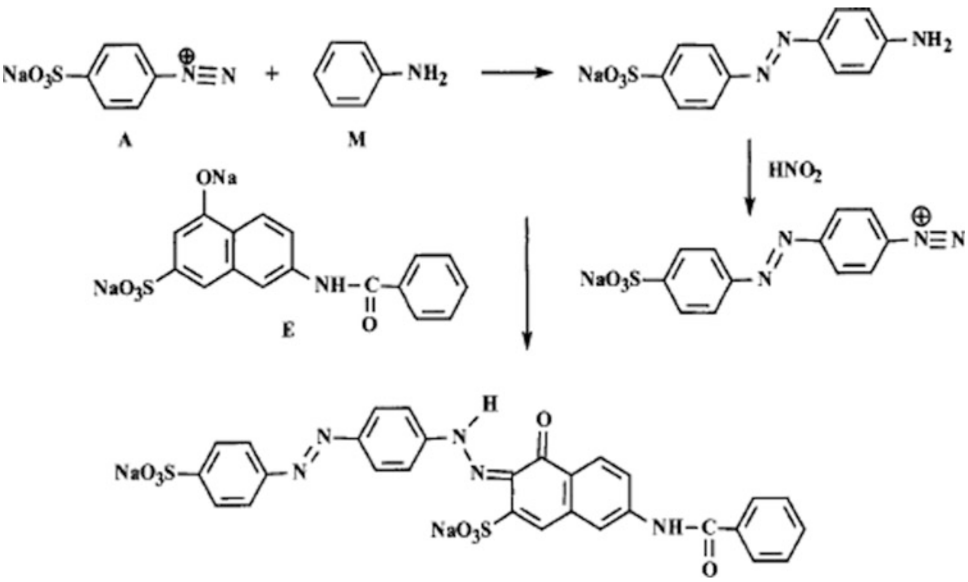


Fig. 13.96 Disazo dye synthesis via an A → M → E process



Triphenylmethane Dyes

Triphenylmethane dyes are usually prepared in two steps: (1) condensation of an *N*,*N*-dialkylaniline with a benzaldehyde compound and (2) oxidation of the resultant leuco base (**27**). The synthesis of C.I. Basic Green 4 (Malachite Green) is given as an example in Fig. 13.106. Alternatively, C.I. Acid Green 50 is prepared in three steps: (1) condensation of *N*,*N*-dimethylaniline and *para*-(*N*,*N*-dimethylamino)benzaldehyde to produce Michler's hydrol (**28**), condensation with R-acid to give an intermediate leuco base (**29**), and (3) oxidation to give the target dye. Historically, PbO_2 has been used as the oxidizing agent. However, concerns about its toxicity have led to the use of a more environmentally friendly agent such as tetrachloro-*para*-benzoquinone (chloranil).

In another synthetic variation, C.I. Acid Violet 17 is prepared in the four steps shown in Fig. 13.107. The different steps in this process are the synthesis of the *N*-arylmethyl intermediate **30** and the diphenylmethane intermediate **31**. Oxidation to the intermediate hydrol and condensation with *N*,*N*-dimethylaniline produce the target dye.

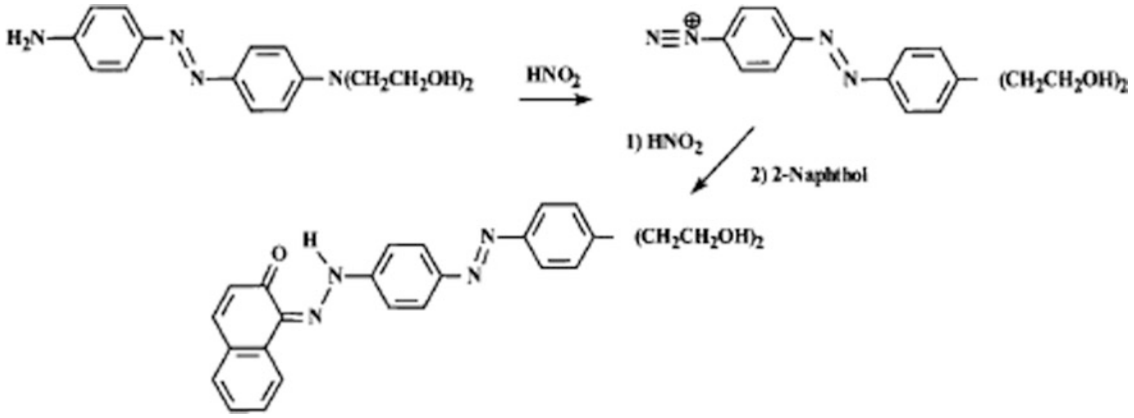


Fig. 13.97 Disazo disperse black dye synthesis conducted inside cellulose acetate fibers

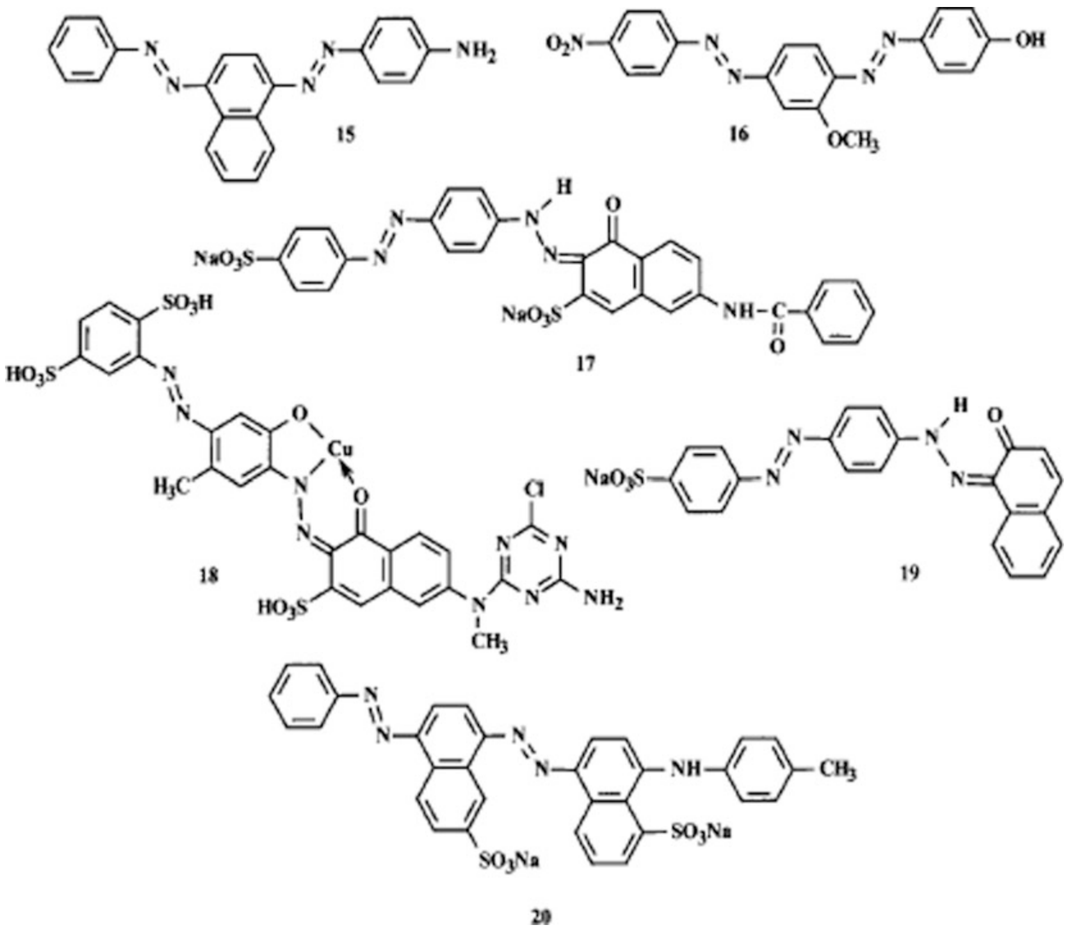


Fig. 13.98 Examples of disazo dyes prepared by a type A → M → E synthesis

Structurally related dyes are synthesized by condensing phenols with phthalic anhydride to give a colorless intermediate lactone (**32**) that reacts with alkali to give the colored form. An example of this dye type is phenolphthalein, the synthesis of which is shown in Fig. 13.108.

Xanthene Dyes

Like phenolphthalein, xanthene dyes are prepared in a condensation reaction involving phthalic anhydride. However, resorcinol is employed instead of phenol. The simplest

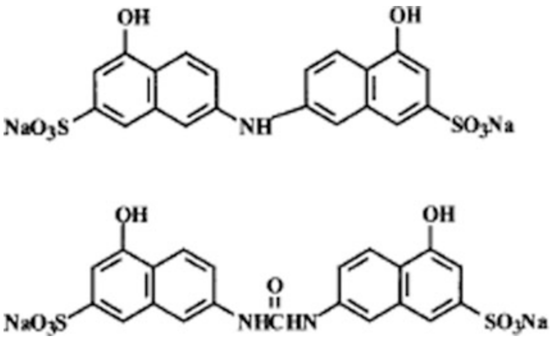


Fig. 13.99 Structures of J-acid imide (*top*) and J-acid urea (*bottom*)

representative of this family is C.I. Acid Yellow 73 (fluorescein), which is made via the sequence of steps shown in Fig. 13.109. Similarly, C.I. Acid Red 92 is made by the condensation of tetrachlorophthalic anhydride and resorcinol followed by bromination.

Anthraquinone and Related Dyes

The commercial preparation of anthraquinone dyes begins with the synthesis of anthraquinone itself. In this regard, the three-step synthesis involves: (1) the oxidation of

Fig. 13.100 Synthesis of Direct Red 83 via an $A^1 \rightarrow Z{\cdot}X{\cdot}Z \leftarrow A^1$ synthesis and metallization

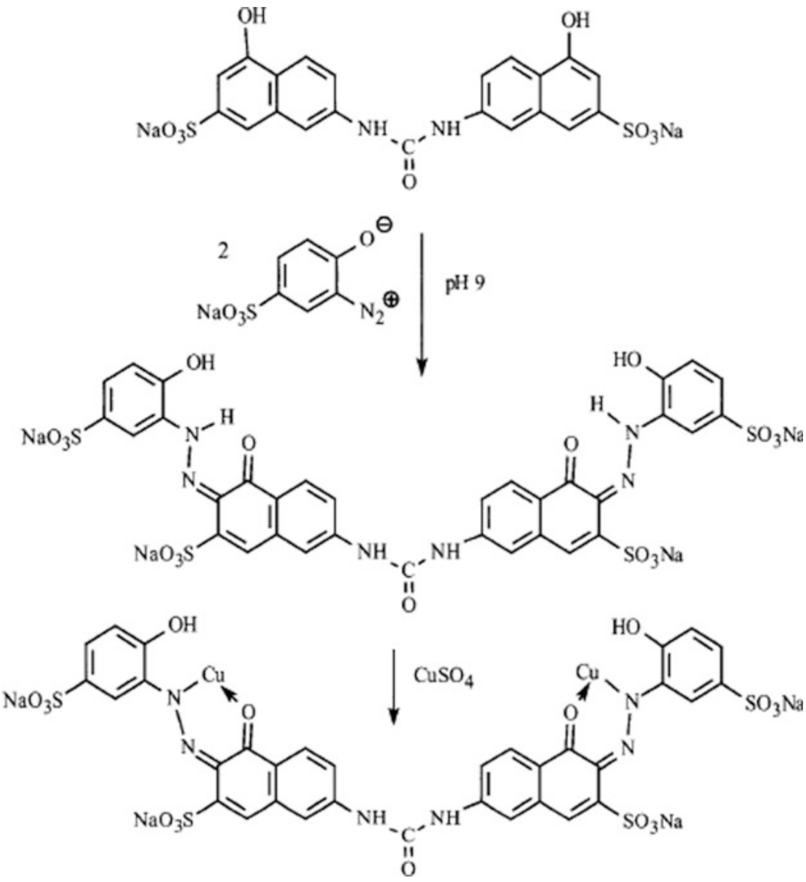


Fig. 13.101 An alternative route to type $A^1 \rightarrow Z{\cdot}X{\cdot}Z \leftarrow A^1$ disazo dyes

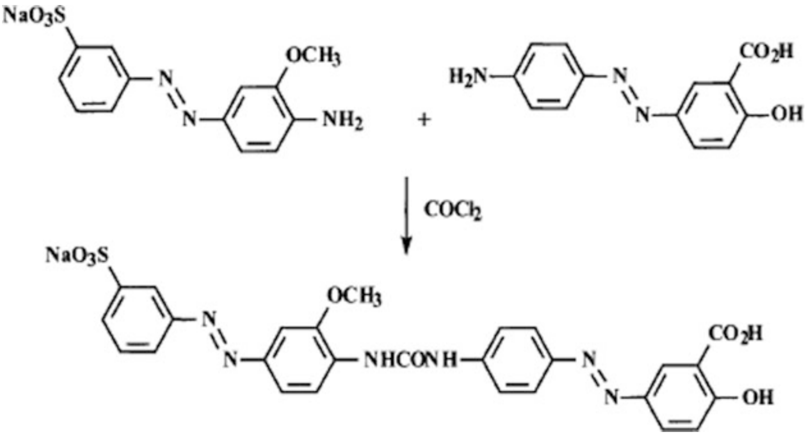
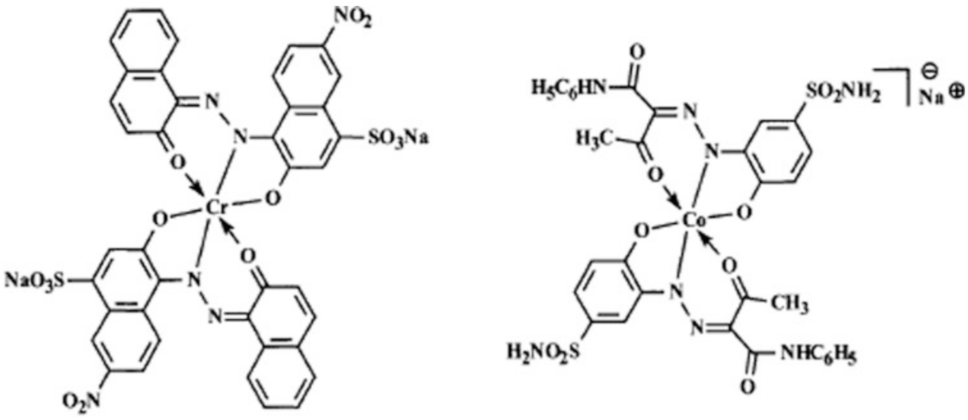


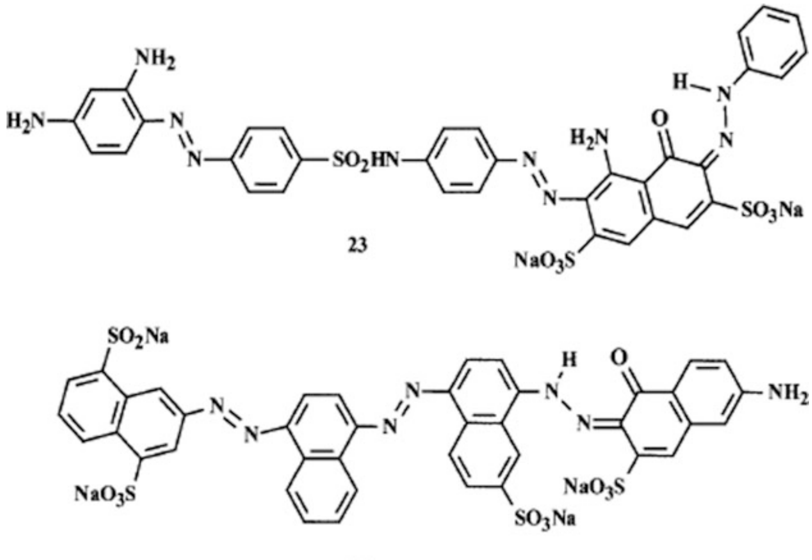
Fig. 13.102 Representative 1:2 bisazo metal complexed dyes



21

22

Fig. 13.103 Examples of trisazo dyes



24

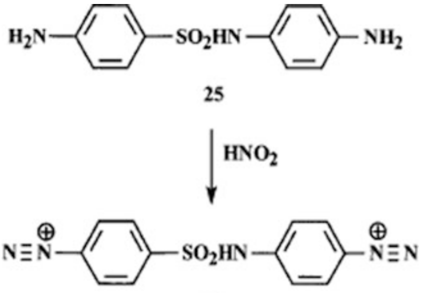


Fig. 13.104 Structures of compounds **25** and **26**

naphthalene to phthalic anhydride, (2) Friedel–Crafts acylation of benzene to give a keto acid, and (3) cyclodehydration using H_2SO_4. See Fig. 13.110. The preparation of 1,4-disubstituted anthraquinones utilizes the intermediates prepared in Fig. 13.111, where R = OH corresponds to quinizarin.

The reduction of quinizarin using sodium hydrosulfite produces leuco quinizarin, which, in turn, undergoes condensation with alkyl- or arylamines and reoxidation to produce blue and green disperse and solvent dyes. Although chemical oxidation can be used, air oxidation is normally sufficient. See steps “A” and “B” in Fig. 13.112 for the general reaction scheme. The use of boric acid in the reduction step follows the course outlined in Fig. 13.113, where the synthesis of C.I. Solvent Green 3 is given as an example [59].

Anthraquinone Disperse Dyes

Examples of dyes prepared using the above methods are shown in Fig. 13.114. The C.I. disperse dyes Red 15, Violet 1,

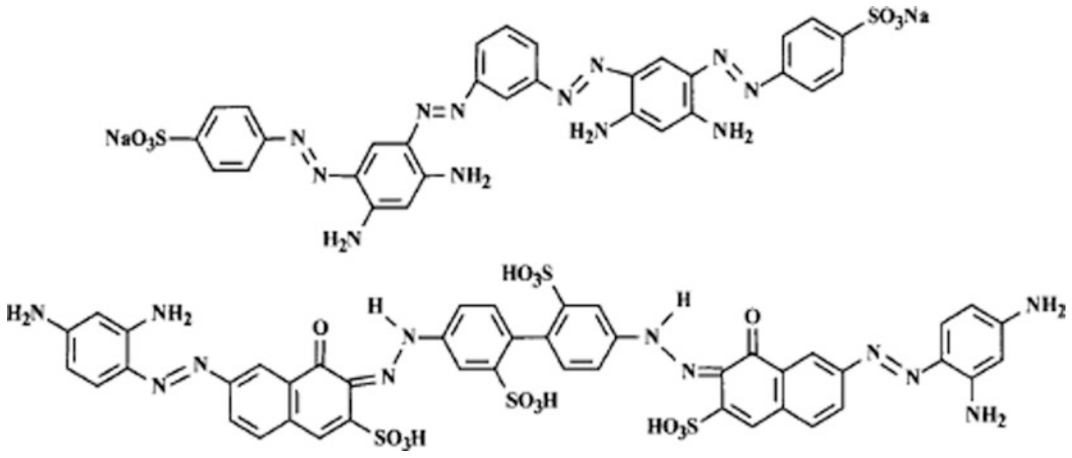


Fig. 13.105 Representative polyazo dyes prepared via the $A^1 \rightarrow Z^1 \leftarrow D \rightarrow Z^2 \leftarrow A^2$ (*top*) and $E^1 \leftarrow M^1 \leftarrow D \rightarrow M^2 \rightarrow E^2$ (*bottom*) methods

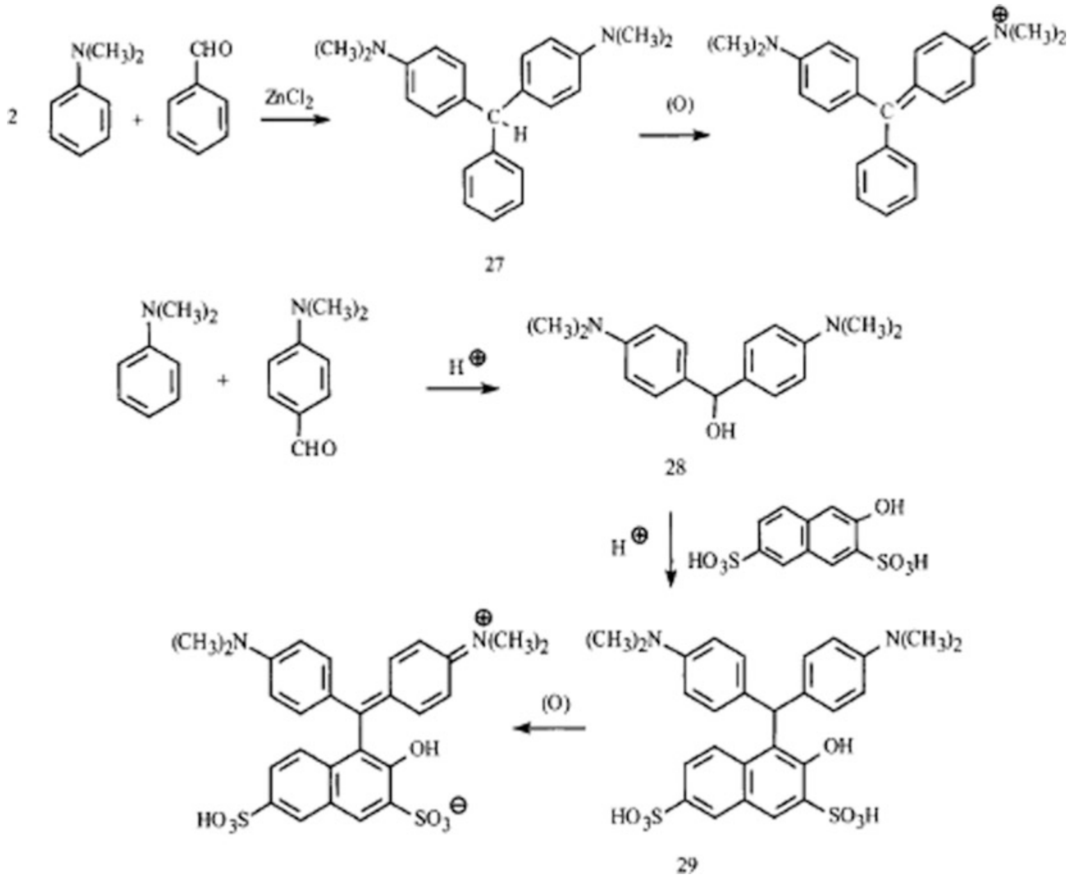


Fig. 13.106 Two-step synthesis of Malachite Green (Basic Green 4). Three-step synthesis of Acid Green 50

Blue 3, Violet 27, Blue 19, and Blue 23, are prepared from leucoquinizarin. When unsymmetrical dyes such as Disperse Blue 3 are made, the use of a mixture of two amines in the condensation step gives the corresponding symmetrical dyes as by-products. In this case, Disperse Blue 23 would be one of the by-products.

The synthesis of Disperse Red 4 employs the dibromoanthraquinone intermediate **33**, which is hydrolyzed to

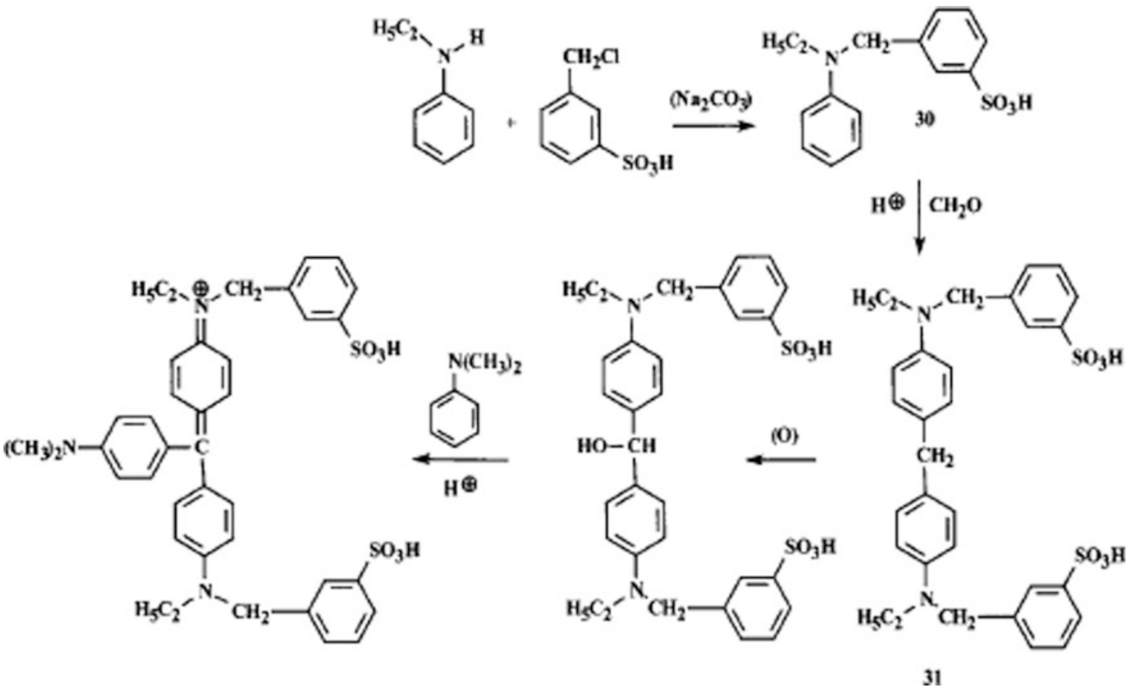


Fig. 13.107 Synthesis of Acid Violet 17 via diarylmethane intermediate **31**

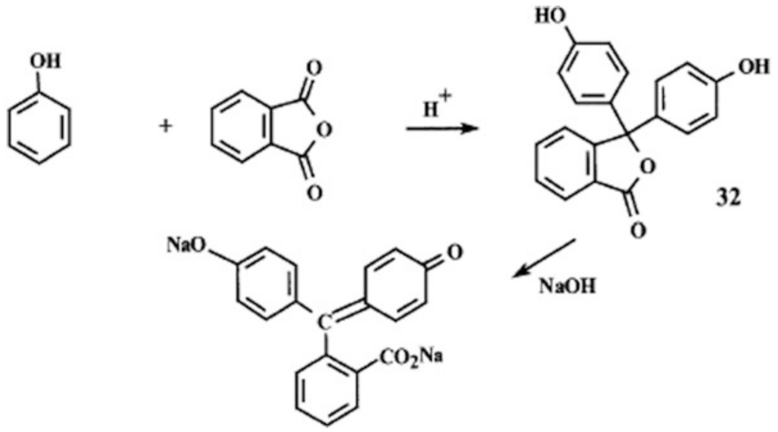


Fig. 13.108 Synthesis of phenolphthalein via colorless lactone **32**

compound **34** and converted to the target dye upon alcoholysis. See Fig. 13.115. The synthesis of Disperse Violet 26 is conducted in two steps: (1) chlorination of Disperse Violet 1 in the 2, 3-positions using SO_2Cl_2 and (2) condensation with phenol.

Disperse dyes containing substituents in both anthraquinone rings are often prepared from dinitroanthrarufin (DNA) and dinitrochrysazin (DNC), the structures of which are shown in Fig. 13.116. Examples of these dyes are C.I. Disperse Blue 56 and Blue 77. The former dye is made by reduction of DNA followed by bromination, and the latter is made by condensing aniline with DNC. The DNC condensation shows that nitro groups in the α-position can be displaced like a halogen.

The dichlorinated precursor for Disperse Violet 26 can be used to make turquoise blue dyes such as C.I. Disperse Blue 60, as shown in Fig. 13.117. In this sequence, the chloro groups are replaced by cyano groups, using NaCN, and the resultant intermediate (**35**) is hydrolyzed to give the corresponding imide (**36**), which in turn is alkylated to give the target dye.

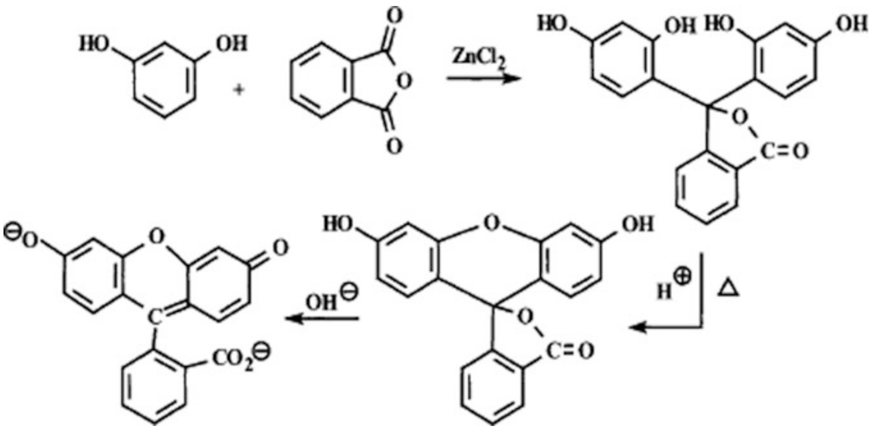


Fig. 13.109 Synthesis of Acid Yellow 73, a xanthene dye

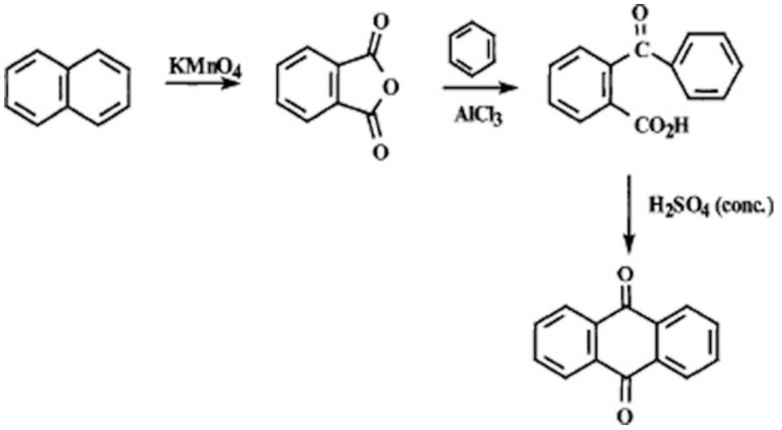


Fig. 13.110 Synthesis of the anthraquinone ring system

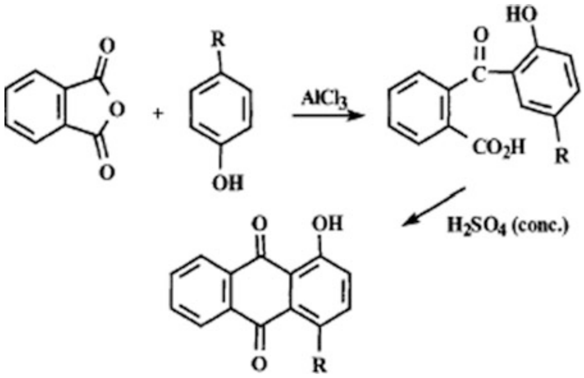


Fig. 13.111 Synthesis of key substituted anthraquinone intermediates, where R = Cl, OH

Anthraquinone Acid Dyes

A key intermediate in the synthesis of anthraquinone acid dyes is bromamine acid. This compound is made via the sequence shown in Fig. 13.118. Acid dyes made from this intermediate include C.I. Acid Blue 25, C.I. Acid Blue 40, and C.I. Acid Blue 127.

The synthesis of C.I. Acid Blue 127 takes place according to the route shown in Fig. 13.119. A key step in the synthesis is the formation of diamine **37**, which is produced in two

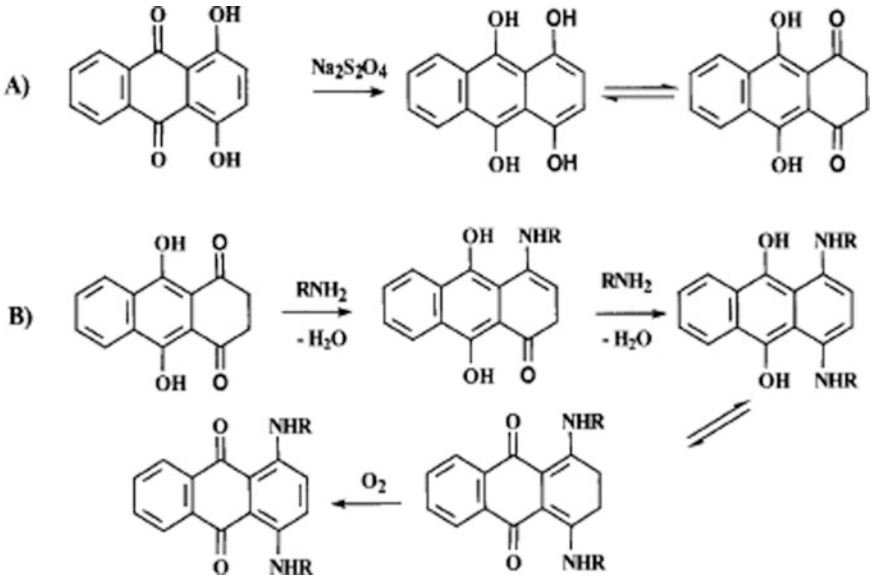


Fig. 13.112 General reaction scheme for the synthesis of 1,4-diaminoanthraquinone dyes

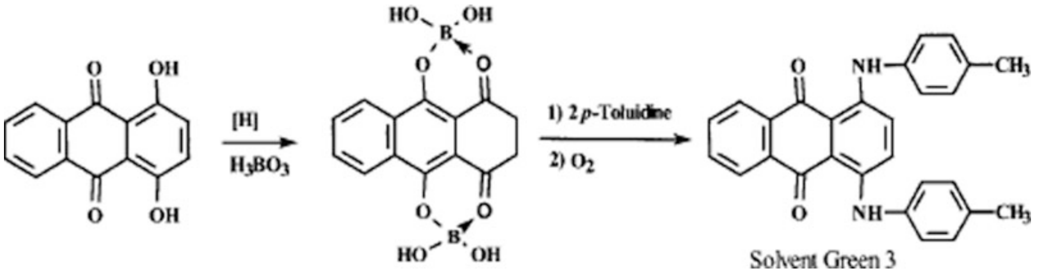
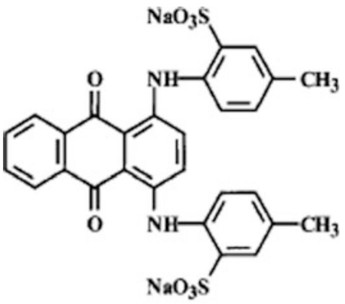


Fig. 13.113 Boric acid catalyzed synthesis of Solvent Green 3

steps from *N*-sulfomethylaniline: (1) condensation with acetone and (2) hydrolysis to remove the protecting group. At this point, one molecule of diamine **37** is condensed with two molecules of bromamine acid to form the dye.

Another important dye is C.I. Acid Green 25. This dye is made by the sulfonation of C.I. Solvent Green 3 (Fig. 13.113). Because the benzene rings are more reactive than the anthraquinone system, sulfonation occurs there preferentially.



Acid Green 25

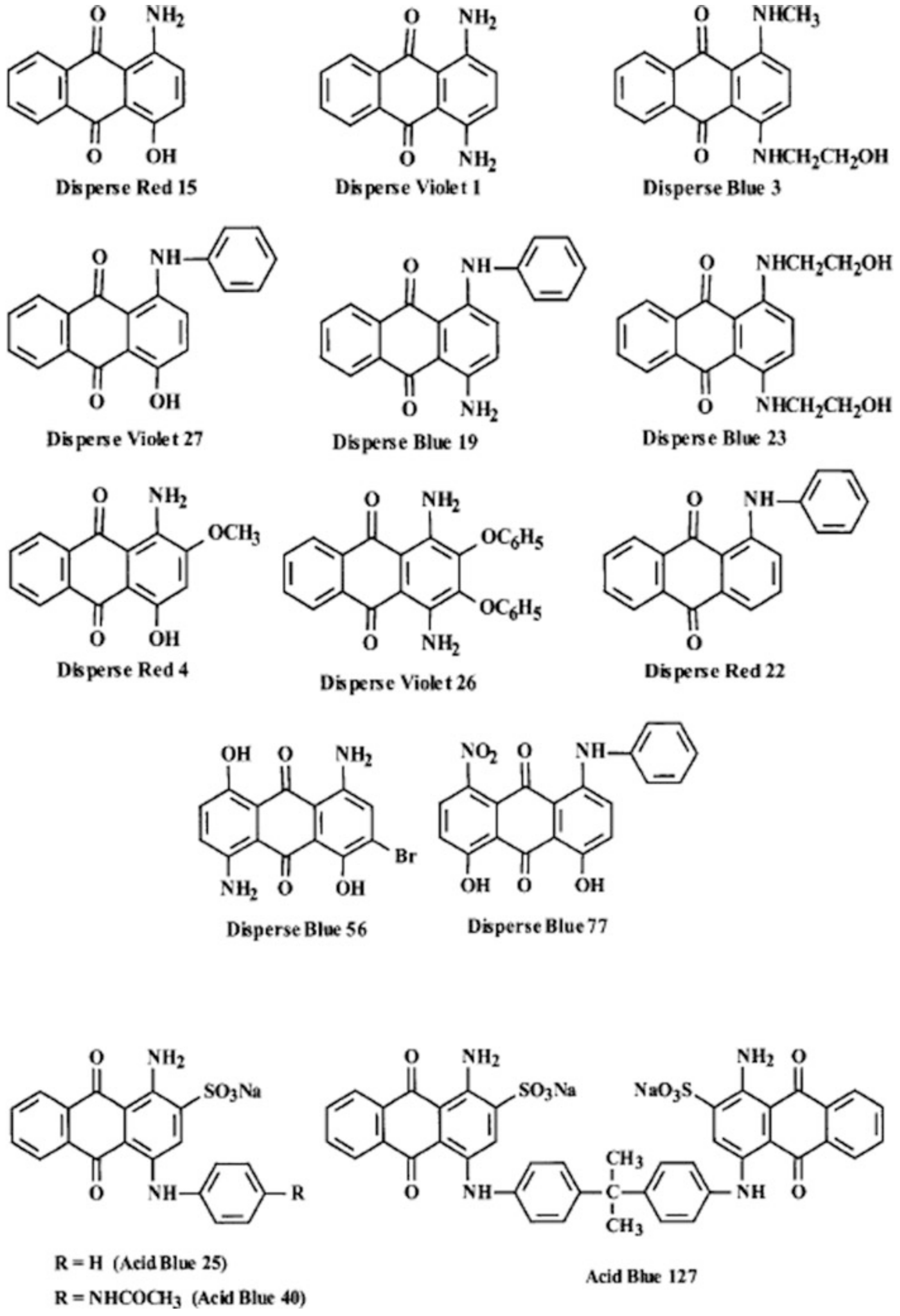
Anthraquinone Basic Dyes

Dyes of this type include C.I. Basic Blue 22 and Basic Blue 47. The synthesis of Basic Blue 22 is shown in Fig. 13.120, as an example of the type of chemistry required. The sequence begins with the preparation of *N*,*N*-dimethylpropylenediamine, which in turn is combined with methylamine and condensed with leucoquinizarin. Oxidation gives the key intermediate **38**, which is alkylated using methyl chloride to produce the dye.

Anthraquinone Reactive Dyes

Three examples of dyes of this type are C.I. Reactive Blue 19 (**39**), Reactive Blue 2 (**40**), and Reactive Blue 4 (**41**). All three dyes can be synthesized by condensing the appropriate arylamine with bromamine acid. In the case of the high-volume dye Reactive Blue 19, arylamine **44** is the key

Fig. 13.114 Representative 1,4-disubstituted anthraquinone dyes



Disperse Red 15

Disperse Violet 1

Disperse Blue 3

Disperse Violet 27

Disperse Blue 19

Disperse Blue 23

Disperse Red 4

Disperse Violet 26

Disperse Red 22

Disperse Blue 56

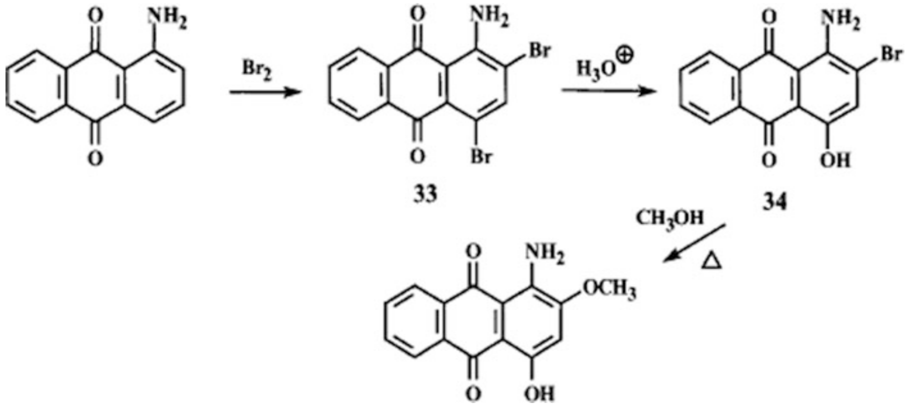
Disperse Blue 77

R = H (Acid Blue 25)

R = NHCOCH$_3$ (Acid Blue 40)

Acid Blue 127

Fig. 13.115 Synthesis of Disperse Red 4



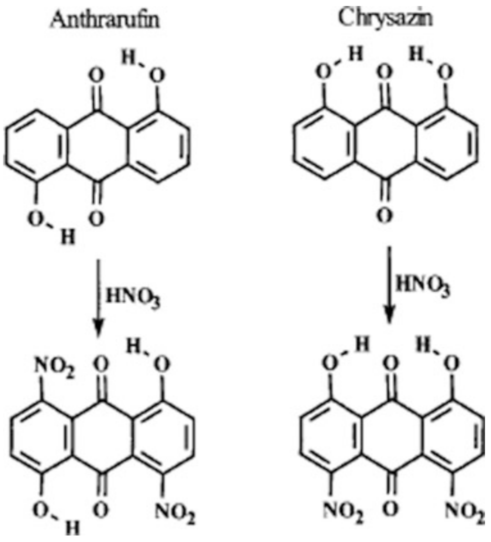


Fig. 13.116 Formation of DNA and DNC

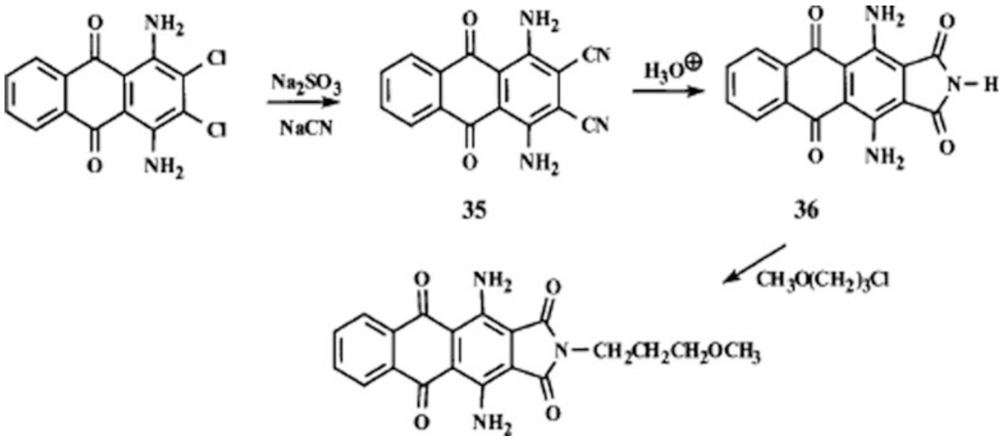


Fig. 13.117 A three-step synthesis of Disperse Blue 60

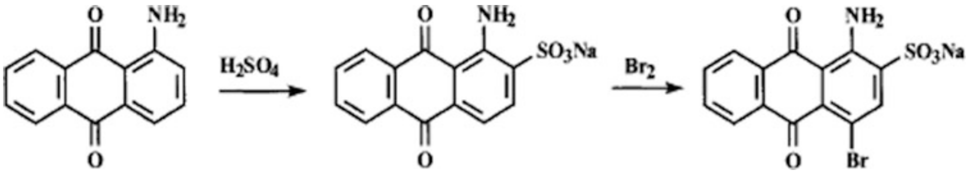


Fig. 13.118 A two-step synthesis of bromamine acid

intermediate, and its synthesis is shown in Fig. 13.121. Chlorosulfonation and then reduction of the intermediate sulfonyl chloride produce the sulfinic acid **42** Alternatively, the reduction step can be conducted with $Na_2S_2O_4$. Alkylation of the sulfinic acid with 2-chloroethanol or ethylene oxide (a more toxic agent) produces Compound **43**. Treatment of this compound with hot H_2SO_4 gives simultaneous hydrolysis of the acetamido (−NHAc) group and sulfonation of the hydroxyethyl ($-CH_2CH_2OH$) group to give key intermediate **44**.

Similarly, dyes **40** and **41** are prepared by condensing 2,5-diaminobenzenesulfonic acid with bromamine acid,

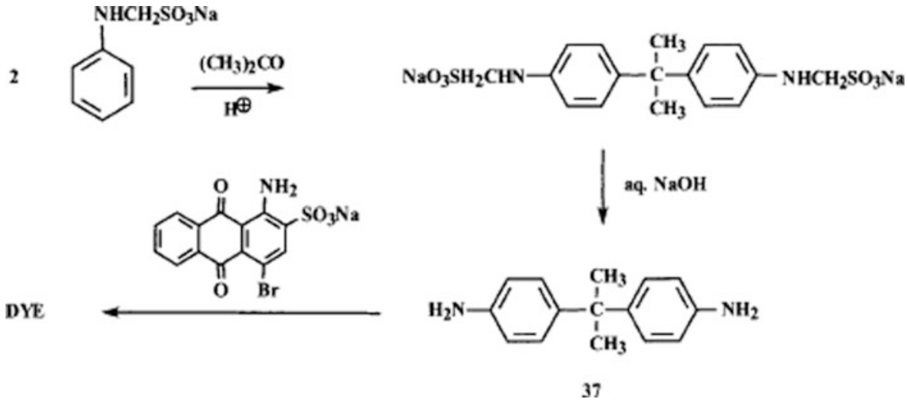


Fig. 13.119 Synthesis of Acid Blue 127

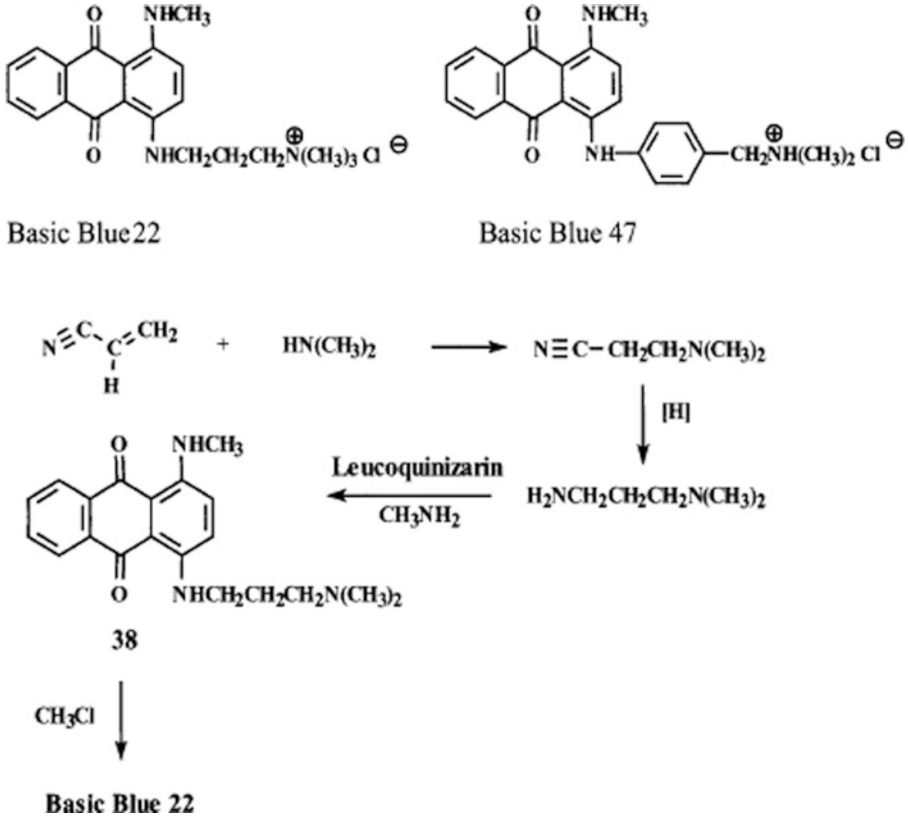


Fig. 13.120 Synthesis of Basic Blue 22

which reacts first at the less hindered amino group, followed by a reaction with cyanuric chloride to introduce the reactive group. These steps produce dye **41** and dye **40** is formed by reacting **41** with a mixture of sulfonated anilines. See Fig. 13.122.

Vat Dyes

The synthesis of vat dyes covers the full gamut of simple to complex chemistry. We have chosen examples to illustrate the broad spectrum of possible structures and synthetic methods. Emphasis is placed on anthraquinone vat dyes, because they dominate the number of commercial dyes.

Anthraquinone

The simplest anthraquinone vat dyes are benzoylated amines such as C.I. Vat Yellow 3 (**45**) and Vat Yellow 33 (**46**). The syntheses are shown in Figs. 13.123 and 13.124.

Anthraquinone vat dyes containing a thiazole ring include C.I. Vat Yellow 2, the synthesis of which is shown in Fig. 13.125. In this case, at least two approaches are

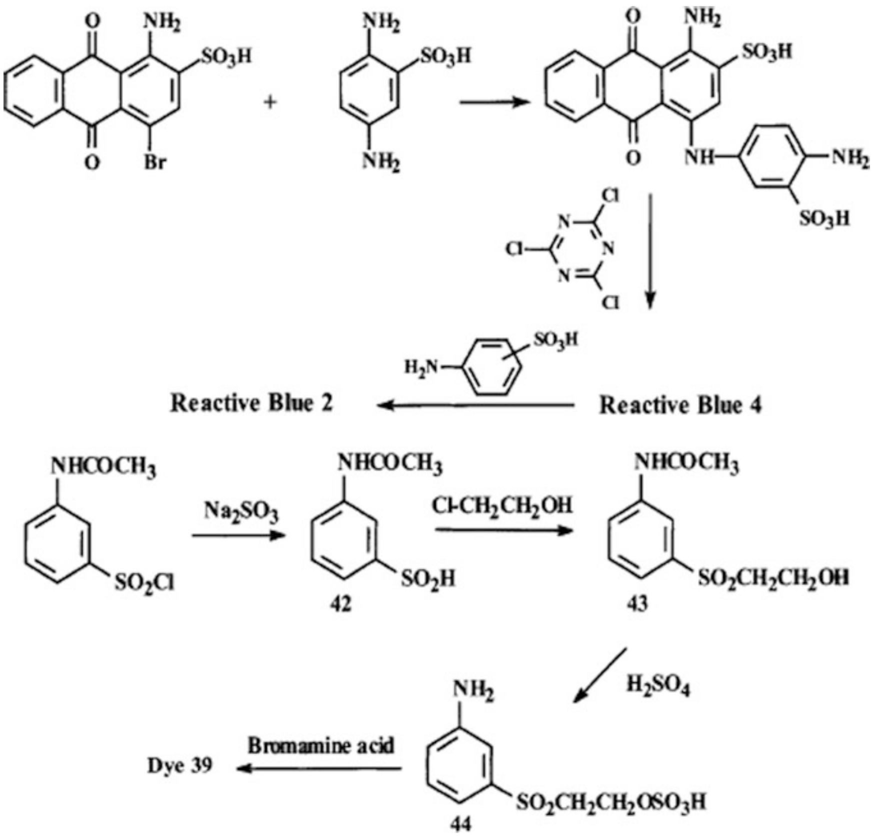


Fig. 13.121 Synthesis of three reactive blue dyes from bromamine acid

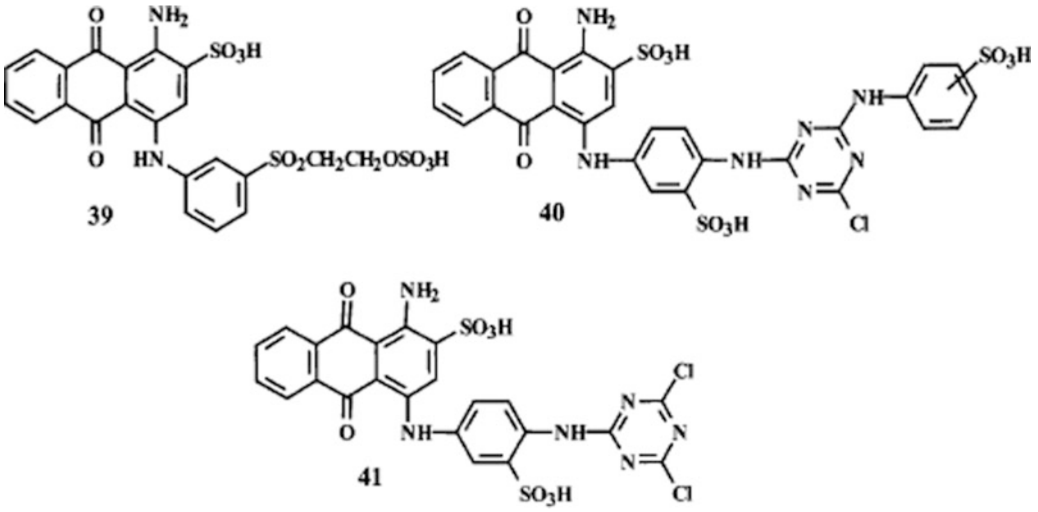
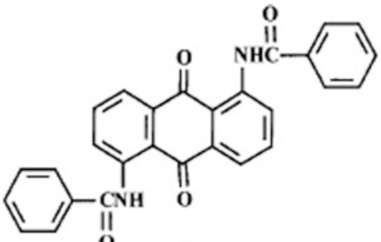


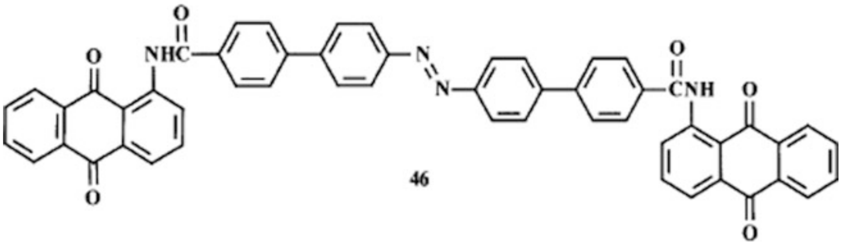
Fig. 13.122 Structures of Reactive Blue 19 (**39**). Reactive Blue 2 (**40**), and Reactive Blue 4 (**41**)

possible. In the first, 2,6-diaminoanthraquinone is condensed with benzotrichloride in the presence of sulfur and the initial product is oxidized without isolation to give the target dye. Alternatively, the starting diamine can be chlorinated and converted to the corresponding dithiol (**47**). At this point condensation with benzaldehyde followed by oxidation (e.g., air or dichromate) gives the dye.

Important vat dyes containing a carbazole moiety include C.I. Vat Brown 3 and Vat Black 27. These dyes are made according to the method shown in Fig. 13.126 for Vat Brown 3.



45



46

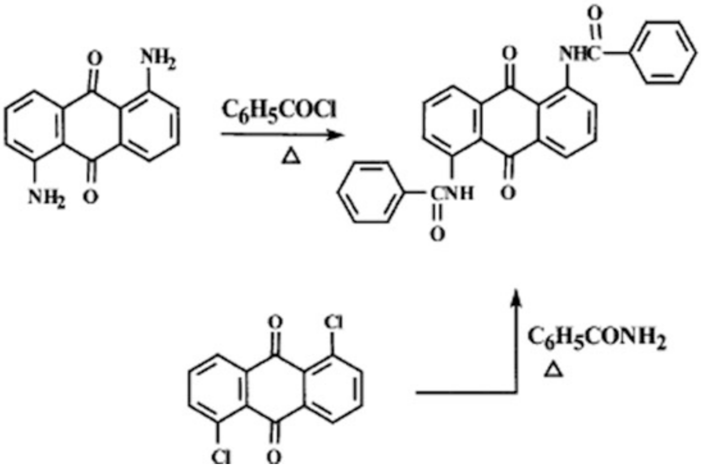


Fig. 13.123 Two approaches to the synthesis of Vat Yellow 3

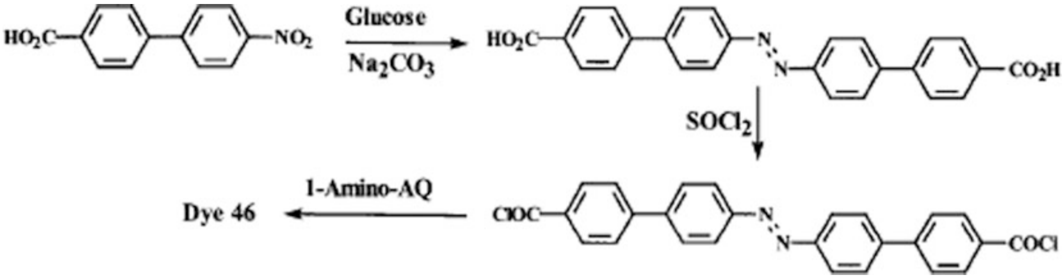


Fig. 13.124 Synthesis of Vat Yellow 33, where AQ = anthraquinone

The synthesis employs an Ullmann-type condensation reaction between compounds **48** and **49** followed by acid-induced cyclization using H_2SO_4.

Vat dyes that do not contain all of the elements of the anthraquinone moiety include benzanthrone-based vat dyes such as C.I. Vat Orange 1 and Vat Green 1, which are made according to the routes shown in Figs. 13.127 and 13.128. The first synthesis is a three-step process: (1) dibenzoylation of naphthalene, (2) Lewis acid-induced cyclization to the benzanthrone system, and (3) dibromination. The second

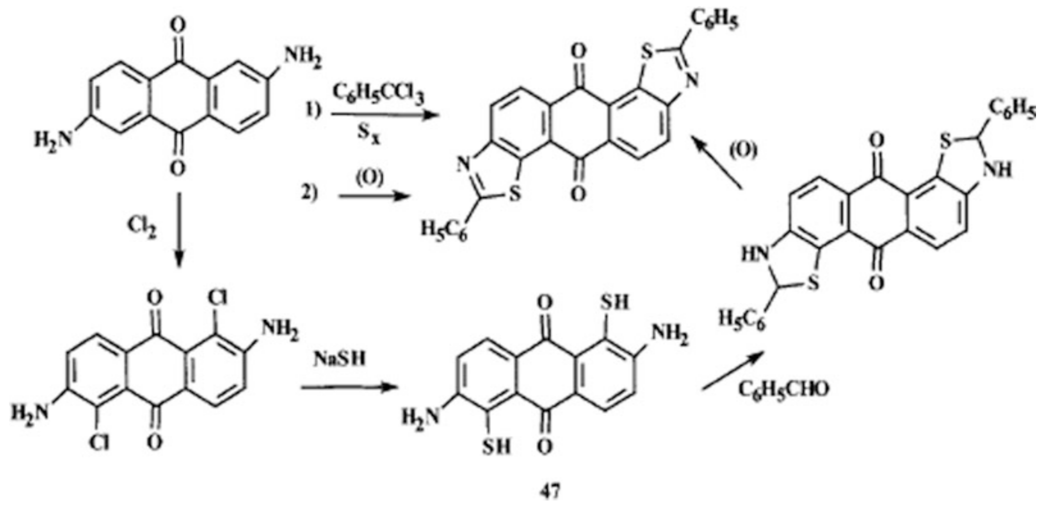


Fig. 13.125 Alternatives to the synthesis of Vat Yellow 2

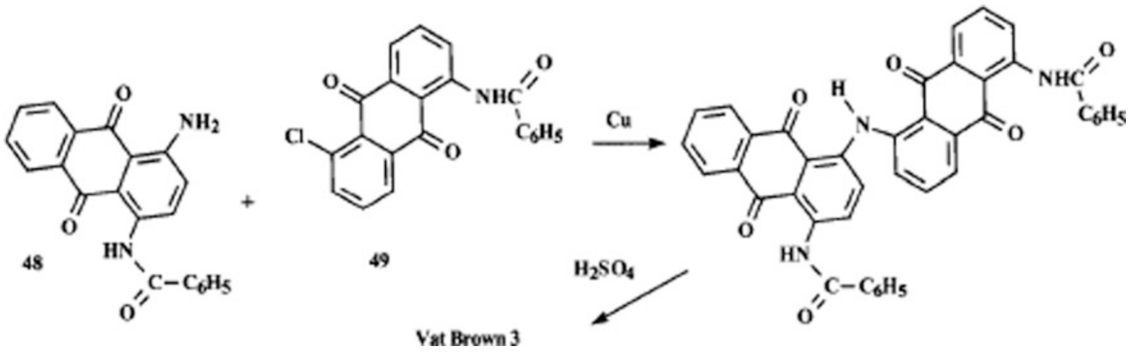


Fig. 13.126 Synthesis of carbazole-based anthraquinone vat dyes

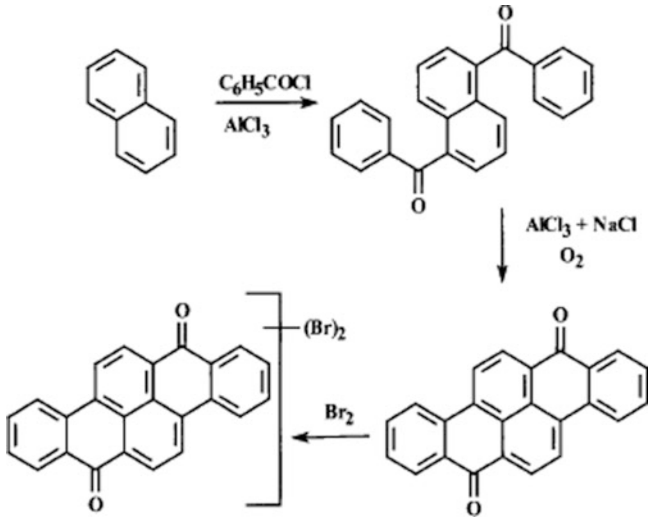


Fig. 13.127 Three-step synthesis of Vat Orange 1

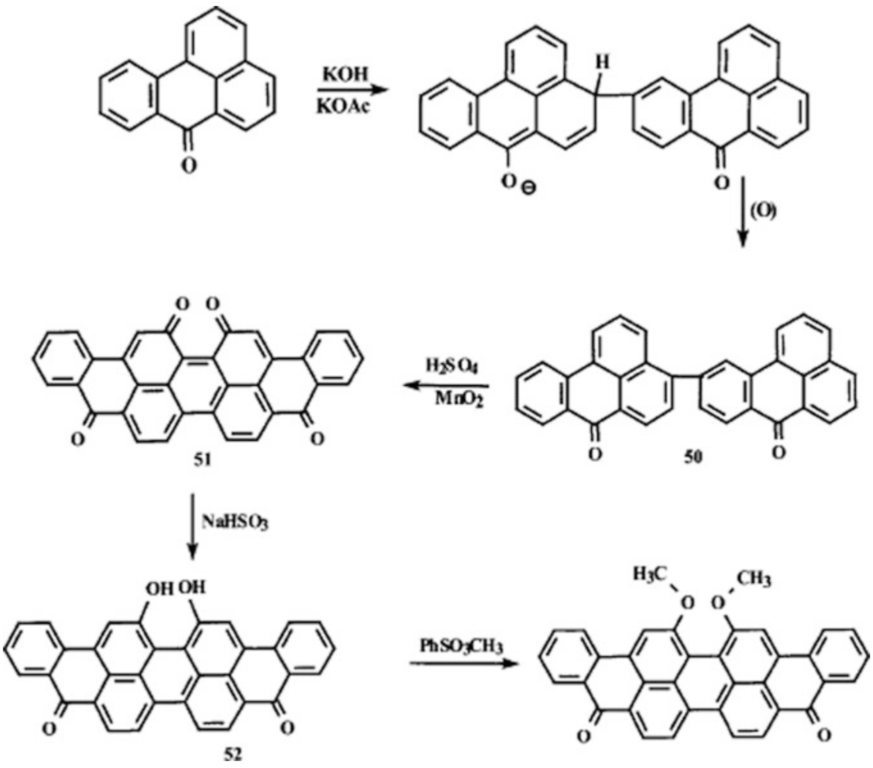
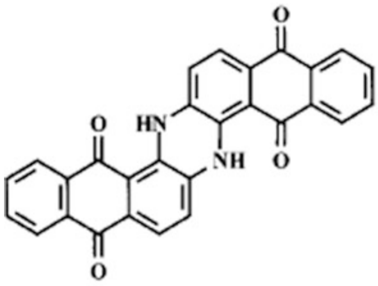
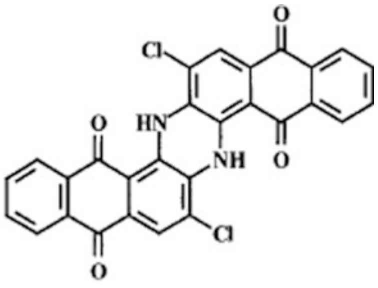


Fig. 13.128 Synthesis of Vat Green 1 from benzanthrone

synthesis is a four-step process: (1) oxidative-coupling of benzanthrone in the presence of alkali to give compound **50**, (2) H_2SO_4-induced ring closure to give compound **51**, (3) reduction to compound **52**, and (4) methylation to give the target dye.



Vat Blue 4



Vat Blue 6

Other important anthraquinone vat dyes belong to the family known as indanthrones. Important examples of this structural type are C.I. Vat Blue 4 and Vat Blue 6. Vat Blue 4 is made by heating 1-amino or 2-aminoanthraquinone at 220–230 °C in a KOH/H_2O mixture. The Vat Blue 6 synthesis is a much longer process that requires the synthesis of 2-chloro-3-aminoanthraquinone [60]. The resultant amine is brominated and converted to the target dye via an Ullmann reaction.

Indigoid and Thioindigoid

By far the most important member of these vat dye families is C.I. Vat Blue 1 (indigo). Its synthesis can be achieved via the four-step method shown in Fig. 13.129. The method

Fig. 13.129 Commercial preparation of synthetic indigo

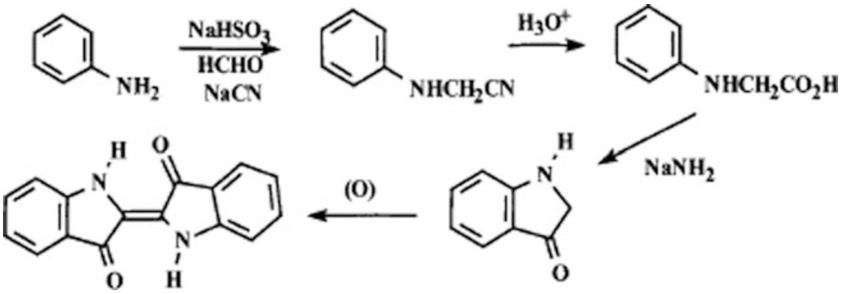


Fig. 13.130 Synthesis of thioindigoid dyes

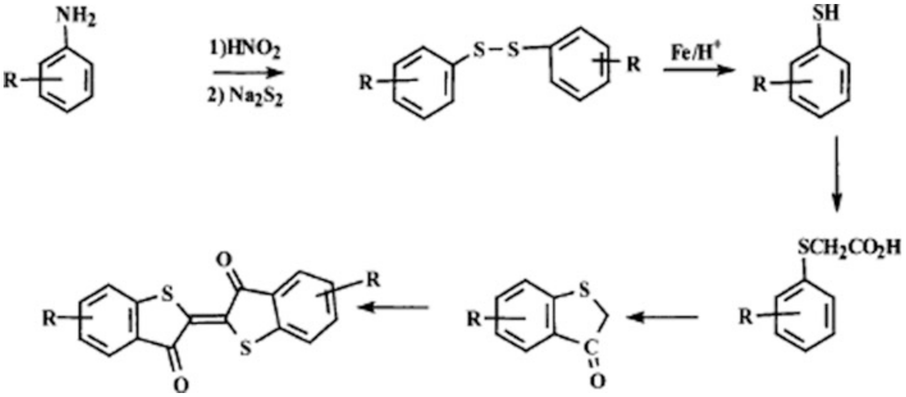
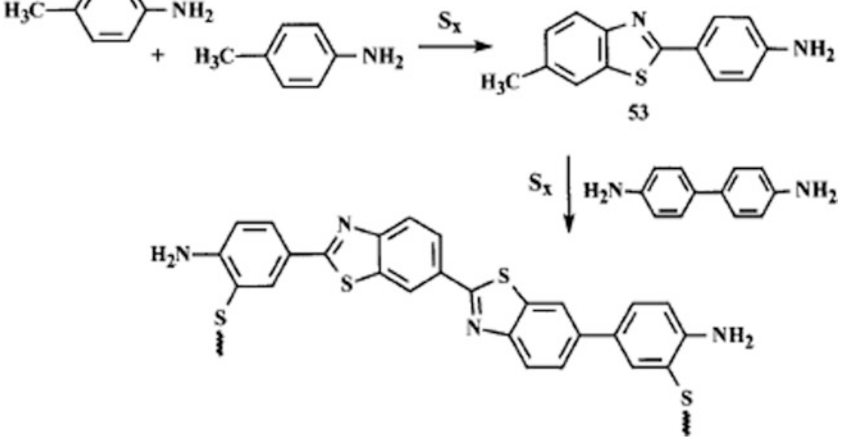


Fig. 13.131 Synthesis of Sulfur Yellow 4 by the sulfur bake process



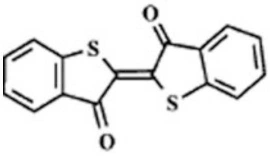
shown is known as the Heumann–Pfleger synthesis [61], where the key intermediate, *N*-carboxymethylaniline, is fused with $NaNH_2$. The cyclic product of the fusion step undergoes air oxidation to give indigo.

Thioindigoids are similarly prepared, in that the synthesis of carboxymethyl intermediates is conducted. The resultant cyclic ketones are much less air sensitive, making oxidation with a chemical agent important. However, this also means that unsymmetrical thioindigoid systems can be synthesized (see Fig. 13.130). Although many have been made, few are in commerce today. Examples are C.I. Vat Red 1 and Vat Red 41.

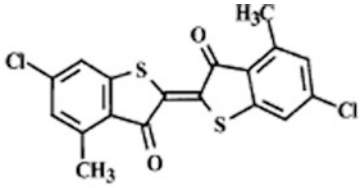
Sulfur Dyes

Earlier we mentioned that sulfur dye chemistry, although quite old, is still much less well defined than for the other classes of dyes. It is clear, however, that many sulfur dyes are produced by the sulfur bake process and that compounds containing the benzothiazole group (e.g. **53**) are formed in route to the final dyes. For instance, the synthesis of C.I. Sulfur Yellow 4 follows a course of the type outlined in Fig. 13.131. In this regard, heating a mixture of *para*-toluidine and sulfur produces a 2-(*para*-aminophenyl) benzothiazole. The sulfur bake process has also been used to make C.I. Sulfur Orange 1, where benzothiazone intermediate **54** is produced along the way [62]. See Fig. 13.132.

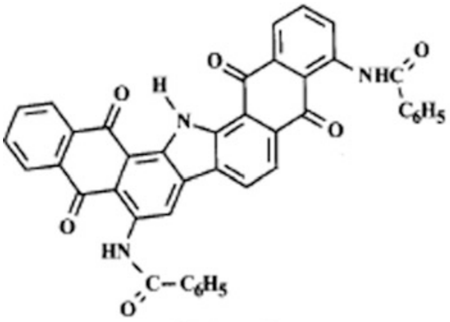
Sulfur blue dyes are often made using an organic solvent such as *n*-butanol, in what is known as the solvent reflux process. Examples are C.I. Sulfur Blue 9 and Sulfur Blue 13. In this case, intermediate structures are indophenols (e.g., **55**). See Fig. 13.133. Similarly, sulfur dyes containing benzothiazine groups can be made from tetrahalogenated benzophenones. See Fig. 13.134.



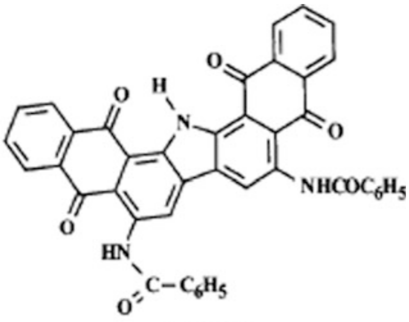
Vat Red 1



Vat Red 41



Vat Brown 3



Vat Black 27

Fig. 13.132 Synthesis of Sulfur Orange 1 by the sulfur bake process

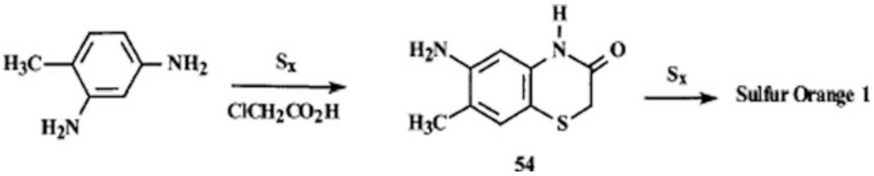
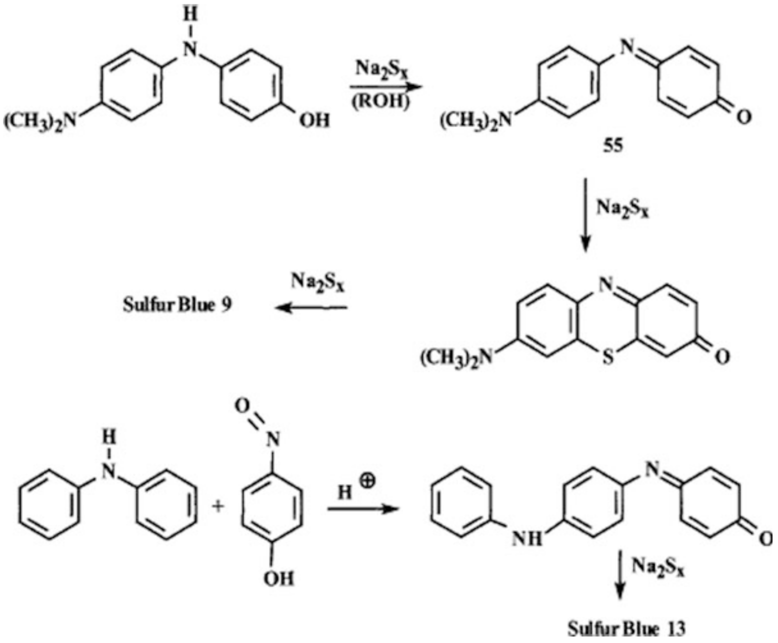


Fig. 13.133 Synthesis of blue dyes by the solvent reflux process



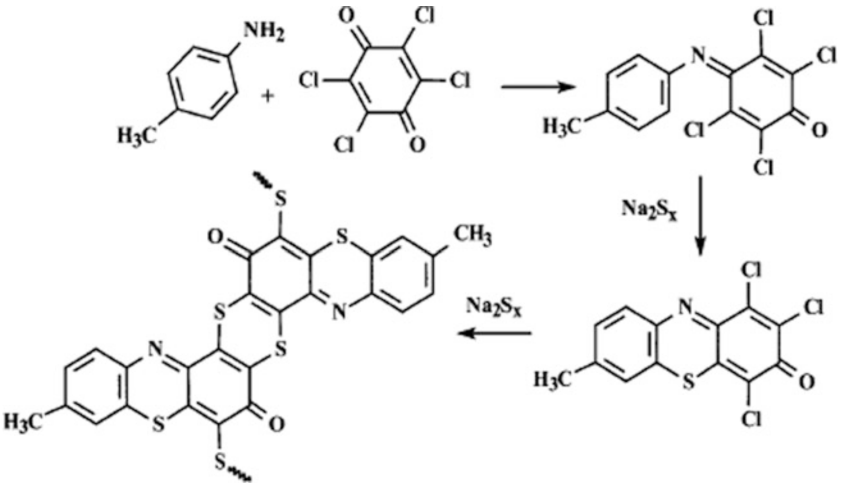


Fig. 13.134 Sulfur dye synthesis from tetrachlorobenzoquinone

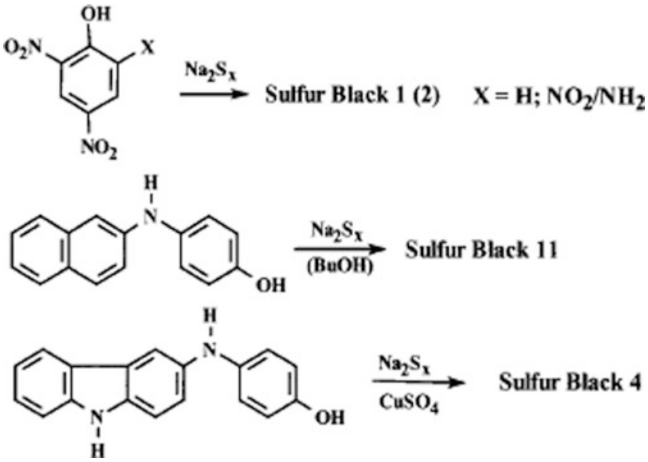


Fig. 13.135 Synthesis of sulfur black dyes

Sulfur black dyes are synthesized according to the methods shown in Fig. 13.135. In these examples sodium polysulfide is the sulfurizing agent employed.

Phthalocyanine Dyes

The synthesis of the CuPc system is achieved as shown in Fig. 13.136. Here it can be seen that any of four precursors can be used. Disulfonation gives C.I. Direct Blue 86 and tetrasulfonation gives C.I. Acid Blue 249.

The chlorosulfonation of the CuPc system opens the door to the synthesis of reactive dyes, as shown in Fig. 13.137. In this case, aminochlorotriazine **56** reacts with a CuPc–SO_2Cl intermediate to give a monochlorotriazine reactive dye (**57**), which in turn can be used to make the cationic reactive dye **58**.

Fluorescent Brighteners (Colorless "Dyes")

Many fluorescent brighteners are derivatives of 4,4′diaminostilbene-2,2′-disulfonic acid (**59**), an example of which is C. I. Fluorescent Brightener 32 (Fig. 13.138). In this case, successive reactions involving diamine **59** with two molecules of cyanuric chloride and two molecules of aniline followed by hydrolysis of the final chloro groups give the target compound.

Structurally related fluorescent brighteners containing a benzotriazole moiety are made according to the route shown in Fig. 13.139. In this case, diamine **59** is tetrazotized, coupled to 2 molecules of 1,6-Cleve's acid, and the intermediate disazo stilbene structure (**60**) is oxidized to C.I. Fluorescent Brightener 40. Nowadays, monosulfonated benzotriazole brighteners are more important [63].

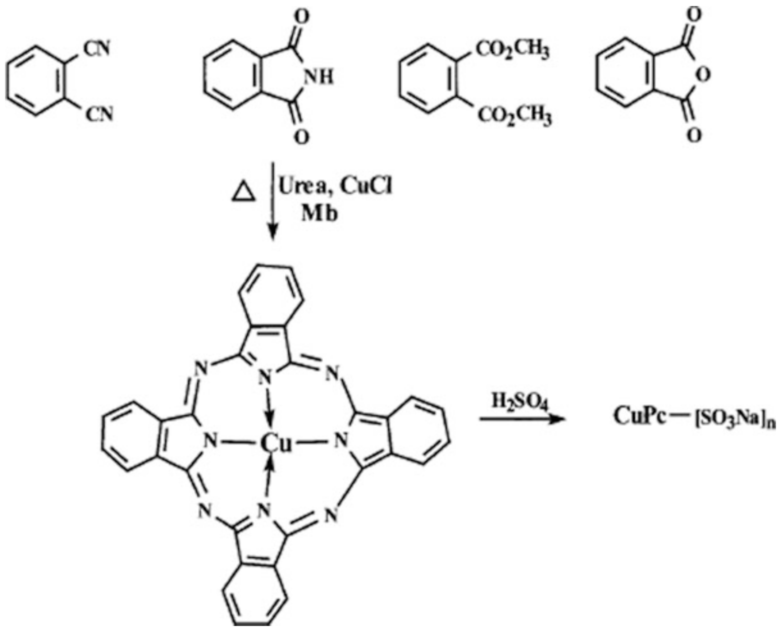


Fig. 13.136 Synthesis of phthalocyanine and sulfonated derivatives

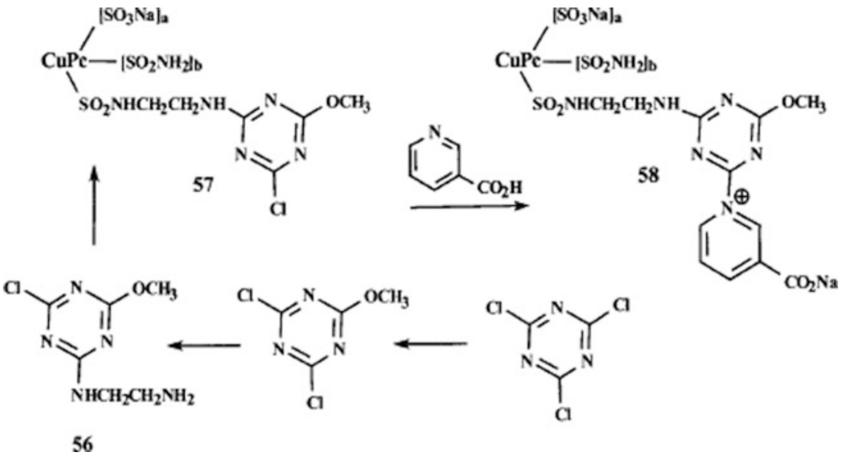


Fig. 13.137 Synthesis of CuPc-based reactive dyes

The synthesis of one example is shown in Fig. 13.140 for C.I. Fluorescent Brightener 46.

Examples of hydrophobic fluorescent brighteners include C.I. Fluorescent Brighteners 199, 130, 236, and 162. The synthesis of these compounds is shown in Figs. 13.141, 13.42, 13.43, and 13.144. In the first of these examples, a *bis*-stilbene structure is made in two steps from *bis*-chloromethyl-xylene, using the traditional reaction of a phosphorus ylide with an aldehyde as the key step in the sequence.

In the second example, the synthesis of a coumarin-type fluorescent brightener is illustrated. Here, *meta*-hydroxy-*N*,*N*-diethylaniline is condensed with ethyl acetoacetate followed by cyclization of the intermediate keto ester **61**. The latter compound undergoes acid-catalyzed cyclization and dehydration to give C.I. Fluorescent Brightener 130. See Fig. 13.142.

A fluorescent brightener containing coumarin and triazole groups is made according to the method shown in Fig. 13.143. The synthesis begins with the preparation of amino-coumarin **62**, which in turn is coupled to Tobias acid with concomitant loss of the SO_3H group and then oxidized to give C.I. Fluorescent Brightener 236.

The final example is for a naphthalimide structure that is made from acenaphthene (**63**) in the four-step sequence shown in Fig. 13.144: (1) sulfonation, (2) chromate oxidation to give the naphthalic anhydride (**64**), (3) condensation with *N*-methylamine, and (4) replacement of the sulfonic acid group in a reaction with methoxide. This process gives C.I. Fluorescent Brightener 162.

Fig. 13.138 Synthesis of fluorescent brightener 32

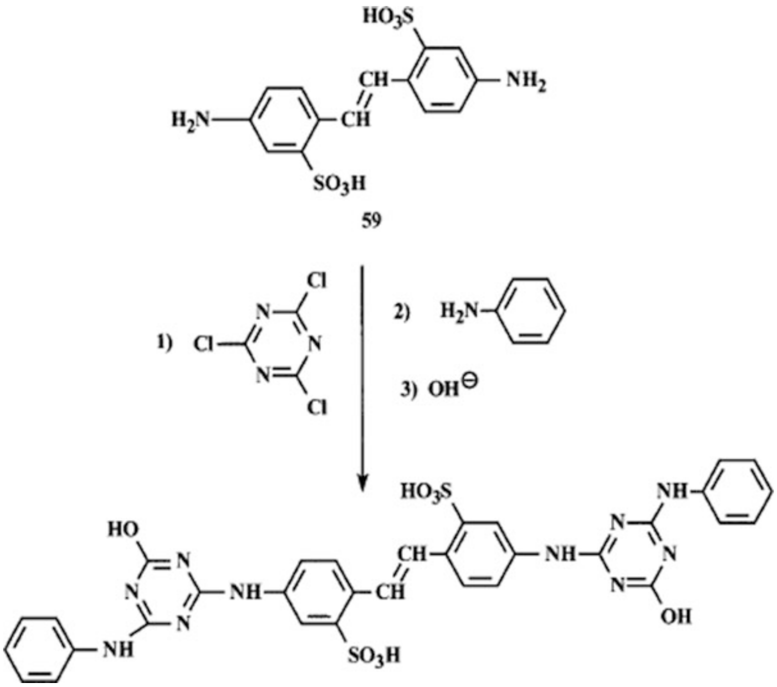
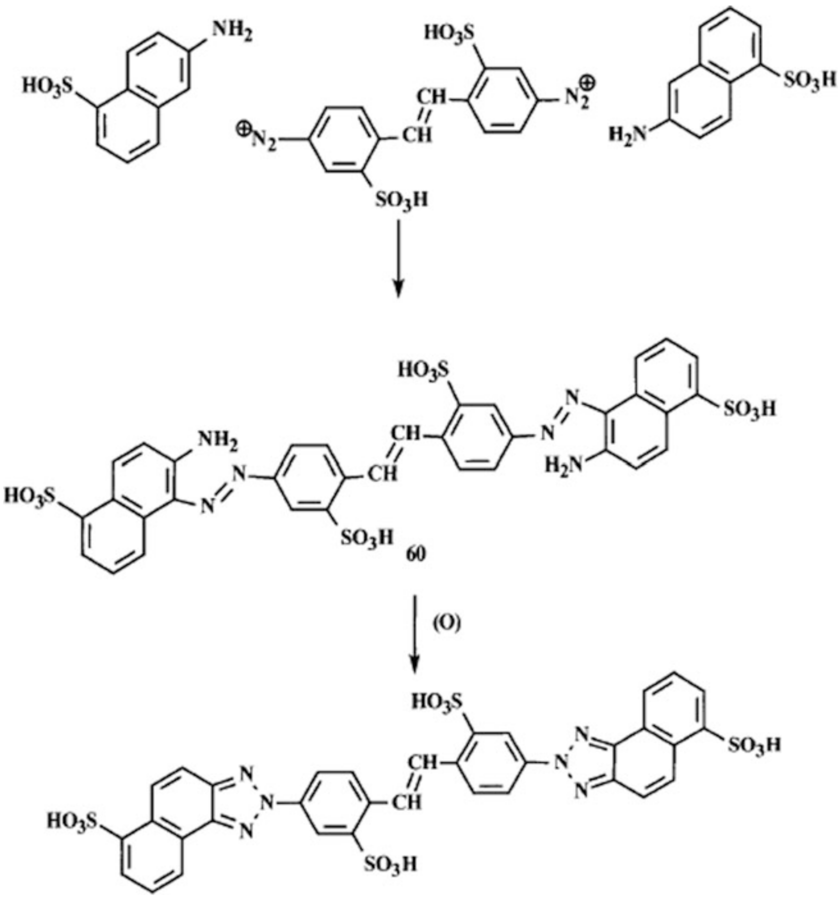


Fig. 13.139 Synthesis of fluorescent brightener 40



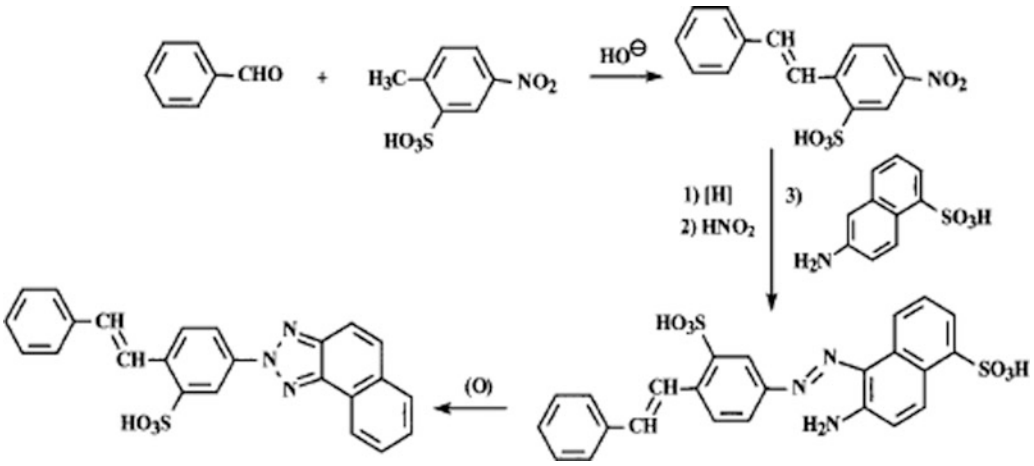


Fig. 13.140 Synthesis of fluorescent brightener 46

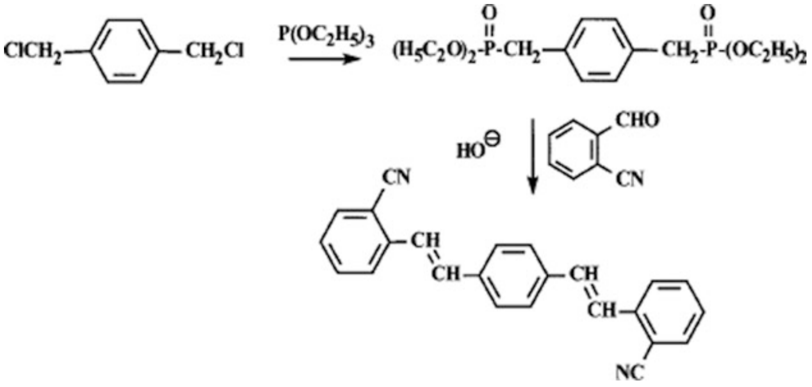


Fig. 13.141 Synthesis of fluorescent brightener 199

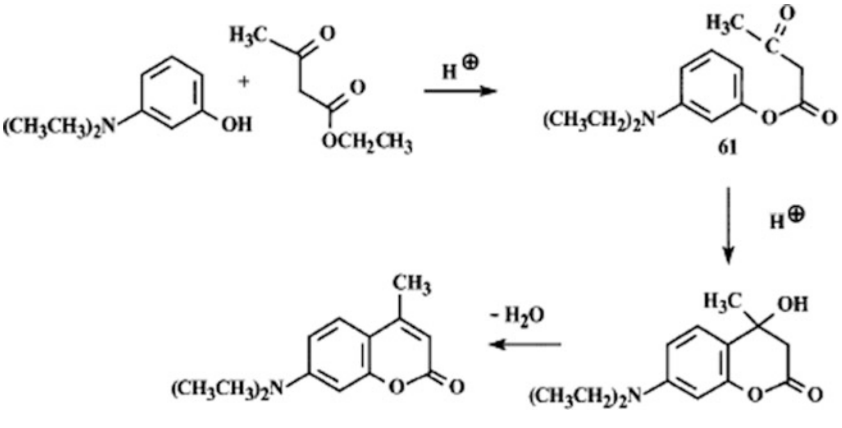


Fig. 13.142 Synthesis of fluorescent brightener 130

Production and Sales

During the 1990s, the large international companies began to form alliances with producers around the world. Hoechst AG, which had done little research on disperse dyes since the 1970s, signed an agreement in 1990 with Mitsubishi of Japan and gained access to a strong line of disperse dyes. BASF AG and Mitsui signed agreements for vat dyes. ATIC resulted from a joint venture between ICI and Atul of India. Finally, a major break came in January 1995, when Bayer AG and Hoechst AG, the parent companies in Germany,

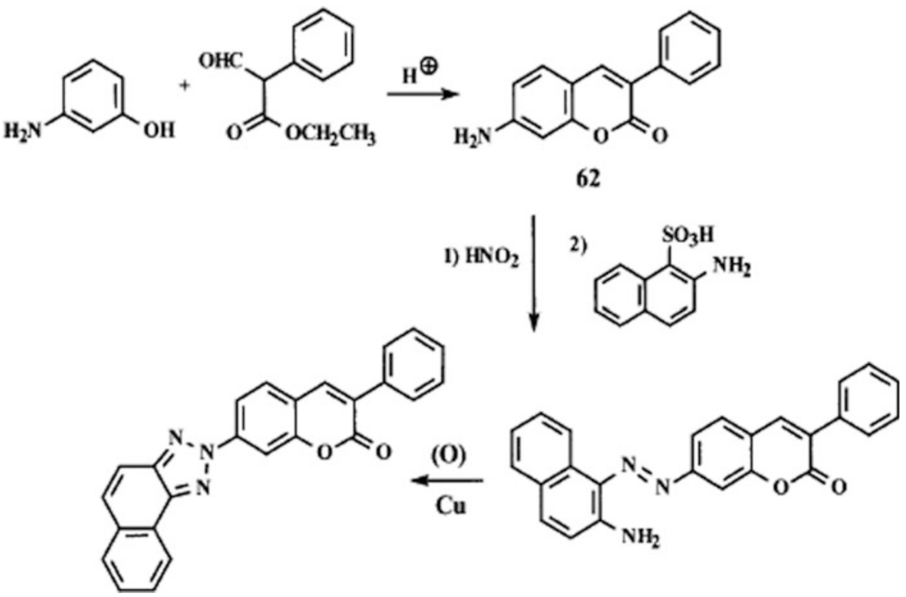


Fig. 13.143 Synthesis of fluorescent brightener 236

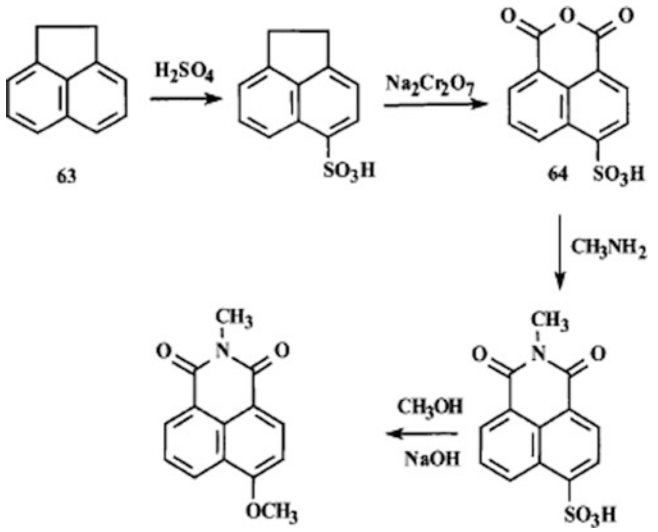


Fig. 13.144 Synthesis of fluorescent brightener 162

announced the formation of DyStar, a worldwide consolidation of their textile dye businesses, which included the US Hoechst Celanese, and Bayer. Within a short time, BASF acquired the textile dyes business of ICI/Zeneca. Swiss companies Ciba and Clariant (derived by consolidating Sandoz and portions of Hoechst in 1995) announced a merger of the textile dyes business but canceled the venture in 1998. Crompton & Knowles (C&K) emerged as the sole US-based major company, but the company struggled during the late 1990s and was sold to Yorkshire Group PLC of the United Kingdom. Yorkshire Pat-Chem and C&K became Yorkshire Americas.

Globalization and establishment of NAFTA meant fewer textile dyes were needed and manufactured in the United States during the late 1990s. The market shrank from 232 million pounds ($955 million) in 1994 to 214 million pounds ($689 million) in 1998 with further cuts expected. Imported dyes expanded but prices fell. Some 1.1 million pounds of disperse dyes were brought in with a value of $5 million in 1992. In 1999, 5.7 million pounds with a value of $10

Table 13.6 World market textile dyes 2002 volume[a]

	Volume tons	%
Reactive	179,381	28
Direct	15,986	3
Vats	18,663	3
Indigo	35,159	6
Sulfur	92,873	15
Disperse	175,845	28
Acid	23,257	4
Metal complex	17,202	3
Chrome	5,519	1
Cationic	18,568	3
Naphthol	8,942	1
Phthalogen	591	0
Pigment prep.	46,885	7
Total	638,871	100

[a] Market Survey DyStar 2002

million were imported. For each class of dyes, you can find expansion of imports for fewer and fewer dollars. The latest year when consumption was publicly revealed is given in Table 13.6.

References

1. Edelstein SM (nd) Historical notes on the wet-processing industry. Dexter Chemical Co., Bronx, NY
2. Zahn J (1967) Bayer Farben Rev 12:32; 11:75
3. Edelstein SM, Borghetty HC (1969) The Plictho of Gioanventura Rosetti. MIT Press, Cambridge, MA
4. Robinson R (1957) The life and work of Sir William Henry Perkin. Proceedings of the Perkin Centennial 1856–1956, AATCC, vol 41
5. Lehner S. America's debt to Perkin. ibid. 269
6. Buyer's guide. AATCC Rev 1(7):17 (2001)
7. Raghavan KSS (1982) Text Dyer Printer 2(20):21
8. Mock GN (2001) AATCC Rev 1(12):18
9. Compiled from: The Colour Index, 5, 4th edn (1992)
10. Crossley ML (1938) Am Dyest Rep 27(3):124
11. Crossley ML (1939) Am Dyest Rep 28(3):487
12. Welham AC (1986) J Soc Dyes Colour 102:126
13. Beffa F, Back G (1984) Rev Prog Color 14:33
14. Allen RLM (1971) Colour chemistry (Chapter 7). Thomas Nelson & Sons, London
15. Hueckel M (1969) Text Chem Color 1(11):510
16. Raue R (1984) Rev Prog Color 14:187
17. Aspland JR (1993) Text Chem Color 25(6):21
18. Allen RLM (1971) Colour chemistry (Chapter 5). Thomas Nelson & Sons, London
19. Shore J (1990) Colorants and auxiliaries–organic chemistry, properties and applications, vol 1 (Colorants). Society of Dyers and Colourists, Manchester, England, p 20
20. Aspland JR (1991) Text Chem Color 21(11):21
21. Shore J (1991) Rev Prog Color 21:23
22. Haley TJ (1975) Clin Toxicol 8:13
23. Prival MJ, Bell SJ, Mitchell VD, Peirl MD, Vaughan VL (1984) Mutat Res 136:33
24. Ecological and Toxicological Association of Dyes and Organic Pigments Manufacturers (1996) ETAD Information Notice No. 6. Text Chem Color 28(6):11
25. Freeman HS, Esancy MK, Esancy JF, Claxton LD (1991) CHEMTECH 21(7):438
26. Dawson JF (1984) Rev Prog Color 14:90
27. Annen O, Egli R, Hasler R, Henzi B, Jakob H, Matzinger P (1987) Rev Prog Color 17:72
28. Weaver MA, Shuttleworth L (1982) Dyes Pigments 3:81
29. Egli R (1991) The chemistry of disperse blue dyes, past and present. In: Peters AT, Freeman HS (eds) Colour chemistry (Chapter 1). Elsevier Applied Science, London
30. Reinert G (1997) Rev Prog Color 27:32
31. Renfrow AHM, Taylor JA (1990) Rev Prog Color 20:1
32. Aspland JR (1992) Text Chem Color 24(5):31
33. Carr K (1995) Reactive dyes, especially bireactive molecules. In: Peters AT, Freeman HS (eds) Modern colorants: syntheses and structures (Chapter 4), vol 3. Blackie Academic and Professional, London
34. Wood WE (1976) Rev Prog Color 7:80
35. Guest RA, Wood WE (1989) Rev Prog Color 19:63
36. Aspland JR (1992) Text Chem Color 24(4):27
37. Baumgarte U (1987) Rev Prog Color 17:29
38. Jones F (1989) Rev Prog Color 19:20
39. Egerton GS (1949) J Soc Dyes Colour 65:764
40. Freeman HS, Peter AT (eds) (2000) Colorants for non-textile applications. Elsevier, Amsterdam
41. Gregory P (1991) Colorants for high technology. In: Peters AT, Freeman HS (eds) Colour chemistry. Elsevier Applied Science, London
42. Freeman HS, Sokolowska J (1999) Rev Prog Color 29:8
43. Bauer W, Ritter J (1993) Tailoring dyes for ink-jet applications. In: Yoshida Z, Shirota Y (eds) Chemistry of functional dyes (Chapter 8.1). Mita Press, Tokyo, p 649
44. Carr K (2000) Dyes for ink jet printing. In: Freeman HS, Peter AT (eds) Colorants for non-textile applications (Chapter 1). Elsevier, Amsterdam
45. Gregory P (1990) Dyes Pigments 13:251
46. Bradbury R (2000) Thermal transfer printing. In: Freeman HS, Peter AT (eds) Colorants for non-textile applications (Chapter 2). Elsevier, Amsterdam
47. Bamfield P (2001) Chromic phenomena: technological applications of colour chemistry (Chapter 1). The Royal Society of Chemistry, Cambridge, UK, p 50
48. Inokuchi H (1993) Organic semiconductors: a still fashionable subject. In: Yoshida Z, Shirota Y (eds) Chemistry of functional dyes (Chapter 7.1). Mita Press, Tokyo, p 521
49. Bamfield P (2001) Chromic phenomena: technological applications of colour chemistry (Chapter 4). The Royal Society of Chemistry, Cambridge, UK, pp 245–256
50. Pavlopoulos T (2000) Laser dyes. In: Freeman HS, Peter AT (eds) Colorants for non-textile applications (Chapter 7). Elsevier, Amsterdam
51. Moura JCVP (2000) Biomedical applications of dyes. In: Freeman HS, Peter AT (eds) Colorants for non-textile applications (Chapter 5). Elsevier, Amsterdam
52. Corbett JG (1985) Rev Prog Color 15:52
53. Corbett JG (1967) J Soc Dyes Colour 83:273
54. Corbett J (2000) Hair dyes. In: Freeman HS, Peter AT (eds) Colorants for non-textile applications (Chapter 10). Elsevier, Amsterdam
55. Waller D, Zbigniew HJ, Filosa M (2000) Dyes used in photography. In: Freeman HS, Peter AT (eds) Colorants for non-textile applications (Chapter 3). Elsevier, Amsterdam
56. Moury DT (1948) Chem Rev 42:213

57. Zollinger H (1991) Color chemistry (Chapter 7), 2nd edn. VCH, Weinheim
58. Fierz-David HE, Blangey L (1979) Fundamental processes in dye chemistry, 5th edn. Interscience, New York, p 250
59. Gordon PF, Gregory P (1983) Organic chemistry in colour (Chapter 2). Springer, Berlin
60. Zollinger H (1991) Color chemistry (Chapter 8), 2nd edn. VCH, Weinheim
61. Püntener A, Schlesinger U (2000) Natural dyes. In: Freeman HS, Peter AT (eds) Colorants for non-textile applications (Chapter 9). Elsevier, Amsterdam
62. Hallas G (1990) Chemistry of anthraquinoid, polycyclic, and miscellaneous colorants. In: Shore J (ed) Colorants and auxiliaries–organic chemistry, properties and applications (Chapter 6), vol 1 (Colorants). Society of Dyers and Colourists, Manchester, England, p 20
63. Siegrist AE, Heffi H, Meyer HR, Schmidt E (1987) Rev Prog Color 17:39

The Chemistry of Structural Adhesives: Epoxy, Urethane, and Acrylic Adhesives

14

Kirk J. Abbey and Dennis J. Zalucha

Introduction

Adhesives have been used successfully in a variety of applications for centuries. Today, adhesives are more important than ever in our daily lives, and their usefulness is increasing rapidly. In the past few decades there have been significant advances in materials and in bonding technology. People now routinely trust their fortunes and their lives to adhesively bonded structures and rarely think about it.

At the same time, the subject of adhesives and adhesion continues to receive much attention by both industrial and academic researchers as evidenced by many measures including the continued growth of membership and attendance at the annual meeting of The Adhesion Society [1]. The scientific literature continues to grow at a rapid pace. New books devoted to general and specialized aspects of adhesion and adhesives continue to appear [2–23]. New patents on adhesive compositions and processes are granted almost daily. Figure 14.1 shows U.S. Patent activity for the past 10 years as discerned by specifically claimed text. There are more patents that mention the use of one of these types of structural adhesives elsewhere in the patent for assembly of some type of article or use other wording with equivalent meaning. Figure 14.2 shows a sample of the most active world patenting organizations irrespective of adhesive chemistry. There are scientific journals devoted to the science and technology of adhesives and their use (Table 14.1). Specialized Internet sites have also appeared in recent years (Table 14.2). A significant portion of the publications dealing with adhesives is concerned with epoxy, urethane, and acrylic structural adhesives as they are used in a wide variety of commercially important applications. This chapter reviews some of the chemistry of these adhesive types as it applies to structural applications, that is, to those applications where the adhesive bond must carry a load while being resistant to dimensional changes, also known as creep.

Adhesion

Materials are generally defined as adhesives by what they do. Almost any organic polymer and even many inorganic materials can function as adhesives in some situations. However, whether they are organic polymers or inorganic, or perhaps even solders or brazing alloys, all adhesives share common traits in performing their functions.

1. An adhesive, by surface attachment only, transfers and distributes mechanical loads among the components of an assembly.
2. At some time in the course of the bond formation, the adhesive must be liquid or behave as a liquid in order to wet the adherends.
3. An adhesive carries some continuous, and often variable, load throughout its life.
4. An adhesive must work with the other components of the assembly to provide a durable product that is resistant to degradation by elements of the environment in which it will be used.

The expectations of the user are extremely important in determining whether an adhesive is “good” or “bad.” Adhesives are judged on the ability of the whole assembly to meet the user’s expectations, which will, in turn, depend on the way the assembly is loaded and tested and on what and where the weakest points of the assembly are located [8]. Adhesion is not an intrinsic property of any polymer but is rather a property of the whole assembly. Structural adhesives are distinguished from nonstructural adhesives by the magnitude of the load that they carry.

K.J. Abbey (✉)
Thomas Lord Research Center, 110 Lord Drive,
Cary, NC 27511, USA
e-mail: kirk_abbey@lord.com

D.J. Zalucha
Midland, MI, USA

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_14,

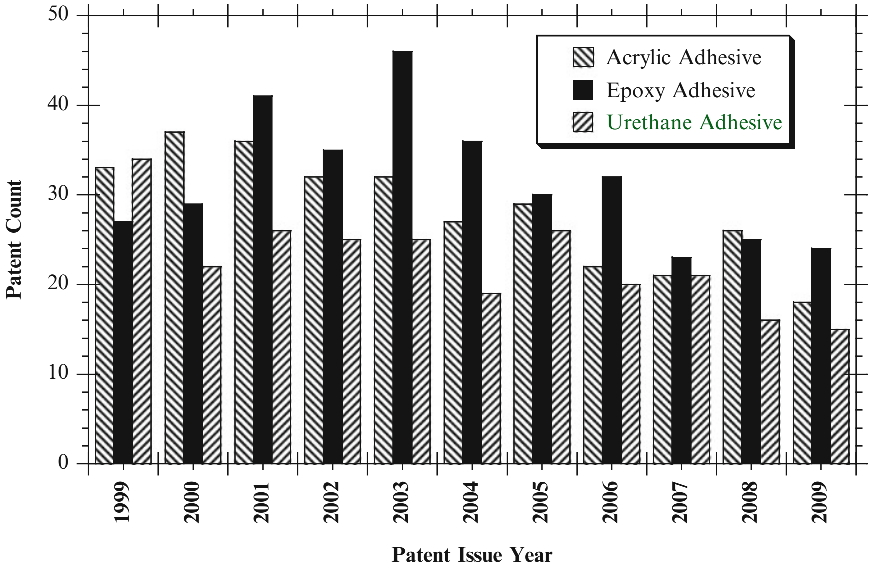


Fig. 14.1 U. S. patents issued from 1999 to 2009. Search terms used in claims were "acrylic adhesive," "epoxy adhesive," and "urethane adhesive" summed with "polyurethane adhesive"

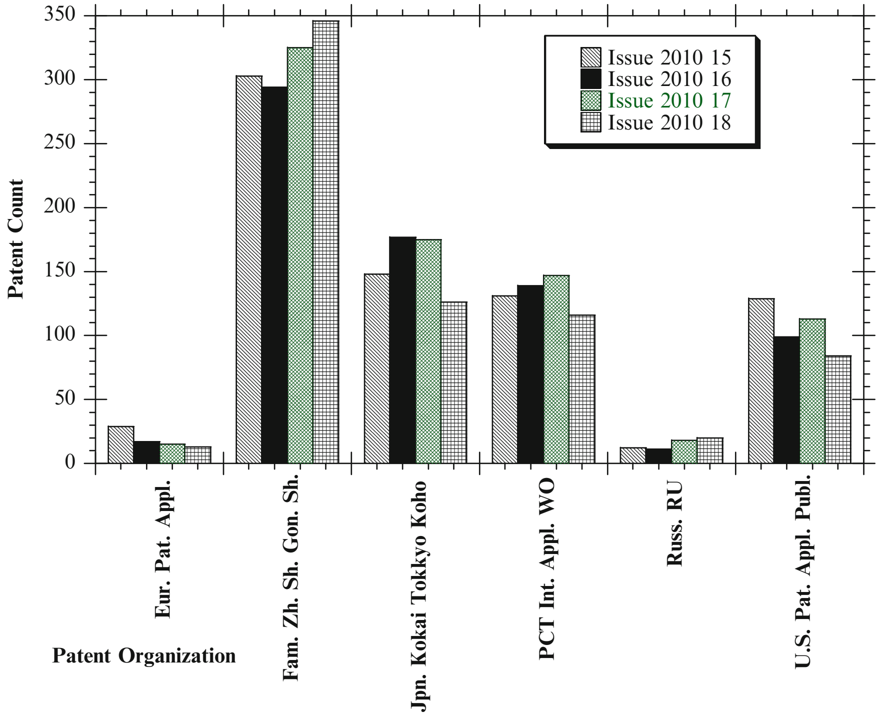


Fig. 14.2 Sampling of citations based on *CA Selects Plus: Adhesives*, for July–August 2010 [24]

Table 14.1 Journals devoted to adhesion and adhesives

Title	Publisher
The Journal of Adhesion	Taylor & Francis, Inc.
Journal of Adhesion Science and Technology	Brill Academic Publishers
International Journal of Adhesion and Adhesives	Elsevier
CA Selects Plus: Adhesives	American Chemical Society
Adhesives & Sealants Industry	Business News Publishing Co.

Curing

The chemistry of a structural adhesive is designed to do at least two important things. First, the adhesive must at some time pass through a fluid state in order to wet the adherends. Second, the adhesive in its final state in the bond line must be a solid, high-molecular-weight polymer that is able to carry and transfer mechanical forces. In almost all cases, the polymer matrix of a structural adhesive will be cross-linked. The chemistry must make possible some manufacturing process for the assembly that allows for the liquid state,

Table 14.2 Internet web sites focused on adhesives

Web address	Publisher
http://www.adhesivesandsealants.com	VertMarkets, Inc.
http://www.adhesiveshq.com	Verdex Group, Thomas Publishing
http://www.specialchem4adhesives.com	SpecialChem, S. A.
http://www.adhesionsociety.org	The Adhesion Society

and there must be some mechanism for passing from the liquid to the solid, load-bearing state. The process of going from the liquid to the solid state is usually termed “curing.”

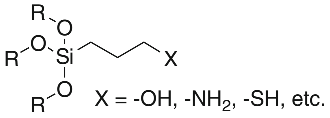
There are three general ways in which adhesives are cured. In the first “hot-melt” method, the adhesive can be applied in a molten state and allowed to cool and solidify in the bond line. In the second method, the adhesive can be applied as a solution or dispersion and the carrier liquid allowed to evaporate, leaving behind the high-molecular-weight polymer. In the third method the adhesive consists of a low-viscosity fluid containing reactive groups that undergo polymerization in the bond line to build the molecular weight sufficiently to carry a load, entailing in most cases forming a cross-linked network. This method is the one most commonly used with epoxy, urethane, and acrylic structural adhesives. The chemistry, methods, and mechanisms for accomplishing this polymerization will be covered in more detail in the following sections.

Adhesion Mechanisms

Focusing on the third cure method, once the liquid adhesive has been applied to the adherends and intimate contact and wetting have been established, the liquid mass is cured by polymerizing it to a solid, high-molecular-weight, load-bearing state. It now transfers load among the components of the assembly. The mechanisms of adhesion can be grouped into three or four categories. Kinloch identifies four categories and devotes an entire chapter to elaborate these in detail [7]. In this review, three mechanisms will be invoked to explain the adhesion of one material to another: (1) mechanical interlocking, (2) electrostatic attraction, and (3) the formation of chemical bonds across the interface. All three mechanisms may play some role in any given bonding situation although often one contributes much more than the others.

Mechanical interlocking is usually invoked when describing the adhesion of ice to glass. Silver amalgam dental fillings are held in place largely by mechanical interlocking. Although there probably are some exceptions, mechanical interlocking usually is not a major factor in bond formation with structural adhesives because, unlike water freezing, most materials contract on cooling or curing. Interdiffusion of polymer chains (i.e., entanglements) may also be considered to be mechanical interlocking at the atomic scale, but it requires mutually compatible and essentially uncross-linked compositions to occur. This is likely to be most important in the welding of plastics.

The other extreme in bonding is the formation of direct covalent chemical links across the interface. These bonds would be expected to be quite strong and durable, but they require special attention not only to the chemistry of the adhesive, but also to that of the substrate. It is necessary that there be mutually reactive chemical groups tightly bound on the adherend surface and in the adhesive. There is evidence that such bonds can be formed under controlled conditions. Silane coupling agents are one example of using specific reactive groups to promote the formation of direct chemical bonds.



By far the dominant adhesion mechanism, particularly in the absence of covalent linkages, is the electrostatic attraction of the polar groups of the adhesive to polar groups of the adherends. These are mainly forces arising from the interaction of permanent dipoles, including the special cases of hydrogen bonding (10–25 kJ/mol) and Lewis acid–base interactions (<80 kJ/mol) [25, 26]. These forces provide much of the attraction between the adhesive and the adherend and also provide a significant portion of the cohesive strength of the adhesive polymer. These interactions are generally classified as attractive long-range forces that drop off inversely to the sixth power of the distance, $1/r^6$, when r is very large compared to the dipole charge separation distance; that is, $r > 1.5$ nm. However, at shorter distances attractive interactions rise more rapidly especially when dipoles are aligned.

In two articles on the cohesive and adhesive strengths of polymers, Mark [27, 28] derives some estimates of what adhesive bond strengths might be achievable with covalent bonds or polar forces across an interface. He concludes that the bonds actually achieved in real life are only a small fraction of what he estimates for the situation in which covalent chemical bonds are the main contributors to adhesion. He further proposes that even if there are a significant number of covalent chemical bonds across the interface, the failing strength of the bond still will depend on the strength of the polar bonds. The polar bonds will fail at a lower strength than the covalent bonds, and the applied load then will be concentrated on the covalent bonds. The measurable mechanical strength of a partially covalent adhesive bond still will be dominated by the polar forces. The implication is that although increasing the proportion of covalent bonds across the interface can enhance durability, the ultimate load-carrying capacity probably will not be significantly affected.

Surfaces

Adhesives must function solely through surface attachment. Therefore, the nature and the condition of the adherend surface are crucial to the formation of strong and durable bonds. By “surface” we usually mean that region of a material which interacts with its surroundings. There is some region of a bonded assembly where the adhesive and the adherend interact, but only rarely is this a sharp boundary. Usually it is a very diffuse, somewhat ill-defined region of interaction that has become known as an interphase rather than an interface. For example, the interaction with a freshly cleaved single crystal of zinc might occur over only a few atomic layers or a few nanometers. Rough or porous surfaces present more surface area than smooth ones of the same dimensions, and the adhesive might reach a depth of several hundred nanometers on a porous adherend such as wood or paper.

One very important aspect of surfaces is that they rarely have the same chemical composition as the bulk material and often seem to be entirely unrelated to the bulk. The surface usually consists of several regions having no clear boundaries. A metal alloy might have a well-defined bulk composition, but at the surface there probably will be a region that is still metallic but is of different chemical composition because of alloying elements or impurities that have segregated at the surface. On top of this region there probably will be a layer of oxides and hydroxides formed by reaction with the atmosphere. There also will be many other contaminants such as nitrogen, sulfur, and halogen compounds formed by interaction with the pollutants in the atmosphere. Finally there will be several layers of adsorbed water. The surface of a metal also might be contaminated with rolling oils, cutting lubricants, drawing compounds, or corrosion inhibitors. Mechanical working of the metal might even mix these contaminants with the other surface materials to create something like an inhomogeneous “frosting” on the surface.

Engineering plastics display some of the same surface phenomena as metals, in that the surface is very different from the bulk. The manufacturing process often introduces anisotropy so that the mechanical properties of the material are different in different directions. In addition, it is common to find that components of the plastic have accumulated at the surface. Low-molecular-weight polymers or oligomers, plasticizers, pigments, mold release agents, shrink control agents, and other processing aids as well as adsorbed contaminants often are present.

The nature of the surface of an engineering plastic can change rapidly in response to its surroundings. The bulk of the material might be in the glassy state, but because of the concentration of low-molecular-weight material and contaminants, the surface region can be quite mobile. Exposure of the surface to a polar environment, such as by wiping with a polar solvent, can cause polar groups in the plastic surface to preferentially orient themselves outward. Exposure to a nonpolar medium can bring out the nonpolar nature of the surface.

The cured adhesive itself can be expected to be inhomogeneous particularly when arising from a formulated composition containing fillers, rubber-toughening agents, and other additives. Low-molecular-weight materials can be drawn into the surface of a porous adherend, leaving higher-weight polymer and fillers behind. The polar or nonpolar nature of the adherend influences the orientation and morphology of adhesive polymers in the interphase. Compounds at the adherend surface can catalyze or inhibit polymerization. Solvents in the adhesive can swell the adherend or dissolve portions of the adherend surface.

The interphase region is complex, and its composition is usually unknown. Primers or surface treatments often are used to improve control of the interphase and provide increased adhesion, durability, and resistance to aggressive environments. The chemistry of primers and surface treatments is as varied as the chemistry of the adhesives, but they will not be considered further in this chapter.

Any bonded construction consists of at least two adherends and one adhesive and contains at least two interphase regions. It is important to remember that the performance of the construction, its durability, its mechanical properties, and its response to tests and challenges, are all properties of the entire assembly. The successful use of adhesives depends on taking account of all parts of the construction and the process. Whereas the adhesive is just one part of the assembly, its chemistry plays an important role in the bonding process.

The following sections discuss the chemistry of some major classes of thermosetting, structural adhesives.

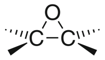
Epoxy Structural Adhesives

Introduction

Epoxy adhesives command a large portion of the structural adhesives market. Many people are probably familiar with epoxy structural adhesives, as these are the typical two-part adhesives found in hardware stores and supermarkets. Epoxy adhesives owe their popularity with both the general public and industry to their ease of use, their relative safety, and their compatibility with many adherends (Fig. 14.3). The various chemical reactions involving the epoxy ring provide a fertile field for the development of a wide range of properties [29]. Even a half century after the first epoxy patents were issued [30], new patents on epoxy adhesive technology continue to appear every month (Fig. 14.1).

Fig. 14.3 Construction of a bicycle frame from carbon fiber composite tubing using an Araldite™ epoxy adhesive (Courtesy of Huntsman Advanced Materials)

Epoxy adhesives get their name from the portion of the adhesive containing 1,2-epoxy, epoxide, or oxirane ring. This three-member ring consists of two carbon atoms joined to an oxygen atom. The highly strained geometry of this moiety with a strain energy of 114 kJ/mol accounts for its reactivity with many nucleophilic or electrophilic compounds.

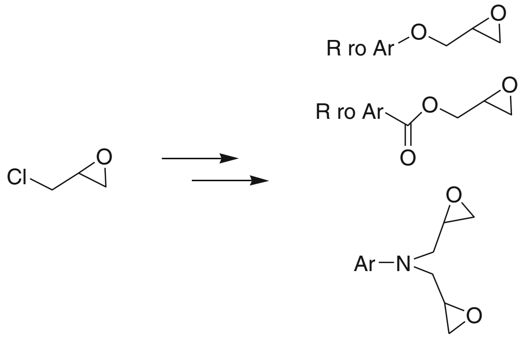


Epoxy Ring

Typical epoxy resins used to formulate epoxy adhesives have at least two epoxy rings, usually at the ends of a relatively short-chain prepolymer. The epoxy groups then are reacted with other epoxy groups in a chain-growth polymerization or with another curative in a step-growth polymerization to produce a polymer network, which can be either thermoplastic or thermoset. The polymer linkages created by reaction of the epoxy ring are polar and provide adhesion to a variety of polar surfaces. With the proper backbone polymers and curing agents, cured epoxy adhesives can be very tough and resistant to chemical degradation. Bonded assemblies can have a high degree of durability and environmental resistance. Epoxy adhesives are widely used in the transportation industry where they can greatly reduce the amount of welding required.

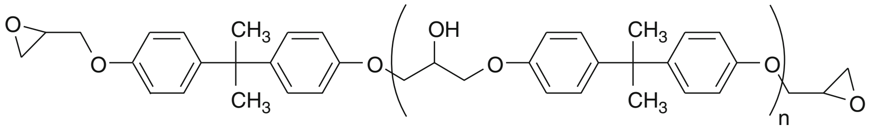
Commercial Epoxy Resins

Epichlorohydrin, or 1-chloro-2,3-epoxypropane, is the key raw material in the manufacture of the most widely used epoxy resins known as glycidyl ethers, amines, and esters. The most commercially important epoxy resins for structural application, for coatings, or for adhesives are the diglycidyl ether of bisphenol A (DGEBA, also referred to as bisphenol A diglycidyl ether, BADGE) and its higher homologues. Pure DGEBA crystallizes after a short time in storage, melting point 42 °C. Because of this, most liquid grades used in epoxy adhesives contain 10–15% higher oligomers, average $n \approx 0.15$–0.20 as well as some side products. Typical commercial resins have molecular weights of about 370 and epoxide equivalent weights in the range of about 180–190, and will have some small percentage of material that is less than difunctional. Various grades of liquid DGEBA resins are manufactured by several companies including Dow Chemical (D.E.R.™) [31], Momentive Specialty Chemicals (formerly, Hexion Specialty Chemicals, Resolution Performance Products, and Shell Chemical, EPON™) [32], Reichhold (Epotuf™) [33], Polystar (ChemRes™) [34], and Huntsman (Araldite™) [35].



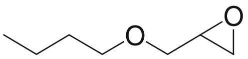
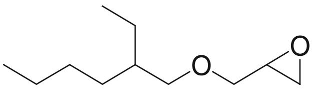
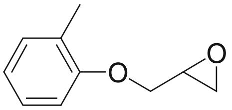
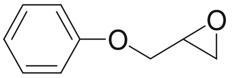
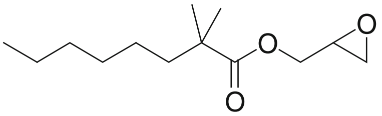
Epichlorohydrin derived epoxy resins

The higher-molecular-weight epoxy resins can be produced commercially either through the direct reaction of bisphenol A, NaOH, and epichlorohydrin or by chain extension of the DGEBA with bisphenol A. For resins with $n > 2$, the number of secondary hydroxyls exceeds the number of oxirane groups. They are often used as the reactive site for coatings. The value of n can be as high as 90 [36]. The higher-weight polymers give greater toughness but with a greater tendency for swelling by solvents. Their high T_g's and higher melt viscosity make them less suitable for most adhesive applications.



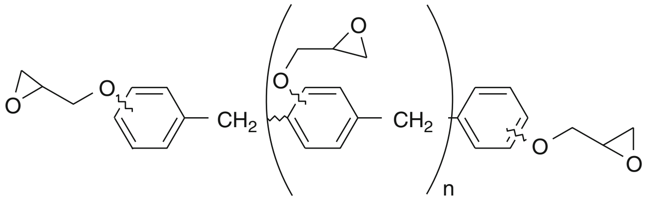
Bisphenol A resin

Table 14.3 Select examples of monofunctional epoxy diluents

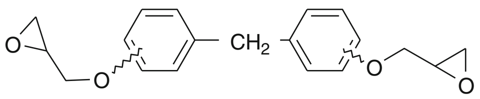
Composition	Nominal structures	Trade name
n-Butyl glycidyl ether		Heloxy Modifier 61[a] 6741[b]
2-Ethylhexyl glycidyl ether		Heloxy Modifier 116[a] 6746[b]
o-Cresyl glycidyl ether		Heloxy Modifier 62[a] 6742[b] Araldite® DY-K[c]
Phenyl glycidyl ether		Heloxy Modifier 63[a] 6740[b]
Glycidyl ester of neodecanoic acid		Cardura™ Glycidyl Ester[a] 6770[b] Glydexx™ N 10[d]

[a] Momentive Specialty Chemicals (formerly Hexion Specialty Chemicals/Resolution Performance Chemicals/Shell Chemicals)
[b] Pacific Epoxy Polymers, Inc.
[c] Huntsman
[d] ExxonMobil Chemical

Another important class of aromatic epoxy resins is the epoxy novolac resins. Novolacs are polymers made by condensing phenol or cresol with formaldehyde in the presence of an acid catalyst. The result is a chain of phenol or cresol groups held together by methylene bridges in a mixture of *ortho* and *para* isomers. The hydroxyl groups then can be reacted with epichlorohydrin and dehydrohalogenated with base to produce an epoxy functional novolac. The simplest, containing only two rings, is the diglycidyl ether of bisphenol F. It is a liquid resin having an epoxy equivalent weight (EEW) of about 165.



Epoxy phenol novalac



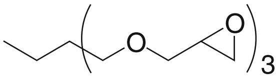
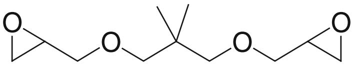
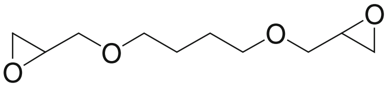
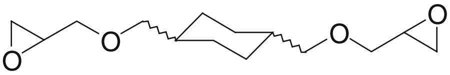
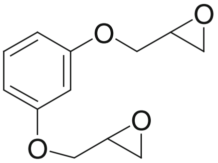
Diglycidyl ether of bisphenol F

Unlike that of the higher molecular weight bisphenol A resins, the EEW of the epoxy novolac resins remains relatively constant with increasing molecular weight because all of the aromatic rings contain hydroxyl groups that can be epoxidized. Higher-molecular-weight epoxy novolac resins can produce adhesives with higher cross-link densities. The value of *n* for epoxy novolac resins used in adhesives is usually in the range of about 0.2–3.5.

Various aliphatic and aromatic monoglycidyl ethers are available that can be used as reactive diluents of many formulated epoxy structural adhesives. They are used to lower the viscosity of the composition without introducing low-molecular-weight unreactive species that might migrate out of the adhesive after it is cured. These compounds are made in much the same way that the polyfunctional epoxy resins are made. Some of the common reactive diluents are given in Table 14.3.

Butyl glycidyl ether has the lowest viscosity and gives the greatest viscosity reduction for the same weight concentration. However, it has a higher vapor pressure than either cresyl glycidyl ether or the higher alkyl glycidyl ethers, and is more likely to cause problems with skin sensitivity and toxicity. The cresyl glycidyl ether is particularly effective at reducing the tendency of liquid epoxy resins to crystallize on storage. For safety reasons the higher-molecular-weight,

Table 14.4 Select polyfunctional diluent resins

Composition	Nominal structures	Trade name
Trimethylolpropane triglycidyl ether		6752[a] Heloxy™ Modifier 48[b]
Neopentylglycol diglycidyl ether		Araldite DY-T[c] 6749[a] Heloxy™ Modifier 68[b]
1,4-Butanediol diglycidyl ether		6749[a] Heloxy™ Modifier 67[b] Araldite® RD-2[c]
Cyclohexane dimethanol diglycidyl ether		6757[a] Heloxy™ Modifier 107[b]
Resorcinol diglycidyl ether		6769[a]

[a]Pacific Epoxy Polymers, Inc.
[b]Momentive Specialty Chemicals (formerly Hexion Specialty Chemicals/Resolution Performance Chemicals/Shell Chemicals)
[c]Huntsman

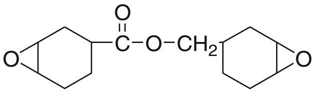
lower-vapor-pressure materials find more use even though they might be less efficient than those with a higher vapor pressure.

Aliphatic and other specialty aromatic polyfunctional glycidyl ethers are also widely used either for viscosity reduction or improving one or more particular properties. Table 14.4 lists examples of some of these.

The diglycidyl ether of 1,4-butanediol is a low-viscosity difunctional epoxy that is a somewhat less efficient diluent than the monoepoxides but does offer the advantage of having two reactive sites. Neopentylglycol diglycidyl ether and cyclohexane dimethanol diglycidyl ether impart greater rigidity than the 1,4-butanediol diglycidyl ether. Resorcinol diglycidyl ether is a very reactive material useful for reducing viscosity and gives higher cross-link densities than bisphenol A resins themselves.

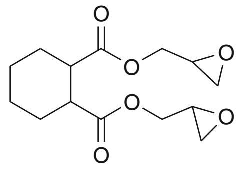
Epoxidized oils such as soybean oil or linseed oil, which are made by oxidizing the internal unsaturation in the fatty acid chain, find much use as plasticizers, particularly in vinyl resins. An internal epoxy ring is much less reactive than a terminal epoxy ring and usually does not take part in typical epoxy curing reactions under mild conditions. These materials can be used as plasticizers in epoxy adhesives but cannot really be considered reactive diluents.

The cycloaliphatic compounds such as 3′,4′-epoxycyclohexylmethyl 3,4-epoxycyclohexane carboxylate (TGY-6110 from Tyger Scientific, Inc. or Uvacure 1500 from Cytec) are made by direct epoxidation of the corresponding cyclohexene with peracetic acid. They produce high T_g polymers too brittle for use as the sole epoxy compound in adhesives. They are marketed towards the ultraviolet cationic cure coatings applications.



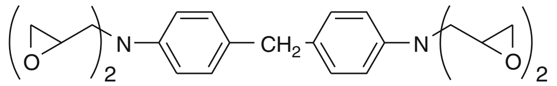
3’,4’ Epoxycyclohexylmethy
3,4-epoxycyclohexane carboxylate

Tougher polymers of somewhat lower T_g are obtained from glycidyl esters such as the diglycidyl ester of hexahydrophthalic acid (Epalloy* 5200, Emerald Performance Materials Co.). The glycidyl esters are prepared by reacting the corresponding cycloaliphatic carboxylic acid with epichlorohydrin and dehydrohalogenating with NaOH.



Diglycidyl ester of hexahydrophthalic acid

There is also a variety of other specialty epoxy resins used in limited quantity for applications demanding some specific performance property. Tri- and tetra-functional materials such as triglycidyl-*p*-aminophenol, triglycidyl isocyanurate, or *N*,*N*,*N*′,*N*′ tetraglycidyl-4,4′-diaminophenyl methane (tetraglycidyl methylene dianiline, Huntsman MY 720) have been used in adhesives requiring high heat resistance and good chemical resistance.



Tetrglycidyl methanedianiline

A typical sales specification for an epoxy resin used in adhesives will include the epoxy equivalent weight (EEW), also sometimes called the weight per epoxy (WPE), viscosity, and density at some specified temperature and the average functionality or number of epoxy groups per molecule. Sometimes a specification on total chlorine is included, which gives some indication of bound chlorine not removed by the NaOH in the dehydrohalogenation process. This bound chlorine is the result of epichlorohydrin side reactions during the epoxy synthesis. A high chlorine content indicates that there will be a large number of molecules of low functionality because each chlorine remaining represents one epoxy ring not formed in the dehydrohalogenation step.

Epoxy Cure Chemistry

Epoxy structural adhesives rely on the chemical reactions of the epoxy group with other reactants to pass from a liquid, wetting state to a solid, load-bearing state. There are a number of ways in which this is done, but all fall under one, or some combination, of three general schemes: step-growth polymerizations through reaction with curing agent, chain-growth polymerizations initiated by Lewis acids, or chain-growth polymerizations initiated by Lewis bases.[1] Often the cure times of the slower step-growth curing adhesives are shortened by including Lewis acid or Lewis base catalysts.

It would be tempting to consider the step-growth and chain-growth polymerization reactions as if they were independent and one could have the choice of either in any particular situation. The truth is that there are aspects of both types of polymerization in the cure of almost every epoxy structural adhesive. Such multiple-cure reactions often make it difficult to calculate the stoichiometry of an epoxy adhesive formulation. One type might predominate, depending on the formulation and cure conditions, but the effects of the other could not be completely discounted. The significance of this statement can be seen by looking at the two generalized reactions.

A typical step-growth reaction is one in which a single epoxy ring reacts with the active hydrogen of the curing agent. The general reaction is:

Model epoxy reaction

In order to form a high-molecular-weight polymer, both the epoxy-containing material and the curative must be at least difunctional. If both behave as difunctional materials, the resulting polymer is linear, and then it is necessary to drive the reaction nearly to completion to obtain a high-enough molecular weight to be useful. Neither of the reactants can contain a significant amount of monofunctional material. Monofunctional impurities or reactive diluents will act as chain terminators and limit the ultimate molecular weight obtainable.

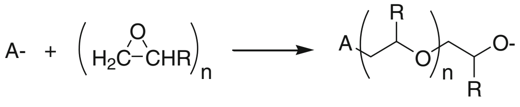
Chain-growth polymerizations occur through the reaction of epoxy rings with the active site on a growing chain and not with each other or with a second curing agent. After being initiated by a Lewis acid or a Lewis base, the growing chain will continue to consume epoxy groups and can reach a high molecular weight very rapidly. Because the epoxy groups are reacting with the growing chain, even a monofunctional epoxy compound can be polymerized. If A* is the initiator and M the monomer unit, the general reaction is:

$$A^* + nM \rightarrow A - M_{n-1}M^*$$

Chain-growth reaction

where the asterisk indicates the active site of the growing chain.

Lewis bases initiate anionic chain-growth polymerizations, the generalized reaction being that of a propagating alkoxide anion.

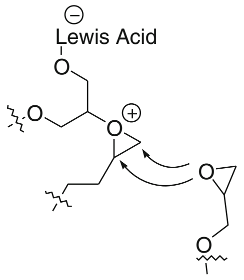


Anionic chain growth

---

[1] Lewis acids and Lewis bases are discussed in "Chain-Growth Polymerizations."

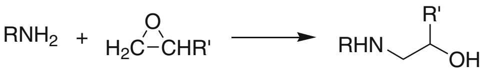
Lewis acids initiate cationic chain-growth polymerizations. There are several possible chain propagation reactions, and the mechanism of cationic chain growth is still open to some debate. However, the propagating species is likely to be an alkylated epoxy cation, or oxonium ion.



In chain-growth polymerizations, epoxy reactants containing more than one epoxy ring per molecule can form tightly cross-linked, three-dimensional networks, as each epoxy group acts as a difunctional reactant.

Step-Growth Polymerization. Only a relative few of the dozens of active hydrogen compounds that undergo reactions with the epoxy ring find widespread use in epoxy structural adhesives. The most common are amines, acid anhydrides, phenols, thiols, and carboxylic acids.

Primary and secondary amines react with epoxy groups to form secondary or tertiary amine linkages. The uncatalyzed reactions proceed at room temperature with the glycidyl ethers and glycidyl esters over many hours with handling strength of >50 psi being achieved in about 2–3 h. The resulting polymer depends on the structure of the reactants and the degree of cure, and thus a variety of final adhesive properties can be had. The reaction is susceptible to general acid catalysis, that is, by hydrogen bonding. Indeed, the reaction is autocatalytic as secondary alcohols are generated during the ring opening reaction. Better hydrogen-bond donors, such as phenols, are commonly used [37–39]. The reaction of the secondary amine with the epoxy group produces a tertiary amine, which can in turn be the Lewis base that initiates an anionic chain-growth polymerization of the remaining epoxy groups, depending on cure conditions.



Epoxy amine reaction

The simple linear aliphatic diamines, $H_2N(CH_2)_nNH_2$, can be used as curatives in adhesives. For small values of n, the short distance between the amine groups can hinder the reaction of the second amine and slow the cure process. Also, the resulting products tend to be brittle for values of n less than about 6.

More flexible, tougher products can be obtained by using liquid diamines or polyamines having more flexible backbones. For example, Jeffamine T-403 (Huntsman) is a low-viscosity liquid (70 mPa/s) having a molecular weight of about 440 and an amine hydrogen equivalent weight of about 81. It is a poly(propylene oxide) triamine made from the polyether initiated by trimethylol propane, and is promoted as a flexible cross-linker for epoxy systems. Their lower reactivity relative to other aliphatic amine may be a consideration in their usage.

Various polyamines also are useful as curatives for epoxy adhesives. An example of a simple polyamine is diethylenetriamine (DETA).

$$H_2NC_2H_4NHC_2H_4NH_2$$

Diethylenetriamine (DETA)

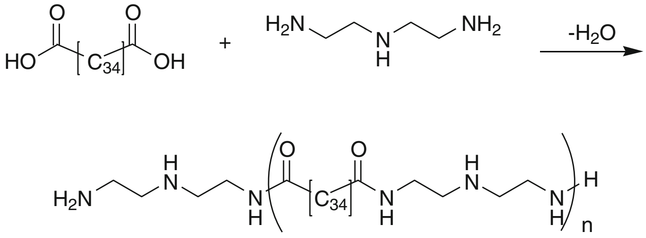
Its higher-weight homologues also are quite useful and are made by adding (CH_2CH_2NH) groups, leading to a homologous series of the form.

$$H_2N-(C_2H_4NH)_n-C_2H_4NH_2$$

where n is two, the material is triethylenetetraamine (TETA), and for $n = 3$ it is tetraethylenepentamine (TEPA), and so on. These three are the most important members of this series for adhesive applications, and are used more often than the simple aliphatic diamines. They often are used in combination with some other curative. Calculating stoichiometry can be difficult with these polyamines. All the active hydrogens might not be available because of steric factors introduced by the first reactions, or because once the first one or two active hydrogens are reacted with epoxy rings, the molecule is anchored into the chain and cannot readily diffuse to other epoxy groups. Therefore, a modest excess of polyfunctional amine is often added to the adhesive formulation. If the amine is used in too much excess, as with any other imbalanced stoichiometry, the final composition may have too many dangling chain ends, that is, be soft and cheesy, or even have unreacted amine that can leach from the product.

Other series of polyamines can be made too. For example, propyl groups can replace some or all of the ethyl groups, or the compound might be modified by reaction with an excess of a monoepoxide to give a hydroxy functional amine. Such modifications are made to improve adhesive properties and sometimes to lower toxicity of the curing agent or make it easier to handle.

Among the most important amine functional epoxy curing agents are the polyamidoamine resins. These are made from dimerized unsaturated fatty acids by reaction of the dimer acid with a polyamine such as DETA.



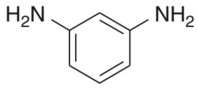
Polyamide Polyamine synthesis

These polyamidoamines are available from several suppliers worldwide (Cognis (Veramide®), Arizona Chemical (Uni-Rez®), Air Products (Ancamide™), and others) and are among the most common curatives in the general-purpose, "do-it-yourself" two-package epoxy adhesives. They have a distinctive odor somewhat like popcorn and are easily recognized in adhesive formulations. The polyamide backbone does contribute to the overall good mechanical properties of the polyamide amine cured adhesives.

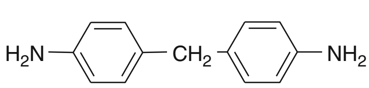
The polyamidoamines are very high-viscosity liquids, some having viscosities over 50,000 mPa/s. Typical amine equivalent weights are 100–150. The polyamidoamines react with bisphenol A epoxy resins at room temperature although the adhesives usually require several hours to reach sufficient molecular weight to carry a load. Cure times can be shortened to a few minutes at about 150 °C.

Epoxy adhesives cured with polyamidoamines are flexible, tough, durable adhesives useful on a wide variety of adherends. They probably have contributed heavily to making the words "epoxy" and "adhesive" equivalent for many people.

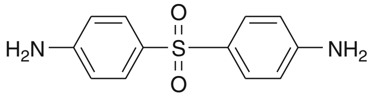
A variety of aromatic polyfunctional amines is also used in curing epoxy adhesives. They generally are slower to react than the aliphatic amines and require a heat cure to be practical. They do provide generally better high-temperature properties than the aliphatic amines. The most commonly used aromatic amine curatives are *m*-phenylenediamine (MPDA, DuPont), methylenedianiline (MDA, Bayer), and diaminodiphenyl sulfone (DDS, Aceto). Albemarle Corporation's Ethacure® 100 is claimed to be a more user-friendly curative that can be used in place of MPDA or MDA.



m-Phenylenediamine



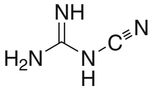
Methylenedianiline



Diaminodiphenyl sulfone

All of the curatives described so far are used to make two-package adhesives in which the curing agent is packaged separately from the epoxy resin. Once they are mixed, they have a limited pot life, usually less than a few hours.

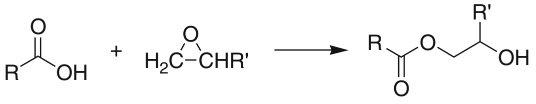
It is possible to make one-package epoxy adhesives that can have very long shelf lives at room temperature but cure rapidly when heated. One amine curative widely used to make single-package heat-cured epoxy adhesives is dicyandiamide (cyanoguanadine), commonly known as dicy.



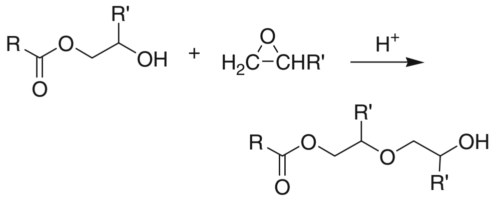
Dicyandiamide

It is made by dimerizing cyanamide in basic aqueous solution, and is a colorless solid melting at 208 °C. Dicyandiamide is soluble in polar solvents, but at room temperature is insoluble in bisphenol A epoxy resins. It can be made into a very fine powder and milled into epoxy resins to form stable dispersions. Because the dicy is insoluble in the epoxy, the only possible reaction sites are at the particle surfaces. Although some reaction certainly occurs over a short time, the adhesives easily can have a useful shelf life of 6 months. On heating to about 150 °C, the dicyandiamide becomes soluble in the epoxy resin, and the adhesive polymerizes rapidly. Cure can be accelerated by incorporation of tertiary aromatic amines or substituted ureas.

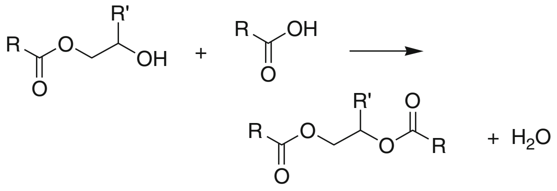
Carboxylic acids can be used to cure epoxy adhesives or otherwise modify epoxy adhesives. The reactions can be complex. If no hydroxyl groups are present initially, the first reaction will be that of the active hydrogen with the epoxy ring to form an ester. This will produce a hydroxyl group on the backbone and allows for competing reactions. The organic acid can catalyze the etherification reaction with the hydroxyl group or undergo a condensation esterification reaction directly with the hydroxyl group.



Epoxy acid reaction



Etherification reaction

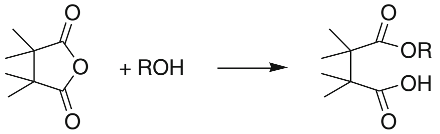


Condensation reaction

These reactions usually are slow at room temperature, and the adhesives must be cured with heat. Tertiary amines or amidines can be used to catalyze the reaction whereby the carboxylate anion is the nucleophile and the ammonium or amidinium ion act as hydrogen-bond donors. Dusek et al. have shown that transesterification only occurs after essentially complete consumption of the carboxylic acid in compositions initially having equivalent amounts of acid and epoxy groups [40, 41]. The transesterification leads to an equilibrium sol–gel composition.

Acid anhydrides also can be used to cure epoxy adhesives although they usually are used only where good service at high temperatures is required. Most of the anhydride-cured epoxy adhesives are cured at high temperature. Because most of the anhydrides are relatively small molecules, the products tend to be tightly cross-linked and can be somewhat brittle.

The first step in the anhydride cure in the absence of a tertiary amine catalyst is ring opening of the anhydride by active hydrogen, perhaps from water or hydroxyl groups already present on the epoxy resin. The resulting acid then reacts as a typical organic acid. In the presence of a tertiary amine, the initial formation of a zwitterionic acylammonium carboxylate salt is possible. The carboxylate anion then reacts with the epoxide as a nucleophile.



Anhydride ring opening by hydroxyl

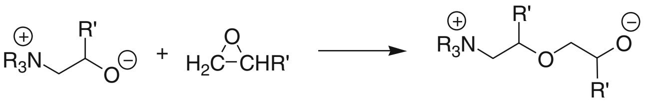
An interesting class of curatives, but with more limited usage, is that of thiols (also known as mercaptans). Thiols react with terminal epoxide groups quite rapidly when a tertiary amine catalyst is present even at temperatures below 0 °C. The reaction is similar to that found with hydroxyl groups and produces a polythioether product. These are the familiar "five-minute" epoxy adhesives and have the characteristic odor of thiols. The rapid cure can be controlled so as to give very good open, or handling, time by the incorporation of very minor amounts of weakly acidic components such as chlorophenol, paraben esters, or carboxylic acids [42]. One limitation on wider application is the commercial availability of suitable polythiols. CapCure™ (Cognis) or GPM™ (Gabriel Performance Products) trifunctional thiols are widely used, but give cured products with low T_g's unsuitable for many structural applications.

Chain-Growth Polymerizations. Chain-growth polymerizations are very important to many commercially successful epoxy structural adhesives. They can be extremely rapid and contribute to the fast cure times needed for high productivity in many manufacturing operations.

A Lewis base is a compound that contains an unshared pair of electrons capable of undergoing chemical reactions. Tertiary amines are examples of Lewis bases, and often are used in epoxy curing agents. In an anionic epoxy polymerization the propagating species is the alkoxide anion generated by the reaction of the Lewis base with an epoxy ring.

Formation of anionic chain initiator

This anion can continue to react with epoxy rings, adding them to the chain until the anion is destroyed in some side reaction.



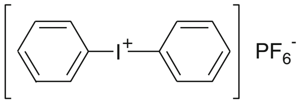
Anionic chain propagation

The product is a polyether, which can be tightly cross-linked when polyfunctional epoxides are used, as each epoxy ring can become part of a different chain.

Lewis acids, compounds with empty orbitals capable of accepting electron pairs, initiate cationic polymerization of epoxy resins. In this case the propagating species is a positive ion. The most commonly used Lewis acids are the boron trihalides, particularly BF_3 and BCl_3. They usually are used in the form of complexes because both are gases at room temperature and are so reactive with epoxy resins that they can be difficult to control. Lewis base complexes with the boron trihalides have much lower reactivity at room temperature but can react quickly on relatively mild heating, depending on the particular complex. Boron trifluoride readily forms complexes with ethers, alcohols, and amines, and several of these complexes are commercially available.

A boron trifluoride ethanolamine complex can be included in the curative portion of a two-package epoxy adhesive. When the two packages are mixed at the time of use, the mixture can have a pot life of hours at room temperature but polymerize in minutes at temperatures of 100–150 °C.

One very interesting new application of Lewis acids in curing epoxy adhesives has appeared within the last 30 years. The Lewis acid initiator for the cationic polymerization is formed by the heat or ultraviolet light-induced decomposition of Lewis acid:Lewis base salts. Several patents by Crivello and coworkers [43–47] describe compounds containing aromatic onium salts such as iodonium or sulfonium, in which the cation is stabilized by the aromatic rings. The field is covered in more detail in a recent review [48].


Diphenyl iodonium salt

The counterion generally is a large stable anion such a $[SbF_6]^-$, $[AsF_6]^-$, $[PF_6]^-$, or, $[BF_4]^-$. The more stable anions are less likely to terminate the growing cationic chains than are typical anions. The salts can be dissociated with heat to release the cation, which appears capable of initiating cationic polymerization of many materials in addition to epoxy rings. If the proper dye sensitizers are added, the cation can be liberated by ultraviolet light to initiate the polymerization. The cations persist for quite some time after the light source is removed. Acids of the form $H^+SbF_5X^-$, where X is a halogen, also have been used to catalyze epoxy reactions [49].

Evolution

The main advances being made on the chemical structure and manufacture of epoxy resins is being driven by microelectronic applications, which are not typically thought of as "structural." For these applications where failure is often induced by mismatched thermal expansion coefficient and fatigue on cycling in temperature, extremely low halide content is essential. The epichlorohydrin route to epoxy resins generally imparts too much residual chlorine content that is not easily removed by subsequent purification steps. A number of alternative routes using hydrogen peroxide have been proposed with various catalysts or coreactants. The patent by Khan et al. [50] proposes one improvement and the patent's background section provides a good review of related processes. The direct oxidation of olefins by molecular oxygen, while limited in scope, can be done to prepare monofunctional epoxies used as reactive diluents [51].

Formulations continue to investigate methods to improve toughness. Several studies of nano-fillers have recently appeared showing improvements with different curing agents and temperatures [52–55]. Nano-dispersed triblock copolymers were also obtained under certain cure temperatures that yielded toughening [56]. Hyperbranched resins terminated with epoxy groups show improvements over linear liquid rubber including lower viscosity at similar toughness [57–60].

Summary

Epoxy structural adhesives have proven to be versatile and reliable compositions. Their widespread use and acceptance is in part due to the varied chemistry of the epoxy ring and the skill of scientists and adhesive formulators in developing high-quality compositions that can produce reliable, reproducible structural joints even when applied by relatively unskilled users. New compositions and chemical reactions continue to be disclosed, and it is certain that the knowledge of epoxy chemistry will continue to grow. Many of the new discoveries will find their way into new epoxy structural adhesives.

Urethane Structural Adhesives

Introduction

The term "urethane adhesive" as it is generally used encompasses a lot of chemistry that is not necessarily urethane chemistry. "Urethane" is the common name for the compound ethyl carbamate. In common usage, "urethane adhesive" generally means an adhesive polymer derived from isocyanate chemistry and reactions of isocyanates with active hydrogen compounds. However, isocyanate reactions do not always lead to urethane linkages, and there are ways of arriving at urethane linkages without involving isocyanates. In this section the common approach is taken; that is, an adhesive that uses reactions of the isocyanate group to bring about polymerization in the bond line is considered a urethane adhesive.

The study of industrial applications of isocyanate chemistry and polymers derived from isocyanates received much attention in Europe, particularly Germany, in the 1930s and during World War II. Patents on aspects of urethane chemistry appeared as early as 1937 [61]. The effort going into understanding isocyanate chemistry and commercializing urethane products continues [62–66].

The isocyanate group consists of a linear arrangement of nitrogen, carbon, and oxygen atoms.

$$-N=C=O$$

Several possible electronic configurations can be drawn, most of which involve a positive charge on the central carbon atom. This partial positive charge on the carbon atom accounts for much of the reactivity of the isocyanate group with nucleophilic groups. Not only does the isocyanate group react with a variety of potential curing agents, but also it is very reactive with many of the adherend surfaces on which urethane adhesives are used.

The same high reactivity that makes the isocyanate a desirable reactant for structural adhesives also renders the unreacted isocyanate more acutely toxic than, for instance, epoxy adhesives. Only a few isocyanates are safe enough and easy enough to handle that they find widespread use in urethane structural adhesives.

Isocyanate Preparation

Organic isocyanates are the major building blocks of urethane structural adhesives. They can be synthesized by a variety of routes, but most of the commercially available isocyanate compounds used in adhesives are made by the reaction of a precursor primary amine or amine salt with phosgene, followed by dehydrohalogenation. The reaction with phosgene usually is carried out at a relatively low temperature of less than 60 °C. The temperature then is raised to 100–200 °C to remove the HCl.

$$\mathrm{RNH_2 + COCl_2} \longrightarrow$$

$$\mathrm{RNHCOCl + HCl} \xrightarrow{\text{heat}} \mathrm{RNCO + 2\ HCl}$$

Isocyanate synthesis by phosgenation

There are many possible side reactions, and the yield depends on the reaction conditions, which usually are specific to the starting materials and desired product. Much work has gone into determining the reaction conditions for manufacturing as clean a product as possible.

The amount of isocyanate in a commercial isocyanate or isocyanate-containing formulation usually will be specified as weight percent isocyanate (as NCO) or as an amine equivalent weight, which is the weight of material containing sufficient isocyanate to react with 1 mol of amine hydrogen.

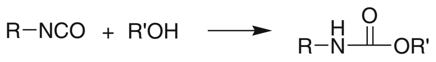
Isocyanate Reactions

Organic isocyanates can undergo a large number of reactions with active hydrogen compounds. One test commonly used to determine the presence of active hydrogen atoms and the number of active hydrogens per molecule is the Tschugaeff–Zerewitinoff analysis or, more commonly, the Zerewitinoff test. An excess of a Grignard reagent, methyl magnesium iodide, is added to the sample to be tested, and the amount of methane evolved in measured.

$$\mathrm{CH_3MgI + RH \rightarrow CH_4 + RMgI}$$

Zerewitinoff reaction

Urethane adhesives take their name from the product of the most common step-growth polymerization reaction used to generate the adhesive polymers. Isocyanates react with hydroxyl groups to create urethane (or carbamate) linkages:



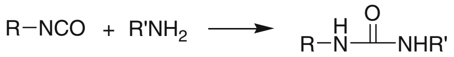
$$\mathrm{R{-}NCO + R'OH} \longrightarrow \mathrm{R{-}\overset{H}{N}{-}\overset{O}{\overset{\|}{C}}{-}OR'}$$

Urethane formation

If both the isocyanate and the hydroxyl-containing material are difunctional, if the mixture is made up to have one isocyanate per hydroxyl, if there are no side reactions, and if the reaction can be driven to completion, a single linear thermoplastic polymer should result. If one or more of the reactants is more than difunctional, it is possible to create an infinite three-dimensional network.

In general, primary hydroxyl groups are faster to react than secondary hydroxyls, which are in turn faster than tertiary hydroxyls, absent catalysts. When reacted with hydroxyl-containing compounds, aliphatic isocyanates tend to be more sluggish than their aromatic counterparts. Urethane linkages made with tertiary hydroxyls tend to be less stable and at high temperature can dissociate into an olefin and an amine with loss of carbon dioxide. Metal compounds, particularly tin compounds such as dibutyl tin dilaurate, and various amines catalyze the isocyanate–hydroxyl reaction.

Isocyanates will react with amines to produce substituted ureas, primary amines being faster than secondary amines.



$$\mathrm{R{-}NCO + R'NH_2} \longrightarrow \mathrm{R{-}\overset{H}{N}{-}\overset{O}{\overset{\|}{C}}{-}NHR'}$$

Urea formation from 1° amine

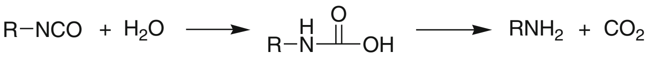
$$\mathrm{R{-}NCO + R'R''NH} \longrightarrow \mathrm{R{-}\overset{H}{N}{-}\overset{O}{\overset{\|}{C}}{-}NR'R''}$$

Urea formation from 2° amine

The ureas are more rigid linkages than the urethane structure, but they also are generally more resistant to heat and chemical degradation. The reaction of isocyanates with amines is generally so rapid that it is nearly impossible to control the reaction well enough to make it useful in formulating urethane adhesives. What usually happens is that the reaction takes place faster than it is possible to mix the adhesive. If it is necessary to include urea linkages in the final production, it can be done by making a prepolymer that has the urea linkages in it already or by taking advantage of the slower reaction of isocyanates with water.

Small (much less than stoichiometric) amounts of diamines sometimes are added to the hydroxyl portion of the adhesive to provide a rapid but limited molecular weight increase as soon as the components are mixed [67, 68]. In this way a two-package adhesive can be made that will flow easily before mixing but will not readily flow after mixing. Such adhesives can be applied to vertical surfaces or overhead, and will remain in place until the bonds are closed and cured.

The reactions of amines and isocyanates are important in adhesives because of the possible reaction of isocyanates with water. Because isocyanates react readily with water, raw materials used in formulating adhesives must be dry, and the compositions must be protected from moisture, including atmospheric humidity, during storage. The first reaction with water is the formation of a carbamic acid, which rapidly loses carbon dioxide to form a primary amine.



$$R{-}NCO + H_2O \longrightarrow R{-}NH{-}C(=O){-}OH \longrightarrow RNH_2 + CO_2$$

Water NCO reaction

The primary amine then can react with another isocyanate to produce a urea. Thus 1 mol of water consumes at least two equivalents of isocyanate, builds molecular weight, and liberates carbon dioxide in the process. If this happens in a closed container such as a drum of adhesive, the result can be explosive, particularly as the reaction mass rises in temperature because of the exothermic reactions. On the other hand, these reactions can be useful in a bond line because under the proper conditions desirable urea linkages can be introduced into the curing adhesive through the reactions of small amounts of water normally present on the adherend surfaces.

The hydrogen atom attached to the nitrogen atom of the urethane group is active enough that it can react with another isocyanate group to produce an allophanate. This is an additional cross-linking mechanism for urethane polymers and can disturb the stoichiometry of the system by consuming an additional isocyanate group for each allophanate formed. Elevated temperatures usually are needed to produce allophanates in uncatalyzed systems; allophanation reactions can be catalyzed by tertiary amines.

$$R''NCO + R{-}NH{-}C(=O){-}OR' \longrightarrow R''{-}NH{-}C(=O){-}N(R){-}C(=O){-}OR'$$

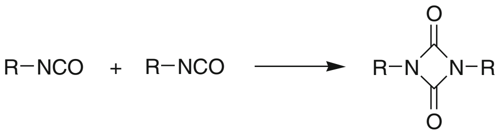
Allophanate formation

In a similar fashion, a urea hydrogen atom can react with an additional isocyanate group to produce a biuret.

$$R''NCO + R{-}NH{-}C(=O){-}NH{-}R' \longrightarrow R''{-}NH{-}C(=O){-}N(R){-}C(=O){-}NH{-}R'$$

Biuret formation

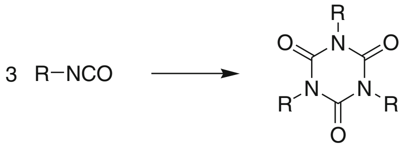
Isocyanate groups also will react with themselves to form a variety of compounds. Two isocyanate groups can react to form a dimer or uretidinedione.



Isocyanate dimerization

These dimers can be dissociated to regenerate the original isocyanates with heat, and some of the dimers, such as the dimer of toluene diisocyanate (TDI), have become commercially important.

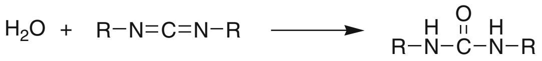
Three isocyanate groups can react to form a trimer or substituted isocyanurate ring. Phosphines or bases such as sodium acetate or sodium formate can catalyze this reaction. The isocyanurate ring is thermally stable, has good chemical resistance, and can enhance the resistance of a urethane adhesive to aggressive environments.



Isocyanates also can react with each other to produce carbodiimides with the loss of carbon dioxide. This reaction requires high temperatures unless catalyzed by specific phosphorus compounds. Formation of carbodiimides normally is not an important cross-linking mechanism in polyurethane adhesives. However, carbodiimides are sold by Angus Chemicals, a division of Dow Chemical (ZOLDINE™ XL-29SE), Nisshinbo Industries (Carbodilite™), and Picassain Polymers, a division of Stahl USA (XL-701). They have been recommended as water scavengers, cross-linkers, and stabilizers for carboxyl functional polyurethanes. The carbodiimide can react with water to give a urea, which still can react with additional isocyanate to produce a biuret.

$$R{-}NCO + R{-}NCO \xrightarrow{\text{catalyst}} R{-}N{=}C{=}N{-}R + CO_2$$

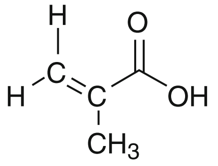
Carbodiimide formation



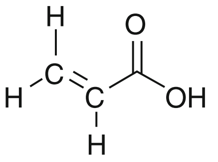
$$H_2O + R{-}N{=}C{=}N{-}R \longrightarrow R{-}NH{-}C(=O){-}NH{-}R$$

Carbodiimide reaction with water

Carbodiimides can react with additional isocyanate groups to form uretone imines, which sometimes are used to modify polyisocyanates used in urethane structural adhesives.



Methacrylic acid



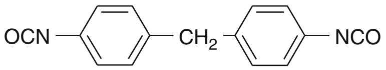
Acrylic acid

Important Isocyanates

Dozens of isocyanate functional compounds have been synthesized, but only a few find much use in urethane structural adhesives. The choices are largely dictated by a combination of performance, price, and safety considerations. Most of the materials used in adhesives are derived from the aromatic isocyanates, toluene diisocyanate (TDI) and 4,4′-diphenylmethane diisocyanate (MDI).

2,4-Toluene diisocyanate 2,6-Toluene diisocyanate

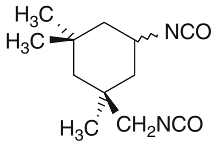
(TDI)



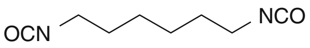
4,4’-Diphenylmethane diisocyanate

4,4-Methylene bis(phenyl isocyanate) (MDI)

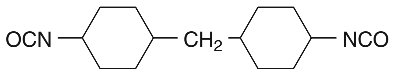
Where color and light stability of the adhesive are important, and cure speed or cost is less important, aliphatic isocyanates are frequently used. Adhesives derived from isophorone diisocyanate (IPDI), hexamethylene diisocyanate (HDI), or 4,4′-dicyclohexylmethane diisocyanate (H$_{12}$MDI) are available.



Isophorone diisocyanate (IPDI)



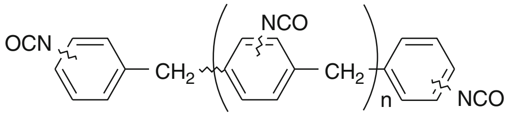
Hexamethylene diisocyanate (HDI)



4,4-Dicyclohexylmethane diisocyanate (H12MDI)

The common commercial TDI is an 80/20 mixture of the 2,4- and 2,6-isomers. The pure 2,6-isomer is available also and is sometimes called TDS. To make TDI, toluene first is nitrated to produce a mixture of the 2,4- and 2,6-dinitro isomers. The dinitrate is reduced to the diamine and reacted with phosgene, which is followed by dehydrohalogenation to give the diisocyanate. Because of its relatively high vapor pressure and toxicity, adhesives rarely contain TDI monomer. Typically, excess TDI is reacted with another material such as trimethylolpropane or a polyester diol or polyether diol to produce a higher-molecular-weight isocyanate functional compound, which is safer and easier to handle than the free TDI. Still, there will be some free TDI present in the adduct, and much work has gone into finding synthesis schemes to minimize the free TDI in prepolymers and adducts [69, 70].

MDI can be considered the first member of a series of polyisocyanates of the general form:

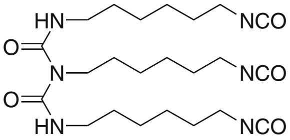


Aromatic polyisocyanates

MDI is the most important member of the series although materials with *n* of 1 or more also are available. The precurser amine is made by condensation of aniline hydrochloride with formaldehyde, followed by reaction with phosgene and dehydrohalogenation. When the aniline is present in excess, the diamine and consequently the diisocyanate are produced in greatest yield, with nearly all of this being the 4,4′-isomer.

MDI is a solid at room temperature with a melting point of about 38°C, which usually is stored and shipped in the molten state for convenience. However, on standing in the liquid state, the MDI slowly dimerizes, and the liquid MDI becomes saturated with dimer at about 1% dimer by weight. The dimer then begins to precipitate. The isocyanate content of the remaining liquid will then remain constant.

HDI can be used as an aliphatic cross-linker in urethane adhesives but presents significant health risks because of its toxicity and high vapor pressure. HDI is more commonly used in the form of its biuret, which is much safer and easier to handle than HDI.



Biuret of hexamethylene diisocyanate

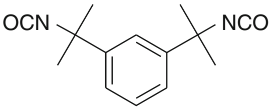
IPDI is another low-viscosity aliphatic diisocyanate that is useful in formulating light-stable polyurethane adhesives, and is somewhat lower in cost than HDI. IPDI has a low vapor pressure at room temperature but, like most low-weight isocyanates, still can present a health hazard. Higher-molecular-weight adducts of IPDI, such as its isocyanurate trimer, are available. The trimer is a high-melting solid (100–115 °C), and one loses the advantage of the low-viscosity liquid in using it but gains in safety and easy handling.

In IPDI, the two isocyanate groups are not equivalent. One is attached directly to the aliphatic ring, and a methylene group separates the other from the ring. Because their environments are different, the reactivity of the two groups are different; and the reactivity can depend on the choice of catalyst. For instance, Ono et al. [71] have shown that when IPDI is reacted with a primary hydroxyl group, the primary NCO is most reactive when the catalyst is 1,4-diazabicyclo [2.2.2]octane. The secondary NCO attached to the ring is the most reactive when dibutyl tin dilaurate is the catalyst.

The saturated analogue of MDI, H_{12}MDI, has found limited use as an aliphatic isocyanate in adhesives. This material is known by a variety of names including Desmodur W™ (Bayer), hydrogenated MDI (or HMDI or H_{12}MDI), reduced MDI (RMDI), and saturated MDI (SMDI). It is a low-viscosity liquid with a fairly high vapor pressure, so it too must be handled with care. In adhesive compositions, the diisocyanate usually is used to make an isocyanate functional prepolymer by reacting excess diisocyanate with a hydroxyl or amine functional polymer such as a polyester diol.

Because it contains two saturated six-member rings, H_{12}MDI can exist in three isomeric forms, in which the orientations of the NCO groups with respect to the rings and each other are different. The *trans–trans* isomer is a solid at room temperature so that over time this isomer can precipitate from the remaining liquid isomers. Samples of H_{12}MDI that have been stored at cool temperatures for a while often have a solid layer of *trans–trans* precipitate in the bottom.

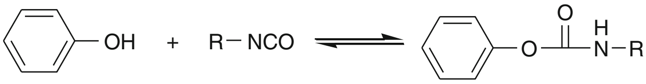
An offering by Cytec Specialty Chemicals, the *meta* isomer of tetramethyl xylene diisocyanate (TMXDI), is interesting because it contains an aromatic ring, but the NCO groups themselves are aliphatic isocyanates and have reaction characteristics typical of aliphatic diisocyanates. It reacts even more sluggishly than the more standard aliphatic isocyanates because of steric interactions, making the reactions easier to control. Compounds such as dimethyl tin dilaurate, lead octoate, or tetrabutyl diacetyl distannoxane have been shown to be effective catalysts for the isocyanate–hydroxyl reaction. The manufacturer claims that it is less toxic than many other isocyanates.



m-TMXDI

Blocked Isocyanates

Blocked isocyanates are compounds formed by the reaction of an organic isocyanate with an active hydrogen compound where the reaction is reversible with moderate heat. The blocked isocyanate can be used in formulating adhesives or other reactive compositions even in the presence of materials that normally would react rapidly with the isocyanate. Phenol is one example of a blocking agent used with isocyanates. There are many other blocking agents in use, including lactams, oximes, and malonates. Even isocyanate dimers such as TDI dimer could be considered blocked isocyanates because they will dissociate with heat to regenerate the isocyanates.



Phenol blocking reaction

There are two major limiting factors to the use of blocked isocyanates in urethane structural adhesives. First, the adherends must be sufficiently heat resistant to withstand the temperatures needed to cause rapid dissociation of the blocked isocyanate. Second, the blocking agent is present to continue competition with the intended curing agent for reaction with the isocyanate. After the adhesive is cured, the blocking agent remains trapped in the bond line and could contribute to poor mechanical properties in the bond, poor resistance to harsh environments, or exude or extract into the environment. Applications on porous substrate, such as wood, or where de-blocking occurs before closure of the bond, such as laminating, are compatible with some of these limitations.

Evolution

Urethane adhesives have also received much attention. New processes are being developed for the phosgene-free

Fig. 14.4 Assembly of an aircraft headliner using an Araldite™ two-part urethane adhesive (Courtesy of Huntsman Advanced Materials)

production of isocyanate-containing raw materials from formamides through reaction with organic carbonates [72, 73]. Methods of stabilizing isocyanate compositions continue to be developed [74, 75]. Formulators continue to produce new mixtures and intermediates from standard materials (Fig. 14.4) [76–78]. New blocking agents for isocyanates are being developed to facilitate the formulation of stable urethane adhesives compositions rapidly curable when heated [79, 80].

Summary

Urethane structural adhesives have proved quite successful in bonding applications that take advantage of their toughness, particularly when the adhesive must function at low temperatures. The reactivity of the isocyanate group may require some extra care in handling and storing the adhesives but the same reactivity provides many opportunities for the polymer chemist and adhesive formulator. The principal reaction used in curing urethane structural adhesives is the reaction of the isocyanate group with primary and secondary hydroxyl groups. Both the isocyanate groups and the hydroxyl groups can be carried on a wide variety of low-molecular-weight oligomers such as polyesters, polyethers, polybutadienes, polyurethanes, and polymer blends so that many of the desired properties of the final adhesive can be built into the polymer before the adhesive is cured.

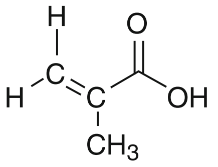
Acrylic Structural Adhesives

Introduction

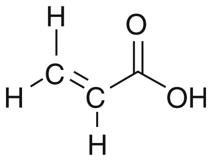
Acrylic structural adhesives are more recently developed products compared to epoxy or urethane adhesives. They first appeared in Europe in the mid-1960s and were commercialized in the United States a few years later [81]. They share many similarities to anaerobic adhesives that preceded them by a few years.[2] The first offerings were rather brittle products of use in limited applications where toughness and flexibility were not critical. However, because these new acrylic structural adhesives polymerized in the bond line through a free-radical chain-growth polymerization mechanism, they did have the big advantage of curing rapidly at room temperature. This polymerization mechanism offered a variety of potential advantages in manufacturing. Bonded structures could be assembled relatively rapidly without the need of curing ovens. Assemblies did not need to be stored for long periods while room-temperature cures were completed. The adhesives could be made at very low viscosities for easy handling and dispensing. In the past 40 years there has been much creative work done with the chemistry of acrylic adhesives so that now they are available in many forms with a wide range of properties.

Acrylic Monomers

All acrylic structural adhesives consist basically of a solution or a mixture of polymers and unsaturated, low-molecular-weight, free-radical-polymerizable monomers with other materials added as needed for the particular intended use. In addition, some precursor of the polymerization initiator will be present. The most common monomers used are methacrylic acid (2-methyl propenoic acid) and its esters or, less commonly, acrylic acid (propenoic acid) and its esters.



Methacrylic acid



Acrylic acid

Several processes for making acrylic acid have been developed [82]. Since the 1980s manufacturers have almost exclusively used a two-step, vapor-phase oxidation of propylene that proceeds through acrolein. The acrylic acid then

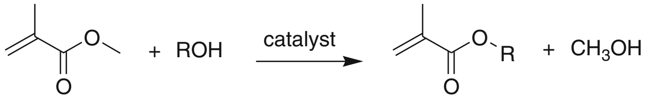
---

[2] See the discussion in [6], pp. 217.

can be esterified with the appropriate alcohols. If current shifts in petroleum prices continue, one of the alternative routes to acrylic acid may become economically preferable.

Unlike acrylic acid, methacrylic acid continues to be produced by a variety of processes including oxidation of ethylene, propylene, or isobutylene [83]. The older commercial process for making methacrylic acid based on propylene proceeded through acetone cyanohydrin. This process is steadily declining in usage because of high waste costs. Rapid development of specialty catalysts in the 1980s by several Japanese companies has moved isobutylene-based production to the forefront. Also, for companies with internal synthesis gas (syngas) production, routes based on ethylene are economically viable.

Direct esterification of methacrylic acid with alcohols using sulfuric acid or other catalysts can be used to prepare methyl methacrylate (MMA) and other esters. Commercial routes for the direct preparation of MMA and some lower alkyl esters also exist. In the 1990s, researchers at Shell developed a direct route to MMA from propyne (methylacetylene), carbon monoxide, and methanol using a Pd(II) catalyst. The limited availability of propyne may slow the expansion of this highly efficient route to high purity MMA. Transesterification of MMA is often the preferred route for the preparation of other esters.



Transesterification of MMA to other methacrylate esters

The choice of monomers that are useful in acrylic structural adhesives is rather limited. Cost always is an important factor, and because acrylic structural adhesives consume only a very small portion of the world's output of acrylic monomer, the formulator usually must rely on acrylic monomers that are made in large quantity for other uses. In addition, the monomers must polymerize readily at room temperature. If a mixture of monomers is to be used, the monomers must copolymerize easily. Finally, the monomers must be good solvents or dispersants for the polymers used in formulating the adhesive. The acrylic monomers finding most use in acrylic structural adhesives are methyl methacrylate and tetrahydrofurfuryl methacrylate. The later, albeit more expensive, has a much higher flashpoint and a generally perceived low odor, yet maintains good solvency.

Curing

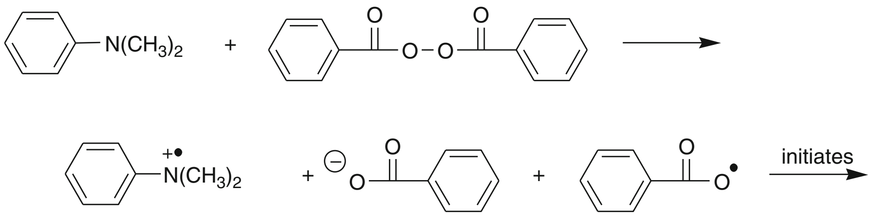
The curing reaction of an acrylic structural adhesive is the chain-growth polymerization of the acrylic monomer. The monomer units are not reactive with each other but react only with a growing chain having an active site on one end. In order to begin a chain, one must generate an initiator in the monomer solution. In the case of acrylic structural adhesives, this initiator nearly always is a free radical, a species having an unpaired, reactive electron. After chains have been initiated, there are three general types of reaction that can occur: propagation, chain transfer, and chain termination. If "*" represents the active site, "A*" represents the initiator, and M is a monomer unit, then the four reaction types are:

Initiation	$A^* + M \rightarrow AM^*$
Propagation	$AM^* + nM \rightarrow AMnM^*$
Chain transfer	$AMnM^* + M \rightarrow AMnM + M^*$
Termination	$RM^* + R'M^* \rightarrow$ RMMR' (combination)
	$RM^* + R'M^* \rightarrow RMC = CH_2 + RMCHCH_3$ (disproportionation)

In addition, a growing chain might undergo chain transfer or be terminated by reaction with a variety of unknown impurities invariably present in any mixture.

The initiating radical is usually created in a redox reaction. Common reactions involve the reduction of an organic peroxide by some reducing agent such as an amine or an ion capable of undergoing a one-electron transfer reaction. Anaerobic adhesives generally rely on metal ions derived from the surface to be bonded as part of the redox system. Acrylic adhesives generally embody the reducing agent in the monomer mixture and place the oxidizer in a monomer free package.

One example of an efficient free-radical initiator generating reaction is the reduction of diacyl peroxides such as benzoyl peroxide (BPO) by tertiary aromatic amines such as *N*,*N*-dimethyl aniline (DMA).



Benzoyl peroxide reaction with N,N-dimethyl aniline

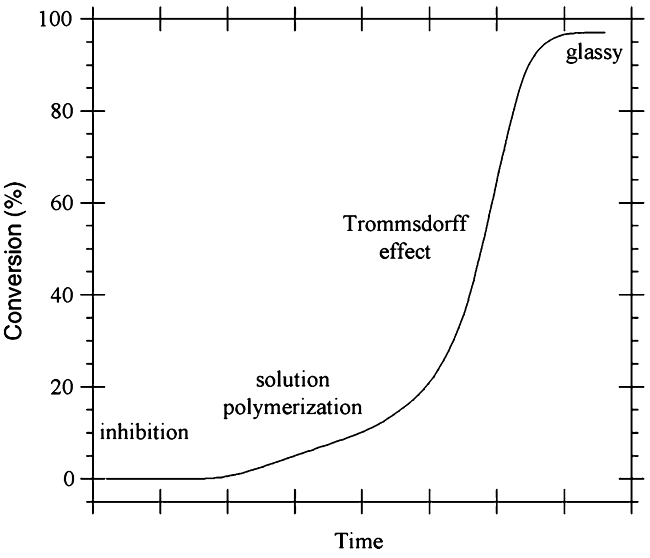
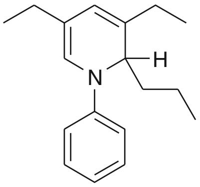


Fig. 14.5 Four stages of free-radical polymerization in bulk

The condensation products of amines and aldehydes have often been used as the reductant. The most common commercial example is that arising from aniline and butyraldehyde for which the active ingredient has been identified as 3,5-diethyl-1,2-dihydro-1-phenyl-2-propylpyridine (DHP). A high purity grade, >85% active, has become available in recent years (Vertellus Specialties Inc., formerly Reilly Industries, Reilly PDHP™ Adhesive Accelerator; R. T. Vanderbilt Company, Vanax™ 808 HP).



DHP

The free-radical polymerization of methacrylate adhesives may show four stages, Fig. 14.5: inhibition, solution polymerization, “gel” polymerization, and glassy polymerization. All commercial monomers have inhibitors added to help prevent premature polymerization during storage. During the inhibition stage, the redox system generates radicals and some minor addition to monomer may occur, but the inhibitors and any dissolved oxygen, a very potent inhibitor of free-radical polymerization, will prevent the formation of polymer and any significant consumption of monomer. Once the oxygen and inhibitors are consumed, very high-molecular-weight polymer will form, $0.1\text{–}1 \times 10^6$, if no efficient chain transfer agents are present. As the reaction progresses, the polymer–monomer solution becomes increasingly viscous and the termination process is retarded. At about 25–30% monomer consumption, the polymer entanglements (gel) become so profound that a rapid rise in observed rate of polymerization is noted as the rate of termination reactions plummets. This rapid acceleration is referred to as the Trommsdorff effect. Because the diffusion of monomer molecules is not appreciably retarded by the “gel,” there is no significant reduction in the propagation rate. Eventually, the polymer–monomer solution will approach the glassy state and even small molecules such as monomer and the redox components cannot move easily. For all practical purposes, the reaction comes to a halt soon thereafter as the smaller, more rapidly diffusing oxygen molecules reenter the bondline and quench the remaining polymer radicals.

Another consequence of oxygen inhibition of free-radical reactions is that the adhesive surface exposed to air will contain a higher concentration of lower-molecular-weight species than the bulk of the adhesive. This can result in the surface being soft, tacky, or even remaining fluid.

Formulation

Although the free-radical chain-growth polymerization brings processing advantages, it also brings limitations. The product of the polymerization of the monofunctional acrylic monomers is a linear thermoplastic polymer. In the case of a copolymer that is largely MMA, the glass transition temperature of the polymer will be well above room temperature if the polymer has been driven to a high-enough molecular weight. The resulting adhesive is brittle, and although its shear strength might be quite high, its ability to withstand peeling forces or impacts will be low. One could use monomer mixtures yielding copolymers having lower glass transition temperatures and expect that the adhesive might be somewhat tougher. However, this parameter is difficult to control because when curing the acrylic adhesive in the bond line, one really is doing a bulk polymerization under uncontrolled conditions and trying to produce a specific polymer with a specific set of material and performance properties. The approach taken by most adhesive manufacturers or formulators is to use “prepolymers,” high-molecular-weight oligomers that can be made under controlled conditions and then dissolved or dispersed in the acrylic monomer. By forming key elements of the polymer chains under controlled conditions, many of the desired properties of the final product can be built into the adhesive when it is formulated and before the material is cured. The prepolymer can cause the solution polymerization stage of polymerization to be skipped, thereby decreasing the cure time substantially.

The simplest approach to improving toughness is to dissolve or disperse a nonreactive rubbery polymer in the monomer mixture that is to be polymerized. Examples of such materials are nitrile rubbers, polyethers, and acrylic rubbers. The choice will depend on cost, desired properties, and the solubility of the polymer in the monomer. As the monomer polymerizes, it will lose its solvating ability for the dissolved polymer. Eventually, when enough of the monomer has been consumed, it can no longer hold the polymer in solution, and the polymer precipitates. The polymer then can segregate as a totally separate phase or as distinct domains heterogeneously dispersed throughout the acrylic polymer matrix, or it can be homogeneously trapped in the acrylic matrix, depending on the rates of reaction, the compatibility with the acrylic polymer, and the relative mobility of the dissolved polymer.

One way of influencing the way in which the added oligomer is distributed in the final cured produce is to provide reactive sites on the oligomer so that it can be incorporated into the acrylic matrix as it forms. For instance, if the added oligomer contains terminal, active, acrylic, or methacrylic unsaturation, it can be easily incorporated into the growing acrylic polymer chains as the adhesive cures. Then it is unable to precipitate as a separate phase and must remain more or less uniformly distributed throughout the matrix. On the other hand, incompatible polymers can be added to deliberately encourage the formation of reinforcing domains. If the oligomer has two or more functional groups capable of participating in the polymerization, there is at least the possibility of forming cross-links in the polyacrylate matrix. Particular advantage has been claimed for combinations of reactive and nonreactive rubbers with attention to their molecular weights [84].

A technology developed at du Pont [85] combines the use of reactive sites on the oligomers with the initiation reaction. The resulting family of acrylic structural adhesives has become popularly known as "second generation acrylics." They consist essentially of solutions of chlorosulfonated polyethylene (formerly made by du Pont as Hypalon, but discontinued in 2010; available from Tosoh Chemical as Toso-CSM™) in acrylic or methacrylic monomers. The chlorosulfonyl groups present on the polymer will react with amine aldehyde condensation products, DHP above, to generate free-radical initiators, at least some of which are claimed to be on the oligomer backbone. The speculation is that the adhesive polymer chains then grow by graft polymerization from various sites on the rubber oligomer. Peroxides and metal ions can be added to accelerate the initiation and polymerization. One of the significant advantages of this type of acrylic structural adhesive is the ability to bond oily or dirty metals with only minimal surface preparation, although some surface preparation usually is necessary to improve bond durability (Fig. 14.6).

Other acrylic structural adhesives can be used on poorly prepared metal surfaces. Inclusion of reactive monomers having phosphate groups can lead to significant improvements in primary bonds to many unprepared metals [86]. Kowa American supplies one such monomer, 2-hydroxyethylmethacrylate acid phosphate.

Fig. 14.6 Applying Fusor™ adhesive for repair of a steel automobile hood (Courtesy of Lord Corporation)

Many other modifications to acrylic adhesives are useful. Improved heat resistance of acrylic adhesives can be achieved by incorporating small amounts of epoxy resin into the formulations [87], or by including cyanate esters and organometallic compounds to provide an additional cross-linking mechanism [88]. Lower sensitivity of oxygen can be achieved by incorporation of a wax that migrates to the monomer–air interface providing a diffusion barrier or by using more efficient initiators [86, 89]. Innovations in acrylic adhesive have led to new initiator reactions that provide rapid initiation and cure while at the same time allowing for improved shelf life [90]. Acrylic adhesives cured by ultraviolet light have been made from mixtures of acrylic functional urethane prepolymers mixed with free radically polymerizable monomers and photo initiators [91–93]. Refined formulations with significantly improved peel strength have recently been introduced [94].

Organoborane initiators for acrylic adhesives have received much attention arising from the observation that they give exceptional adhesion to low energy surfaces, such as polyethylene and polypropylene without surface preparation [95]. Earlier use of tributylborane–amine complexes in MMA and activated with isocyanates, acid chlorides, or sulfonyl chlorides for a dental resin was reported in 1969 [96]. In the mid-1990s, 3M became active in the area of amine-blocked organoboranes with a focus on polyolefin bonding [97]. A commercial product appeared in the mid-1990s that was later twice improved, Scotch-Weld™ DP-8010 [98]. By 2005, more than 30 patents had since appeared covering particular blocking groups, de-blocking agents, and additives to improve adhesive performance. Later Dow Chemical made a push into this field with more than 16 patents on adhesive composition as well as others in specific automotive applications [99]. It has been commercialized under the LESA trademark [100]. The amine-blocked organoborane adhesives generally require refrigeration for good shelf live. A family of more robust, internally coordinated organoboranes has been disclosed [101, 102].

Summary

Acrylic structural adhesives have the advantage of easy handling and processing and rapid cure at room temperature. The ability of some compositions to adhere to unprepared metal surfaces or low surface energy substrates can allow substantial cost savings in manufacturing processes. The brittleness of the polymers made from the monomers often can be overcome by formulating to include tough, reactive oligomers in the liquid adhesive. The variety of possible initiators and mechanisms for generating them has brought many innovative compositions to market.

Hybrid Adhesives

Adhesives usually are classified as epoxy, urethane, or acrylic, based upon the chemistry used to bring about the polymerization of the liquid adhesive. However, the distinction among the various types of adhesives is not always clear, and many hybrid adhesives have been developed and marketed. The objective usually is to take advantage of the desirable mechanical properties or chemical resistance of one polymer while retaining the processing attributes of a different cure system [103].

It is common in acrylic structural adhesives to use oligomers that have a desirable backbone and are terminated with free-radical-polymerizable bonds. A variety of isocyanate-terminated polyurethanes can be adapted for use in acrylic structural adhesives by reacting the terminal isocyanates with a hydroxy functional acrylic monomer such as 2-hydroxyethyl methacrylate [86, 104, 105]. A recent acrylic/siloxane hybrid adhesive using a trialkylborane initiator has recently been described [106].

Sometimes the system is formulated so that multiple-cure mechanisms are possible and can occur sequentially or simultaneously. Compositions that rely on both epoxy and urethane chemistry are examples [68, 107, 108]. These are compositions containing the DGEBA, an isocyanate or isocyanate-terminated prepolymer, amines or other reactants for either epoxy or isocyanates, and catalysts.

Some of the more interesting and innovative work has occurred in areas combining aspects of more than one chemistry type. For instance, moisture-curable thermoplastic adhesives have received much attention. Hot-melt adhesives have been developed that contain active, moisture-curable isocyanate groups. The compositions provide rapid processing on assembly lines because a reasonable bond is formed as soon as the thermoplastic adhesive cools from the melt. However, bond strength and performance improve with time as the composition is slowly cross-linked to a thermoset by reaction of the isocyanates with atmospheric moisture [109, 110].

Evolution

The fundamental chemistry of the structural adhesives described here can change very little. Vinyl and acrylic monomers polymerize by chain-growth polymerization initiated by free radicals or ions. Isocyanate and epoxy compounds react with compounds containing active hydrogen in step-growth polymerizations. Epoxy-containing compounds undergo chain-growth polymerization initiated by certain Lewis acids and Lewis bases. These reactions will remain the most important ones for the polymerization of these raw materials.

Most of the raw materials used in large quantities in structural adhesives are used because they are widely available, relatively safe, and inexpensive. Quite often they are made in large quantities for uses other than adhesives and the adhesive manufacturers have taken advantage of supply and price. Few new basic raw materials are being developed specifically for the adhesive industry although the traditional raw materials are being combined in new ways to enhance desired adhesive properties. New initiators, adhesion promoters, primers, and specialty chemicals are being developed for use in small quantities to provide wider application latitude and improved performance.

The evolution of structural adhesives will certainly continue. Each increment in strength, durability, processing speed and ease, safety, reliability and reproducibility opens new commercial markets, not only to displace older joining methods but also to allow for the manufacture of new structures not possible without adhesives. One area receiving much academic as well as patent activity is “self-healing” systems. One such epoxy system is described by Wu et al. where the healing chemistry can be the same or different such as olefin methathesis [111].

Conclusion

Epoxy, urethane, and acrylic structural adhesives have been commercially successful because each can be used, under the proper conditions, to make reliable, durable, and useful adhesively joined assemblies. The adhesives are classified according to the polymerization reactions used to bring the liquid adhesive to a high-molecular-weight load-bearing state. The reactive sites on the uncured adhesive and the overall polarity of the cured adhesive are important in the adhesion of the polymer to any specific material. The polymerization reactions determine the processing requirements and are important to the mechanical properties and environmental resistance of the cured adhesive. Through the use of specific reactive oligomers it is possible to build certain molecular structures into the composition that will remain in the cured adhesive to provide desired mechanical or chemical properties.

The successful use of adhesive joining in producing any assembly depends on viewing the assembly as a whole from the very beginning of its design. The key is remembering that the adhesive is only one component of the assembly. Adhesion is a property of the whole assembly.

References

1. The Adhesion Society. http://www.adhesionsociety.org, 2 Davidson Hall—0201, Blacksburg, VA 24061, phone: 540-231-7257, fax: 540-231-3971, e-mail: adhesoc@vt.edu
2. Wake WC (1982) Adhesion and the formulation of adhesives, 2nd edn. Applied Science, New York
3. Lee L-H (1984) Adhesive chemistry, developments and trends. Plenum, New York
4. Panek JR, Cook JP (1984) Construction sealants and adhesives, 2nd edn. Wiley-Interscience, New York
5. Adams RD, Wake WC (1984) Structural adhesive joints in engineering. Elsevier, New York
6. Hartshorn SR (ed) (1986) Structural adhesives chemistry and technology. Plenum, New York
7. Kinloch AJ (1987) Adhesion and adhesives science and technology. Chapman and Hall, London
8. Tong L, Steven GP (1999) Analysis and design of structural bonded joints. Kluwer Academic, Boston
9. Ward N, Young T (2001) The complete guide to glues and adhesives. Krause, Iola
10. Gierenz G (ed) (2001) Adhesives and adhesive tapes. Wiley, New York
11. Moore DR (ed) (2001) Fracture mechanics testing methods for polymers, adhesives and composites. Elsevier, Amsterdam
12. Veselovskii RA, Kestelman VN, Veselovsky RA (2001) Adhesion of polymers. McGraw-Hill, New York
13. Pocius AV (2002) Adhesion and adhesives technology. Hanser Gardner, Munchen
14. Pocius AV, Dillard D, Chaudhury M (2002) Surfaces, chemistry and applications: adhesion science and engineering. Elsevier Science & Technology, Amsterdam
15. Pizzi A (ed) (2003) Handbook of adhesive technology. Marcel Dekker, New York
16. Swanson DW (2005) Adhesive materials for electronic applications: polymers, bonding, and reliability. William Andrew, Norwich
17. Fisher LW (2005) Selection of engineering materials and adhesives. CRC, Boca Raton
18. Packham DE (2005) Handbook of adhesion. Wiley, Hoboken
19. Dillard D (2010) Advances in structural adhesive bonding. Woodhead Publishing, Cambridge
20. Adams RD (2005) Adhesive bonding, science, technology, and applications. CRC, Boca Raton
21. Petrie EM (2006) Epoxy adhesive formulations. McGraw-Hill, New York
22. Pizzi A, Mittal KL (2003) Handbook of adhesive technology. Marcel Dekker, New York
23. Yacobi BG, Hubert M (2003) Adhesive bonding in photonics assembly and packaging. American Scientific, Stevenson Ranch
24. Chemical Abstracts Service (2005) American Chemical Society
25. Fowkes FM, Mostafa MA (1978) Acid-base interactions in polymer adhesion. Ind Eng Chem Prod Res Dev 17(1):3–7
26. Fowkes FM (1983) Acid-base interactions: relevance to adhesion science and technology. In: Mittal KL (ed) Physicochemical aspects of polymer surfaces, vol 2. Plenum, New York, p 583
27. Mark HF (1979) Future improvements in the cohesive and adhesive strength of polymers—part I. Adhes Age 22(7):35–40
28. Mark HF (1979) Future improvements in the cohesive and adhesive strength of polymers—part II. Adhes Age 22(9):45–50
29. May CA (ed) (1988) Epoxy resins: chemistry and technology, 2nd edn. Marcel Dekker, New York
30. Castan P (1943) Process of preparing synthetic resins. US Patent 2,324,483
31. Dow Chemical (1999) Dow Liquid Epoxy Resins, Form No. 296-00224-0199 WC_M. http://epoxy.dow.com/index.htm
32. Hexion Specialty Chemical (a Momentive company). http://www.hexion.com/Products/TechnicalDataSheets.aspx?id=2273
33. Reichhold, Inc. http://www.reichhold.com/en/coatings-products.aspx?cat=Brands&pid=24
34. Polystar, LLC. http://www.polystarusa.com/products.php?c=Epoxy+Resins&cid=10
35. Huntsman Corp., EPON Product Overview 2009-08. http://www.huntsman.com/advanced_materials/eng/Home/Coatings/Publications/Coatings/index.cfm?PageID=5898
36. Hexion Specialty Chemical (a Momentive company), Eponol™ Resin 53-BH-35
37. Partansky AM (1968) A study of accelerators for epoxy-amine condensation reaction. Am Chem Soc Div Org Coat Plast Chem 28(1):366
38. Partansky AM (1970) A study of accelerators for epoxy-amine condensation reaction. In: Lee H (ed) Epoxy resins. Am Chem Soc Adv Chem Ser 92:29–47. doi:10.1021/ba-1970-0092.ch004
39. Hine J, Linden S-M, Kanagasabapathy VM (1985) Double-hydrogen-bonding catalysis of the reaction of phenyl glycidyl ether with diethylamine by 1,8-biphenylenediol. J Org Chem 50(25):5096–5099. doi:10.1021/jo00225a021
40. Matejka L, Pokomy S, Dusek K (1982) Network formation involving epoxide and carboxyl groups. Polym Bull (Berlin) 7(2–3):123–128
41. Dusek K, Matejka L (1984) Transesterification and gelation of polyhydroxy esters formed from diepoxides and dicarboxylic acids. In: Riew CK, Gillham JK (eds) Rubber modified thermoset resins. Am Chem Soc Adv Chem Ser 208:15–26. doi:10.1021/ba-1984-0208.ch002
42. Abbey KJ, Pressley MW, Durso SR (1999) Controlled cure of thiol-epoxy systems. In: Speth DR (ed) Proceedings of the 22nd

annual meeting of The Adhesion Society, Panama City Beach, FL, 21–24 Feb 1999
43. Crivello JV (1979) Heat curable compositions. US Patent 4,173,551
44. Crivello JV (1980) Heat curable cationically polymerizable compositions and method of curing same with onium salts and reducing agents. US Patent 4,216,288
45. Crivello JV (1982) UV curable compositions and substrates treated therewith. US Patent 4,319,974
46. Crivello JV, Ashby BA (1982) Methods of adhesive bonding using visible light cured epoxies. US Patent 4,356,050
47. Crivello JV, Lee JL (1984) Photocurable compositions. US Patent 4,442,197
48. Yagci Y, Reetz I (1998) Externally stimulated initiator systems for cationic polymerization. Prog Polym Sci 23(8):1485–1538
49. Tarbutton KS, Robins J (1989) Acid catalyzed, toughened epoxy adhesives. US Patent 4,846,905
50. Khan NH, Razi Abdi SH, Kureshy RI, Singh S, Ahmad I, Jasra RV, Ghosh PK (2007) Catalytic process for the preparation of epoxides from alkenes. US Patent 7,235,676
51. Mikawa M, Uchida S-I (2003) Catalyst for production of epoxides and methods for production thereof and epoxides. US Patent 6,600,056
52. Kinloch AJ, Lee JH, Taylor AC, Sprenger S, Eger C, Egan D (2003) Toughening structural adhesives via nano- and microphase inclusions. J Adhes 79:867–873. doi:10.1080/00218460390242234
53. Kinloch AJ, Mohammed RD, Taylor AC, Eger C, Sprenger S, Egan D (2005) The effect of silica nano particles and rubber particles on the toughness of multiphase thermosetting epoxy polymers. J Mater Sci 40:5083–5086. doi:10.1007/s10853-005-1716-2
54. Liang YL, Oldak RK, Pearson RA (2007) Particle size effect in rubber-glass sphere toughened epoxies. In: Jagota A (ed) Proceedings 30th annual meeting of The Adhesion Society, Inc., The Adhesion Society, Blacksburg, VA, pp 343–345, 2007
55. Ragosta G, Abbate M, Musto P, Scarinzi G, Mascia L (2005) Epoxy-silica particulate nanocomposites: chemical interactions, reinforcement and fracture toughness. Polymer 46 (23):10506–10516. doi:10.1016/j.polymer.2005.08.028
56. Oldak RK, Hydro RM, Pearson RA (2007) On the use of triblock copolymers as toughening agents for epoxies. In: Jagota A (ed) Proceedings 30th annual meeting of The Adhesion Society, Inc., The Adhesion Society, Blacksburg, VA, pp 153–156, 2007
57. Boogh L, Pettersson B, Månson J-AE (1999) Dendritic hyperbranched polymers as tougheners for epoxy resins. Polymer 40:2249–2261. doi:10.1016/S0032-3861(98)00464-9
58. Mezzenga R, Boogh L, Petterson B (2000) Chemically induced phase separated morphologies in epoxy resin-hyperbranched polymer blends. Macromol Symp 149:17–22. doi:10.1002/1521-3900(200001)149:1<17::AID-MASY17>3.0.CO;2-O
59. Mezzenga R, Månson JAE (2001) Thermo-mechanical properties of hyperbranched polymer modified epoxies. J Mater Sci 36:4883–4891. doi:10.1023/A:1011880019156
60. Mezzenga R, Plummer CJG, Boogh L, Månson JAE (2001) Morphology build-up in dendritic hyperbranched polymer modified epoxy resins: modeling and characterization. Polymer 42:305–317. doi:10.1016/S0032-3861(00)00307-4
61. Bayer O (1937) Verfahren zur Herstellung von Polyurethanen bzw. Polyharnstoffen. German Patent 728.981
62. Saunders JH, Frisch KC (1962, 1964) Polyurethanes, chemistry and technology, vols 1 and 2. Interscience, New York
63. Oertel G (ed) (1994) Polyurethane handbook, 2nd edn. Hanser, Munich
64. Randall D, Lee S (eds) (2003) The polyurethane book. Wiley, New York
65. Dunn DJ (2004) Engineering and structural adhesives (Rapra Review Reports, 15(1), Report 169). Rapra Technology Ltd., Shropshire
66. Rapra Technology Ltd (2004) Polyurethane adhesives (Rapra Published Search Number 114). Rapra Technology Ltd, Shropshire
67. Fabris HJ, Maxey EM, Uelzmann H (1973) Urethane adhesive having improved sag resistance. US Patent 3,714,127
68. Goel AB (1988) Sag resistant urethane adhesives with improved antifoaming property. US Patent 4,728,710
69. Baueriedel H (1986) Adhesives based on polyurethane prepolymers having a low residual monomer content. US Patent 4,623,709
70. Bolte G, Henke G, Kruedenscheidt M, Omoruyi A (2003) Low monomer polyurethane prepolymer and process therefore. US Patent 6,515,164
71. Ono H-K, Jones FN, Pappas SP (1985) Relative reactivity of isocyanate groups of isophorone diisocyanate. Unexpected high reactivity of the secondary isocyanate group. J Polym Sci Polym Lett Ed 23:509–515. doi:10.1002/pol.1985.130231003
72. Okawa T (1992) Process for producing isocyanate compound. US Patent 5,166,414
73. Mason RW (2004) Non-phosgene route to the manufacture of organic isocyanates. US Patent 6,781,010
74. Nagata T, Yamashita H, Kusumoto M, Okazaki K (1994) Stabilizing method of isocyanate compounds and isocyanate compositions stabilized thereby. US Patent 5,302,749
75. Rosen R, Bernard J-M, Duffy CS (2007) Color stability of isocyanates. US Patent 7,297,746
76. Ohashi Y, Matsuda H, Nishi E, Nishida T (1997) Moisture curing urethane adhesive composition. US Patent 5,698,656
77. Fukatsu S, Hattori Y (1998) Moisture curable polymer composition and process for production thereof. US Patent 5,767,197
78. Suen W (2010) Two component polyurethane adhesive. US Patent 7,834,123
79. Schoener TE, Housenick JB (2002) Polyurethane reaction system having a blocked catalyst combination. US Patent 6,348,121
80. Ambrose R, Retsch WR Jr, Chasser A (2001) Blocked isocyanate-based compounds and compositions containing the same. US Patent 6,288,199
81. Bader E (1967) Polychloroprene adhesives. US Patent 3,333,025
82. Bauer W Jr (2004) Acrylic acid and derivatives. In: Kirk-Othmer encyclopedia of chemical technology, 5(1):342–369. Seidel A, Ed. Wiley, New York. doi:10.1002/0471238961.010318250201210
83. Wilczynski R, Juliette JJ (2006) Methacrylic acid and derivatives. In: Kirk-Othmer encyclopedia of chemical technology, 16 (5):227–270. Seidel A, ED. Wiley, New York. doi:10.1002/14356007.a16_441
84. Huang J-P, Righettini RF, Dennis FG (2001) Adhesive formulations. US Patent 6,225,408
85. Briggs PC, Muschiatti LC (1975) Novel adhesive compositions. US Patent 3,890,407
86. Zalucha DJ, Sexsmith FH, Hornaman EC, Dawdy TH (1981) Structural adhesive formulations. US Patent 4,223,115
87. Dawdy TH (1984) Epoxy modified structural adhesives having improved heat resistance. US Patent 4,467,071
88. McCormick FB, Drath DJ, Gorodisher I, Kropp MA, Palazzotto MC, Sahyun MRV (2000) Energy-curable cyanate/ethylenically unsaturated compositions. US Patent 6,069,219
89. Righettini RR, Dawdy TH (1999) Free radical polymerizable compositions including para-halogenated aniline derivatives. US Patent 5,932,638
90. Edelman R, Catena W (1999) Rapid curing structural acrylic adhesive. US Patent 5,865,936
91. Usifer DA, Broderick IC (1995) Urethane adhesive compositions. US Patent 5,426,166

92. Usifer DA, Broderick IC (1996) Urethane adhesive compositions. US Patent 5,484,864
93. Hu S (2006) Radiation-curable compositions for optical fiber coating materials. US Patent 7,105,583
94. Maxlok™ acrylic adhesives. http://lordmaxlok.com
95. Imai Y, Fujisawa S, Matsui H, Yamazaki H, Masuhuara E (1973) Japanese Kokai 48-100477
96. Fujisawa S, Imai Y, Masuhara E (1969) Studies on dental self-curing resins: 2. Characterization of the various complexes of Tri-n-butyl borane as an initiator, Iyo Kizai Kenkyujo hokoku. Reports of the Institute for Medical and Dental Engineering, Tokyo Medical and Dental University 3:64–71
97. Zharov JV, Krasnov JN (1996) Polymerizable compositions made with polymerization initiator systems based on organoborane amine complexes. US Patent 5,539,070
98. 3M Technical Literature for DP-8010, Jan 2002
99. Sonnenschein MF, Webb SP, Rondan NG (2004) Amine organoborane complex polymerization initiators and polymerizable compositions. US Patent 6,706,831; and 15 more recent patents by the lead inventor
100. Leaversuch RD, Long-glass PP (2002) makes inroads in automotive front ends. Plast Technol. http://www.plasticstechnology.com/articles/200207cu1.html, Gardner
101. Kendall JL, Abbey KJ. (2004) Internally coordinated organoboranes. US Patent Application
102. Abbey KJ, Kendall JL (2004) Internally coordinated organoboranes: stability and activation in polyolefin adhesives. Polym Mater Sci Eng 91(2):175–176
103. Pohl E, Osterholz FD (1987) Novel vulcanizable silane-terminated polyurethane polymers. US Patent 4,645,816
104. Brownstein AM (1969) Anaerobic adhesive. US Patent 3,428,614
105. Su W-FA (1986) UV curable high tensile strength resin composition. US Patent 4,618,632
106. Sonnenschein MF, Webb SP, Wendt BL (2008) Poly(acrylate/siloxane) hybrid adhesives for polymers with low surface energy. Int. J. Adhesion and Adhesives 28(3):126–134. doi:10.1016/j.ijadhadh.2007.07.001
107. Hawkins JM (1972) Epoxy resin adhesive compositions containing an isocyanate terminated polyurethane prepolymer and a chain extender. US Patent 3,636,133
108. Trieves R, Pratley KGM (1986) Two-component adhesive or sealing composition. US Patent 4,623,702
109. Anderson GJ, Zimmel JM (1999) Thermally stable hot melt moisture-cure polyurethane adhesive composition. US Patent 5,939,499
110. McInnis EL, Santosusso TM, Quay JR (2001) Hot melt adhesives comprising low free monomer, low oligomer isocyanate prepolymers. US Patent 6,280,561
111. Wu DY, Meure S, Solomon D (2008) Self-healing polymeric materials: a review of recent developments. Prog Polym Sci 33:479–522. doi:10.1016/j.progpolymsci.2008.02.001

Synthetic Resins and Plastics

15

Rudolph D. Deanin, Joey L. Mead, Ming Wei, and Bridgette M. Bludhlall

Introduction

Definition

Plastic (adj.) is defined by Webster [1] as "capable of being molded or modeled (e.g., clay) ... capable of being deformed continuously and permanently in any direction with-out rupture." Plastic (n.) a plastic substance; *specifically*: any of numerous organic synthetic or processed materials that are mostly thermoplastic or thermosetting polymers of high molecular weight and that can be made into objects, films, or filaments.

History

Commercial plastics began in 1868 when John and Isaiah Hyatt plasticized cellulose nitrate and molded it into billiard balls, and later into combs, brushes, and other useful articles. Commercial *synthetic* plastics began in 1908 when Leo Baekeland reacted phenol, formaldehyde, and wood flour and molded them into electrical insulators and a growing variety of other product specialties. During the next 30 years, a variety of plastics appeared as interesting specialties. During World War II the U.S. and German governments met the severe demands of advanced military technology by supporting vast research, development, plant construction, and manufacturing of a number of major polymers for plastics and rubber applications. With the end of the war in 1945, all of this technology and production capacity converted to civilian products, and commodity plastics began their tremendous growth, first polyvinyl chloride and polystyrene, then polyethylenes, polypropylene, polyesters, and polyurethanes, along with a constantly expanding range of specialized polymers for more demanding applications.

Advantages of Plastics over Conventional Materials

We have had conventional structural materials—metals, ceramics, glass, wood, leather, textiles, and paper—for thousands of years, during which time we have been able to explore and exploit them thoroughly to full maturity. By comparison, plastics are relatively new so that we are still learning to develop and use them. Their rapid growth is due to the fact that they offer many advantages over conventional materials.

Processability and Product Design. It is much easier to convert plastic materials into an almost unlimited range of products.

Modulus. Plastics cover a broad range from extremely rigid to stiffly flexible to very soft and rubbery, sometimes even within a single chemical family.

Elasticity. Plastics recovery from deformation is superior to almost all conventional materials.

Impact Strength. Plastics are much less brittle than ceramics, glass, and paper, and some families can be made nearly unbreakable.

Lubricity and Abrasion Resistance. Specific plastic materials offer outstanding self-lubricating performance and abrasion resistance.

Thermal Insulation. Plastics offer very good insulation against heat and cold, and can be foamed to increase their insulating qualities.

R.D. Deanin • J.L. Mead (✉) • M. Wei • B.M. Bludhlall
Department of Plastics Engineering, University of Massachusetts Lowell, 1 University Avenue, Lowell, MA 01854, USA
e-mail: Joey.Mead@uml.edu

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_15,

Flame-Retardance. Many plastics are less flammable than wood and paper, and most can be formulated to make them much more resistant to burning.

Electrical Insulation. Plastics are generally excellent electrical insulators, however, they can also be formulated to provide semi-conductivity, and high or low dielectric constant and loss.

Transparency/Opacity. Some plastics can approach or equal the transparency of glass. All can be formulated to a wide range of translucency to opacity.

Color and Appearance. Most conventional materials have a very limited color range governed by the material. In contrast, plastics can be produced in an almost infinite range of colors, transparent or opaque, and can have textured surfaces to simulate a variety of other materials (e.g., leather).

Chemical Resistance. Plastics are generally far superior to metals in corrosion resistance. Among the 100 families of commercial plastics, they offer a wide variety of chemical resistance, or solubility/reactivity, as desired for different types of products.

Water Resistance. Plastics are far superior to wood and paper.

Permeability. Different families of plastics offer a wide range from highly impermeable barrier materials to membranes of different separation abilities.

Weathering. Some plastics are very resistant to weather. Others have moderate resistance. Still others can be designed to degrade to alleviate collection of solid waste.

Cost. Contrary to popular belief, plastics are not "cheap," but their superior processability often makes the finished product less expensive.

This economic advantage, plus all the other advantages listed above, accounts for the success of plastics as replacements for conventional materials and in leading to new products not previously possible.

Markets for Plastics

The U.S. plastics market has grown past 98.7 billion pounds/year [2, 3]. The largest share (39%) goes to packaging, both rigid packages and film. Close behind is building and construction (26%) mainly piping, plywood, siding, insulation, and flooring. Smaller, more specialized amounts go into agriculture, aircraft, appliances, autos and trucks, electrical and electronics, furniture, glazing, housewares, luggage, marine, medical, office equipment, optical, tools, toys, and miscellaneous industrial and consumer products. Thus, taken as a whole, plastics is one of the largest and fastest growing industries in the United States.

Major Classes of Plastic Materials

The primary classification is the distinction between thermoplastics and thermosets. Thermoplastics are stable large molecules, typical molecular weights of 10^4–10^6, which soften on heating to permit melt processing, and solidify on cooling to give solid finished products; the process is reversible, so they are essentially recyclable within the limits of their thermal stability. Thermosetting plastics are reactive low molecular weight polymers, which may be melted or even poured, shaped into final products, and then reacted further into cross-linked molecules of essentially infinite molecular weight; the process is essentially irreversible, so they are difficult or impossible to recycle. Thermoplastic processing is simpler and more economical, thus thermoplastics account for more than 70% of all polymers. On the other hand, thermosetting plastics permit many special processes and offer outstanding final properties, which accounts for the health of their share of their more specialized markets.

Another important distinction is based on: (1) filled and reinforced plastics and (2) foams.

1. When any family of polymers is combined with particulate inorganic fillers, this produces a major increase in density, modulus, dimensional stability, heat transfer, dielectric constant, opacity, and frequently a decrease in cost. When the fillers are reinforcing fibers, they can further produce a great increase in strength, impact resistance, and dimensional stability. Thus, these properties may depend more upon the use of fillers and fibers, than upon the choice of the particular polymer in which they are used. Recently, numerous nanoscale fillers, including nanoparticles, nanotubes, and nanoclays have been used to reinforce the polymer at relatively low loadings compared to conventional fillers. In this case, good dispersion is very important to produce high quality materials.
2. When a polymer is liquefied, foamed, and solidified to trap the air spaces within it, the material exhibits drastically different properties compared to the base polymer. The most outstanding effects are flotation by closed-cell foams, softness in open-cell foams, impact cushioning, thermal and electrical insulation, and permeability in open-cell foams.

Following these introductory remarks, we turn to a study of the fundamental aspects of Polymer Chemistry (Part I), followed with a discussion of Commercial Plastic Materials (Part II), and conclude with Plastics Processing (Part III).

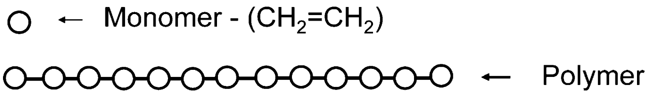


Fig. 15.1 Polymerization

Part I: Polymer Chemistry

Materials can be broadly classified as ceramics, metals, or polymers. Compared to metals and ceramics, polymers exhibit lower moduli, thermal conductivities, and densities. A wide range of products can be manufactured using plastic materials. Polymeric materials are used in automotive, packaging, and consumer goods, just to name a few. The requirements for these diverse applications vary greatly, but through proper control, plastic materials can be compounded or synthesized to meet these varied product requirements. Their lower density offers an advantage over other materials in applications requiring lighter weight.

Molecular Weight

A polymer is prepared by linking a low molecular weight species, called a monomer (such as ethylene), into a long chain, called a polymer (such as polyethylene), similar to stringing together beads to make a necklace (see Fig. 15.1). As molecular weight increases, the properties of the material change. Looking at the alkane hydrocarbon series with the general structure H–(CH_2)$_n$–H, we can see the material change from a gas for values of $n = 1$–4, a liquid for $n = 5$–11, a high viscosity liquid from $n = 16$–25, a crystalline solid for $n = 25$–50, to a tough plastic solid for $n = 1{,}000$–$5{,}000$ [4]. The molecular weight affects both the mechanical and processing behavior of the polymer. In general, higher molecular weights result in improved mechanical properties, but more difficulty in processing. In contrast to low molecular weight species, polymeric materials do not possess one unique molecular weight, but rather a distribution of molecular weights as depicted in Fig. 15.2. Polymer molecular weight can be characterized using two different molecular weight averages, the number average molecular weight, $\overline{M_n}$, and the weight average molecular weight, $\overline{M_w}$. These averages are:

$$\overline{M_n} = \sum_{i=1}^{\infty} \frac{n_i M_i}{n_i}$$

$$\overline{M_w} = \sum_{i=1}^{\infty} \frac{n_i M_i^2}{n_i M_i}$$

where n_i is the number of moles of species i and M_i is the molecular weight of species i.

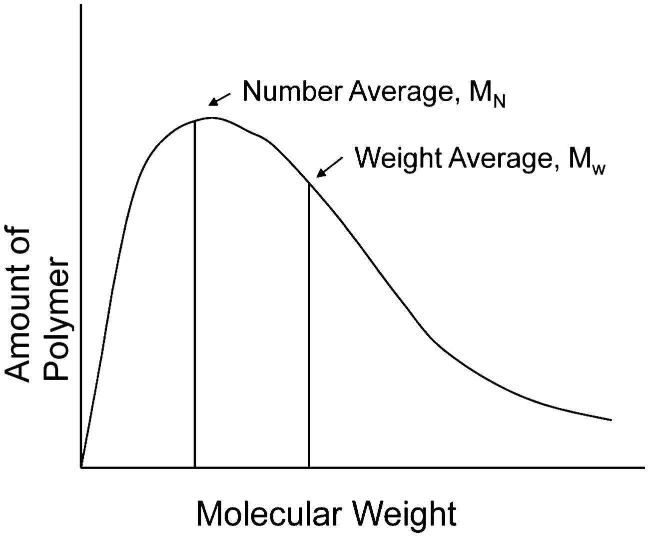


Fig. 15.2 Molecular weight distribution

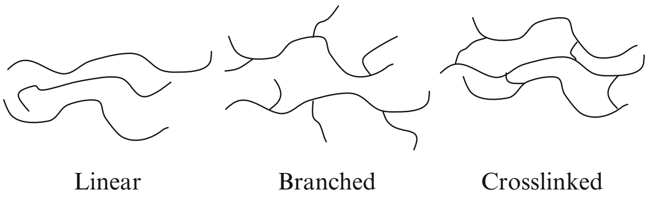


Fig. 15.3 Polymer structures

Chain Structure

Polymerization can produce linear chains, but other structures can also exist. As shown in Fig. 15.3 branched and cross-linked structures can be formed. Linear and branched structures can be shaped and reshaped simply by heating and are called thermoplastics. In the case of a cross-linked structure a three-dimensional network is formed that cannot be reshaped by heating. This type of structure is called a thermoset.

Macromolecular conformations describe the positions of the atoms that occur due to rotation about the single bonds in the main chain [5]. Polymer chains in solution, melt, or amorphous state exist in what is termed a random coil. The chains may take up a number of different conformations, varying with time. Figure 15.4 shows one possible conformation for a single polymer chain. In order to describe the chain, polymer scientists utilize the root mean square end-to-end distance ($<r^2>^{1/2}$), which is the average over many conformations. This end-to-end distance is a function of the bond lengths, the number of bonds, and a characteristic ratio, C, for the specific polymer.

Chemical Structure

The chemical characteristics of the monomer will control the properties of the final polymer. Along the chain axis, primary bonds hold the atoms together and determine molecular properties, such as flexibility and glass transition temperatures. Flexibility of the chain is governed by the ease of rotation about main chain bonds. The presence of methylene units or carbon–oxygen single bonds acts to increase the flexibility of the chain. Groups or interchain interactions (described below) that restrict rotation will decrease the flexibility of the molecule. An extreme example of this principle is the rigid rod polymers. Figure 15.5 shows the structures for two rigid rod polymers. The aromatic groups along the backbone act to stiffen the polymer chain and restrict rotation, causing the polymer to remain straight, much like a log [6]. Materials such as these may exhibit liquid crystalline behavior.

The forces holding the many individual chains together (interchain forces) are determined by secondary bonds, except in the case of thermosets where primary bonds hold the chains together. The type and strength of the secondary bonds (often termed van der Waals forces) [7] will depend on the chemical structure of the polymer. In the case of hydrocarbon polymers, such as polyethylene, the secondary bonds are dispersion forces. For polymers containing carbon and oxygen groups, such as the polyesters, the presence of the –C=O bond results in a dipole from the different electronegativities of the carbon and oxygen atoms. The presence of polarity in the polymer acts to increase the intermolecular forces. In the case of polyamides, hydrogen bonding between the polymer chains leads to high intermolecular forces. The strength of the intermolecular forces affects polymer properties including: viscosity, solubility, miscibility, surface tension, and melting point. As discussed above, interchain forces also affect the flexibility of the chain if they restrict free rotation.

Morphology

Polymer chain structure and processing conditions control the morphology of the solid polymer. Amorphous polymers show no order to the arrangement of the chains. The chains are entangled with each other, much like the strands of spaghetti on one's plate. Polystyrene is an example of an amorphous polymer. If the polymer backbone has a regular, ordered microstructure, then the polymer has the potential to crystallize, however, the material will generally be only semicrystalline. Polyethylene and polypropylene are typical semicrystalline polymers. The amorphous and crystalline chain structures are depicted in Fig. 15.6.

The exact makeup and architecture of the polymer backbone will determine the ability of the polymer to crystallize. Figure 15.7 shows the different types of microstructure that can be obtained for a vinyl polymer. Isotactic and syndiotactic structures are considered stereospecific

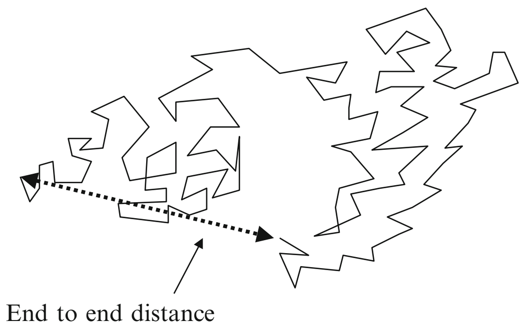


Fig. 15.4 Random coil chain

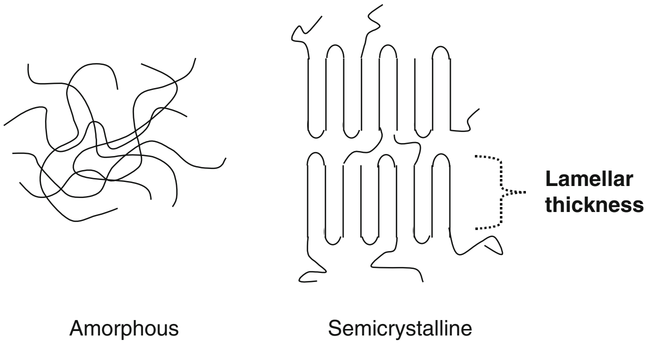


Fig. 15.6 Amorphous and semicrystalline structures

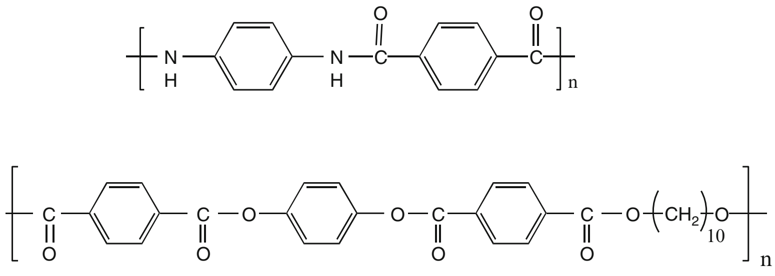


Fig. 15.5 Rigid rod type polymers

Fig. 15.7 Syndiotactic, atactic, and isotactic structures

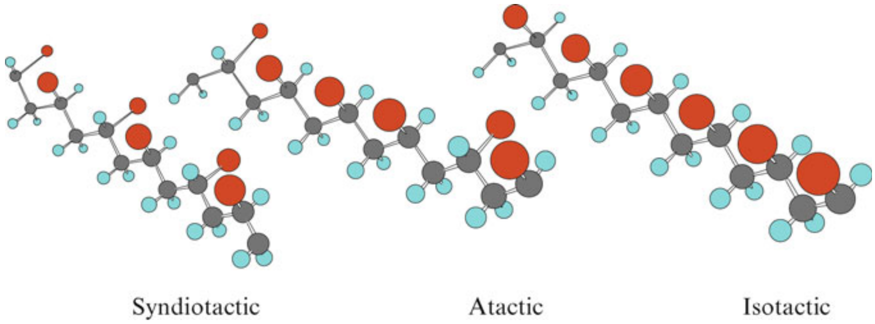
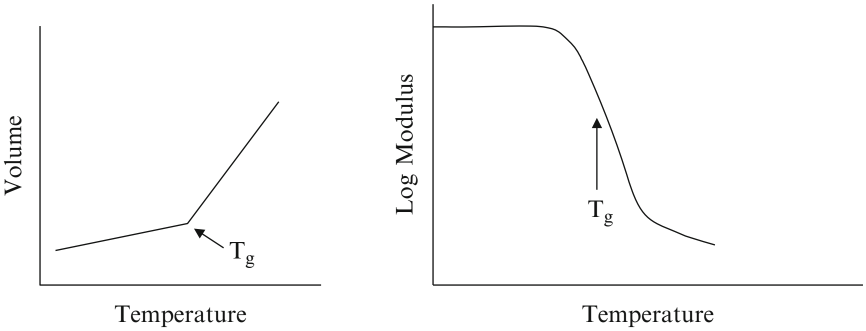


Fig. 15.8 Behavior of volume and modulus at the glass transition temperature



polymers, and their highly regular backbone structure allows them to crystallize. The atactic form is irregular and would produce an amorphous material. This nature of the polymer microstructure can be controlled by using different synthetic methods. As will be discussed below, the Ziegler–Natta catalysts are capable of controlling the microstructure to produce different types of stereospecific polymers.

Transition Temperatures

Glass Transition Temperature (T_g)

Chain flexibility is governed by molecular structure, but is also affected by temperature. As the temperature is reduced, amorphous polymers reach a temperature where large-scale (20–50 chain atoms) segmental motion ceases [8]. This temperature is called the glass transition temperature or T_g. Volume–temperature plots can be used to indicate the T_g. At the glass transition temperature, a several decade change in the modulus occurs and the material changes from a rigid solid to a rubbery material. At higher temperatures, provided the material is a thermoplastic, it becomes a liquid, which can flow and be processed. The behavior in the glass transition region is depicted in Fig. 15.8.

The location of the glass transition temperature will depend on the nature of the polymer. A rubbery material is distinct from a plastic material as a result of the location of its glass transition temperature (T_g). The T_g of a plastic is above room temperature and the T_g of rubber is below room temperature. The flexibility of the chain affects the value of T_g. Flexible groups tend to lower the T_g, while stiffening groups act to increase it. Side groups can also affect the value. The effect of various factors on the value of T_g is shown in Table 15.1. Addition of aliphatic side groups show decreasing T_g's as the length of the side chain increases; however, rigid side groups increase the T_g. Increased polarity of the polymer also increases the glass transition temperature as shown by the polymers polypropylene, polyvinyl chloride, and polyacrylonitrile. Symmetry about the backbone can act to decrease the T_g as illustrated by the pairs polypropylene and polyisobutylene, and polyvinyl chloride and polyvinylidene chloride.

Increases in number average molecular weight, $\overline{M_n}$, and cross-link density both act to increase the T_g. The addition of plasticizers, as in the case of polyvinyl chloride, decreases the value of the glass transition temperature. The glass

Table 15.1 Effect of chain structure on T_g [9]

	T_g (°C)
Side groups	
Polypropylene	−10
Polystyrene	100
Polarity	
Polypropylene	−10
Polyvinyl chloride	85
Polyacrylonitrile	101
Symmetry	
Polypropylene	−10
Polyisobutylene	−70
Polyvinyl chloride	87
Polyvinylidene chloride	−19

transition temperature of the plasticized material can be estimated if the glass transition temperature of the two components (A and B) and their weight fractions are known by [10]

$$\frac{1}{T_g} = \frac{W_A}{T_{gA}} + \frac{W_B}{T_{gB}}$$

where T_{gA} and T_{gB} are the glass transition temperatures of components A and B, respectively, and W_A and W_B are the weight fractions.

One difficulty in obtaining values for the glass transition temperature is its dependence on measurement rate. When experiments are conducted a slow rates, the measured values will be lower than those measured at more rapid rates. Other difficulties include experimental problems and many definitions and interpretations on how to measure the values. Measurement of the temperature at which a step change in the volumetric thermal expansion coefficient occurs, when heating and cooling rates are 1°C/min, is one of the methods [11]. The measured value of the glass transition temperature will increase approximately 3°C (volumetric measurements) to 7°C (maximum in tan δ from dynamic mechanical analysis) for a decade change in rate.

Crystallization and Melting Points (T_m)

A number of polymers may exhibit the ability to crystallize. Such polymers are semicrystalline, meaning they will have regions of amorphous and crystalline material dispersed throughout the part. Both the morphology and degree of crystallinity can be affected by the processing conditions used to manufacture the part. These changes can greatly affect the mechanical behavior of the material. The degree of crystallinity and the melting point are often measured using differential scanning calorimetry (DSC).

The general structure of the crystalline regions is rather complex. Polymer chains appear to fold back and forth into a lamellar structure as shown in Fig. 15.6. These lamellae form layers of ribbon like structures, tied together with amorphous regions.

Semicrystalline polymers do not exhibit a single sharp melting point, but rather melt over a range of temperatures. In addition, the melting point of the material will be higher the greater the lamellar thickness, which can be controlled by the crystallization temperature [12]. Higher crystallization temperatures generally lead to greater lamellar thickness.

The chemical structure of the polymer affects the melting point. Polymers containing polar groups will have higher melting points. Polymers with hydrogen bonding would be expected to have even higher melting points. For example, polyethylene has a melting point of 135°C, while polyamide 6 has a melting point of 265°C. The presence of chain stiffening groups in the backbone will act to raise the melting point of the polymer; contrast the melting point of polyethylene (135°C) with poly-*p*-xylene (400°C). In general, polymers with high interchain forces and rigid chains will tend to have the highest melting points.

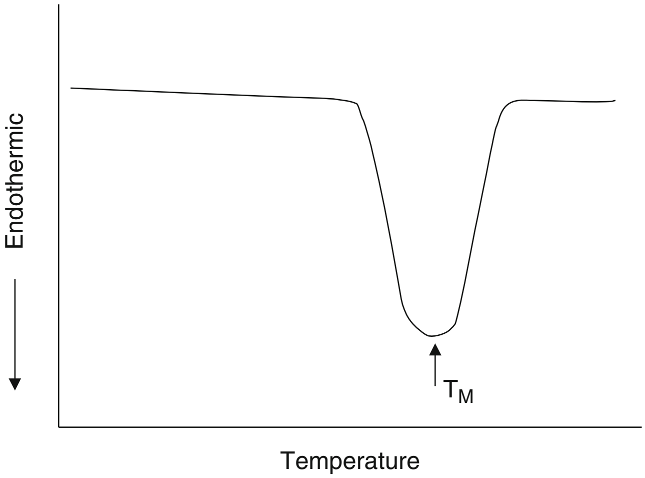


Fig. 15.9 Differential scanning calorimetry (DSC) curve of a melting point

In addition to the melting point, the degree of crystallinity influences the behavior of the polymer. DSC can be used to determine both the melting point and the % crystallinity. Figure 15.9 shows a representative DSC curve for a melting point. The area under the melting peak is related to the % crystallinity of the polymer sample. The percent crystallinity may have a dramatic effect on the mechanical behavior of the material. Above the T_g of the material, the presence of crystalline regions increases the rigidity of the polymer, resulting in an increasing modulus with increasing percent crystallinity [13].

The kinetics of crystallization are a function of the crystallization temperature [14]. As the crystallization temperature is decreased the rate of crystallization increases, up to a maximum. When the crystallization temperature is lowered beyond this point the rate begins to slow down as the molecular motion of the chains is reduced. At temperatures below the T_g crystallization is suppressed.

Polymerization

The two major polymerization approaches are step-reaction and chain-reaction polymerization [15].

Step-Reaction Polymerization

In step-reaction polymerization (also referred to as condensation polymerization) any two polyfunctional (typically bifunctional) monomers of the correct chemical species

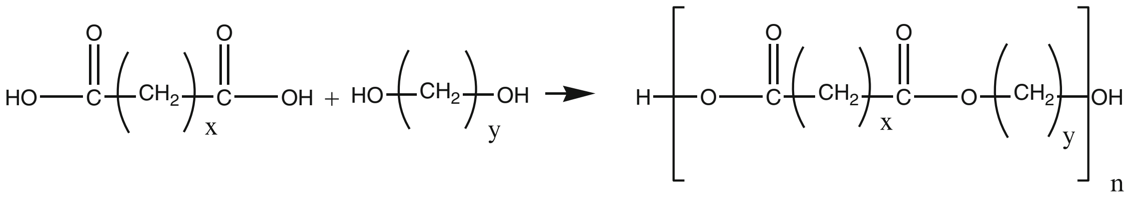


Fig. 15.10 Step-growth polymerization

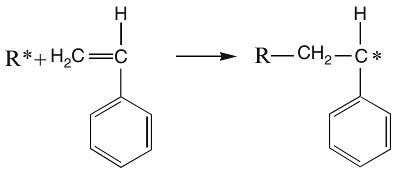


Fig. 15.11 Addition polymerization

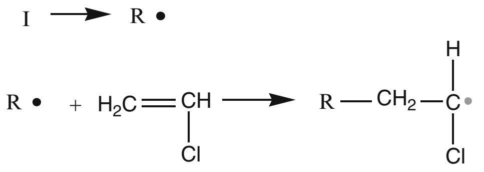
can react. Higher molecular weight species are produced as longer and longer polyfunctional groups react together during the course of polymerization. Monomers may react to form a dimer, then combine to form a tetramer, expressed as x-mer + y-mer → (x + y)mer, where x and y are the number of monomer units for each reactant. Monomer functionalities of two will produce linear polymers, while functionalities of greater than two give branching or cross-linking. The polymer molecular weight increases with time and in order to produce high molecular weight polymers high conversions are usually required. Atoms other than carbon are often found in step-reaction polymers as seen in polyesters and polyamides, which are typical step-reaction polymers. A typical step-growth reaction is shown in Fig. 15.10.

Molecular weight control can be accomplished by stopping the reaction; however, further heating may result in changes in the molecular weight. A stoichiometric imbalance (excess of one of the reactants over another) or a monofunctional reagent can also be used to control the molecular weight [16].

Chain-Reaction Polymerization

In contrast to the slow step-reaction polymerizations, chain-reaction polymerizations are fairly rapid [17]. An initiator is required to start chain-reaction polymerizations (often referred to as addition polymerizations). Free radical, anionic, or cationic species can be used as initiators. The initiator opens the double bond of a vinyl monomer and the reaction proceeds as shown in Fig. 15.11 where * may be a radical, cationic or anionic species. Chain-reaction polymers normally have only carbon in their backbone. Typical chain-reaction polymers are polystyrene and polyvinyl chloride. Once initiated, the chain will continue to add monomer until a termination reaction stops the growth. Termination generally occurs in radical type reactions, but in anionic and some cationic polymerizations termination reactions may not be present and the polymerization is termed a “living polymerization.” This has important commercial implications in that a second monomer may be added to the living end to form block copolymers.

Radical Polymerization. Free radical polymerization consists of three steps, initiation, propagation, and termination or chain transfer. Initiation consists of two steps, decomposition of the initiator to form a radical species, followed by addition of the initiator to the first monomer unit.

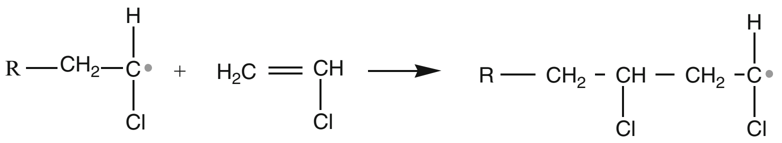


$$I \longrightarrow R\bullet$$

$$R\bullet + H_2C{=}\underset{Cl}{CH} \longrightarrow R{-}CH_2{-}\overset{H}{\underset{Cl}{C}}\bullet$$

Typical initiators include benzoyl peroxide or azobisisobutyronitrile, but other radical generating methods, such as high-energy radiation or photolytic decomposition of compounds, may also be used [12].

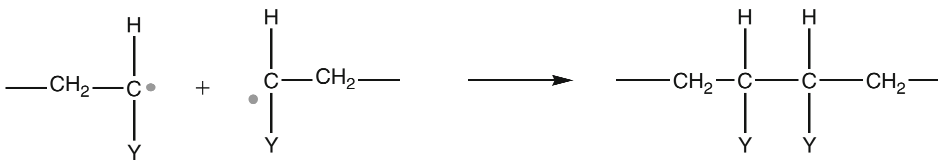
The next step in the reaction is propagation



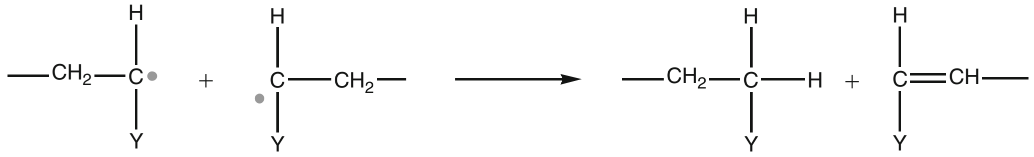
$$R{-}CH_2{-}\overset{H}{\underset{Cl}{C}}\bullet + H_2C = \underset{Cl}{CH} \longrightarrow R - CH_2 - \underset{Cl}{CH} - CH_2 - \overset{H}{\underset{Cl}{C}}\bullet$$

Propagation continues until the radical is terminated. Termination occurs when two radical species meet and react either by coupling or by disproportionation as shown in Fig. 15.12.

In radical polymerizations other reactions occur to prematurely stop chain growth. These reactions are termed chain transfer and when they occur they reduce the molecular weight of the chain. Chain transfer acts to transfer the



Coupling

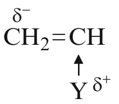


Disproportionation

Fig. 15.12 Termination mechanisms

radical from the growing end of the chain to another species. This may include solvent, initiator, polymer, or a deliberately added chain transfer agent. Propagation now begins from the new radical generated, while the original chain is terminated. When chain transfer to polymer occurs this produces branching along the polymer backbone. Low-density polyethylene (LDPE) is produced via free radical processes, with considerable transfer to polymer. Branching along the polyethylene backbone suppresses the degree of crystallinity, resulting in a lower density material.

Cationic Polymerization. Cationic polymerization follows similar steps as with free radical polymerization. Initiator types are typically acids, such as H_2SO_4 and H_3PO_4, and Lewis acids, such as $AlCl_3$, BF_3, $TiCl_4$, and $SnCl_4$. Lewis acids generally require the presence of a proton source, such as trace amounts of water [18]. Monomers exhibit considerable selectivity in their ability to be polymerized via ionic mechanisms. The nature of the substituent (Y) will influence the electron density of the double bond and thus its ability to polymerize [19]. Substituents that are electron donating, such as alkoxy, phenyl, or alkyl, will increase the electron density of the double bond and allow for polymerization by cationic methods.



$$\overset{\delta^-}{CH_2}=\underset{\underset{Y^{\delta+}}{\uparrow}}{CH}$$

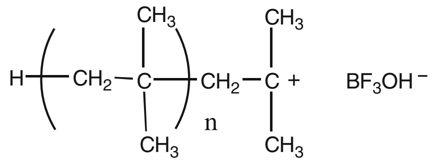


Fig. 15.13 Active cationic propagating species

Typical monomers that may be polymerized by cationic methods include styrene, isobutylene, and vinyl ethers. Unlike radical polymerizations, solvent polarity can influence the rate of polymerization. This is due to the presence of the counterion (see Fig. 15.13). For example, more polar solvents can increase the degree of separation between the growing end and the counterion during the propagation step, increasing the rate of propagation [20]. Chain transfer is the most common chain terminating reaction in cationic polymerization and can include transfer to monomer, solvent, and polymer. Termination by combination with the counterion can also occur in some systems. In some cases, cationic polymerization may be used to prepare stereoregular polymers. Although the exact mechanism is unclear, it is known that stereoregularity varies with initiator and solvent [21]. Lower temperatures also tend to favor more stereoregular polymers.

Anionic Polymerization. Anionic polymerization is similar to cationic polymerizations, except that the propagating species is anionic. Initiator types are typically alkali metals or their compounds. In the case of anionic polymerizations, electron withdrawing substituents, such as cyano, nitro, carboxyl, vinyl, and phenyl, facilitate polymerization by anionic means. Termination in anionic polymerization is generally by chain transfer. If the system is purified so that chain transfer is suppressed, the propagating species may remain active resulting in what is termed “living polymerization” [22]. This allows for the preparation of block copolymers, where one monomer is

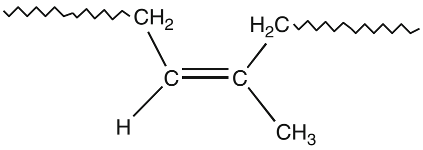


Fig. 15.14 *cis* 1,4 Polyisoprene

polymerized, followed by addition of a second monomer to the living end.

As with cationic polymerization the propagation rate is highly dependent on the solvent. Solvents that reduce the association between the growing chain end and the counterion result in faster rates of propagation [23].

In anionic polymerization of vinyl monomers (nondiene), low temperatures and polar solvents favor the preparation of syndiotactic polymers [24]. Nonpolar solvents tend to favor isotactic polymerization. In the case of diene monomers such as butadiene and isoprene, the use of lithium-based initiators in nonpolar solvents favors preparation of the *cis* 1,4 polymer (see Fig. 15.14).

Coordination Polymerization

In the 1950s, Karl Ziegler discovered a way to polymerize ethylene in a linear structure to produce high-density polyethylene (HDPE) using transition metal compounds and organometallic compounds. Using similar catalysts, Giulio Natta polymerized alpha-olefins, for example propene, with controlled stereoregularity. These catalyst systems are called Ziegler–Natta catalysts, and are widely used for the synthesis of a number of commodity plastics, such as high-density polyethylene and polypropylene [25]. Ziegler–Natta catalysts may be either insoluble (heterogeneous) or soluble (homogeneous) systems.

Ziegler–Natta Catalysts (Heterogeneous). These systems consist of a combination of a transition metal compound from groups IV to VIII and an organometallic compound of a group I to III metal [26]. The transition metal compound is called the catalyst and the organometallic compound the cocatalyst. Typically the catalyst is a halide or oxyhalide of titanium, chromium, vanadium, zirconium, or molybdenum. The cocatalyst is often an alkyl, aryl, or halide of aluminum, lithium, zinc, tin, cadmium, magnesium, or beryllium [27]. One of the most important catalyst systems is the titanium trihalides or tetrahalides combined with a trialkylaluminum compound.

The catalyst system is prepared by mixing the two compounds in the solvent, usually at low temperatures. Polymerization occurs at specific sites on the catalyst surface. There are several proposed mechanisms for polymerization, but the important aspect of both is that the polymerization occurs in coordination with the catalyst. The pi bond of the monomer complexes with the transition metal and is then inserted in between the transition metal and the carbon of the coordinated polymer chain. One of the proposed mechanisms is shown in Fig. 15.15. Isotactic polymers are generally formed with the insoluble catalysts. Syndiotactic polypropylene has been formed with both heterogeneous and homogeneous catalysts.

Metallocene Catalysts (Homogeneous Ziegler–Natta). Solid Ziegler–Natta catalysts suffered from several problems, including the presence of multiple polymerization sites on the catalyst surface and catalyst residue in the final polymer, requiring a secondary purification step [28]. The use of soluble catalysts offered an answer to some of these problems. These catalysts are composed of a metal atom (the active site), a cocatalayst, and a ligand system. Zirconium is the most commonly used metal although other metals such as Ti, Hf, Sc, and Th have been used. The most commonly used ligand is cyclopentadienyl. Methylalumoxane is typically the counterion. Figure 15.16 shows a proposed structure for methylalumoxane and a generalized metallocene structure.

The soluble catalysts can prepare polymers with very good stereospecificity and narrow molecular weight distributions, as a result of the uniformity of the active sites. In fact, these catalysts are often referred to as “single site catalysts” [29]. The polydispersity ($\overline{M_w}/\overline{M_n}$), which is a measure of the molecular weight distribution, is about 2–2.5 for the soluble catalysts and 5–6 for the heterogeneous Ziegler–Natta systems. By proper selection of the catalyst, syndiotactic, atactic and isotactic polypropylene, and higher alpha-olefins can be synthesized. In addition, it has been possible to prepare polypropylene with alternating blocks of isotactic and atactic chains. The resulting material exhibits elastomeric properties.

Polymerization Methods

While polymerization by ionic methods is usually performed in solution, free radical polymerizations can be performed in solution, bulk, suspension, or emulsion [30]. Each of these methods is described below.

Bulk Polymerization

Bulk polymerization is one of the simplest methods of polymerization. It is often used in the polymerization of step-growth polymers [31]. In these types of systems the viscosity remains low for a large portion of the reaction and heat transfer is easily controlled. Chain-growth polymers are more difficult to polymerize by this method due to the rapid and highly exothermic reactions. As the viscosity increases, thermal control becomes more difficult and may result in thermal runaway or localized hot spots.

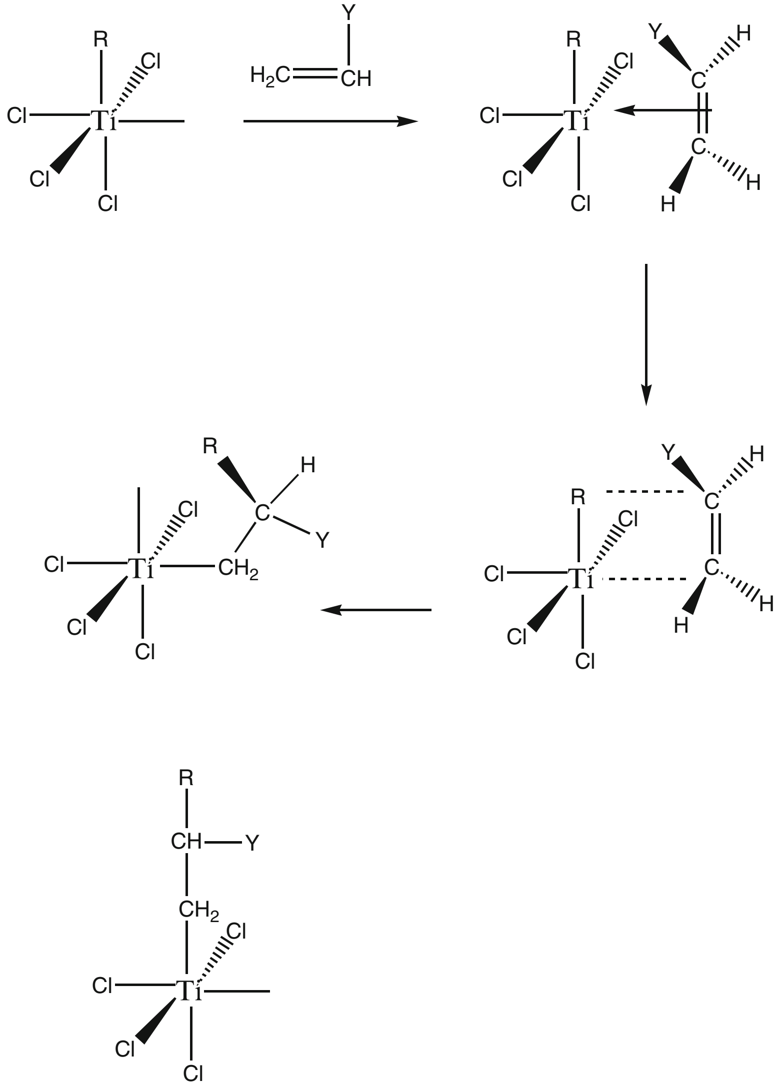
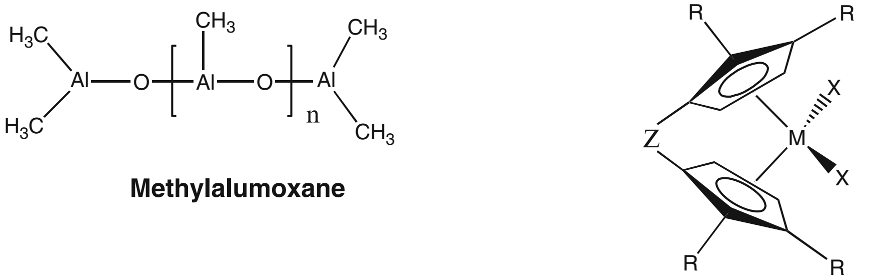


Fig. 15.15 Ziegler–Natta polymerization (monometallic mechanism)



Methylalumoxane

Metallocene

M = metal (Zr, Ti, or Hf)   X = Cl or alkyl

Z = optional bridging group   R = H or alkyl

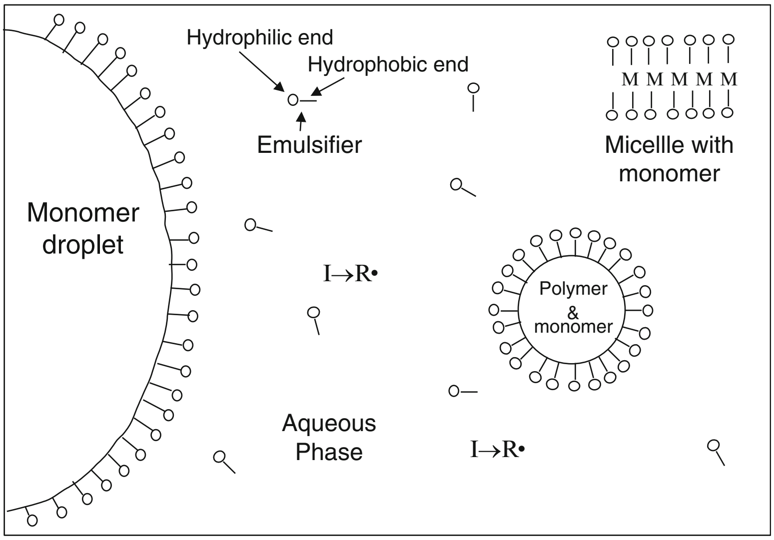
Fig. 15.16 Metallocene catalysts

Commercial use of bulk polymerization for vinyl polymers is rather limited for those reasons. Bulk polymerization may be either homogeneous or heterogeneous, as in the case where the polymer is insoluble in the monomer [32].

Solution Polymerization

Solution polymerization offers improved heat transfer over bulk polymerizations. Proper selection of the solvent is critical to avoid chain transfer reactions. Coupled with

Fig. 15.17 Emulsion polymerization



environmental concerns over organic solvents, the complete removal of solvents from the polymer also poses a potential problem. Recent work has been performed on the use of supercritical carbon dioxide as a solvent, which is easy to remove and poses less environmental concerns [33].

Suspension Polymerization

From an environmental standpoint, the use of water as a solvent is desirable. Unfortunately, many of the monomers of interest are insoluble in water, but suspension polymerization offers a way to utilize water. Suspension polymerization is performed by mechanically dispersing a monomer in an incompatible solvent, most often water. The system is heterogeneous and when polymerization is complete the polymer is collected as granular beads. This method is not suitable for tacky materials, such as elastomers, as the beads will tend to clump together [32].

Monomer droplets are suspended in the water through the use of agitation and stabilizers, such as methyl cellulose, gelatin, polyvinyl alcohol, and sodium polyacrylate [34]. Typical droplet sizes are 0.01–0.5 cm. A monomer soluble initiator is added to begin the polymerization. The kinetics of suspension polymerization are the same as for bulk polymerization, but suspension polymerization offers the advantage of good heat transfer. Polymers such as polystyrene, polyvinyl chloride, and polymethyl methacrylate are prepared by suspension polymerization.

Emulsion Polymerization

Superficially, emulsion polymerization resembles suspension polymerization, but there are a number of important differences. Water is used as the continuous phase and heat transfer is very good for both suspension and emulsion polymerization. In contrast to suspension polymerization, the polymer particles produced in emulsion polymerization are on the order of 0.1 μm in diameter [35]. Another important difference is the presence of an emulsifying agent or soap. At the beginning of polymerization the soap molecules aggregate together in a group of about 50–100 molecules to form what is called a micelle. Some of the monomer enters the micelles, but most of it is contained in monomer droplets. A water-soluble initiator is added, which migrates to the micelles as a result of their large surface to volume ratio and initiates polymerization primarily in the micelles. As polymerization continues, the micelles grow by addition of monomer from the water. The monomer droplets provide additional monomer to the aqueous phase as polymerization continues. Polymerization continues in the micelles until a second radical enters to terminate the reaction. As a result, very high molecular weight polymers may be synthesized [36]. Figure 15.17 shows an overall view of the emulsion polymerization process. Emulsion polymerization is a widely used technique, especially useful for making synthetic rubber, latex paints, and adhesives.

Copolymerization

In many cases the properties of a single type of polymer cannot meet the demands of a particular application. One approach to solving this problem is to combine two monomers into a single polymer through copolymerization. The properties of the resulting copolymer will then depend on the chemical nature of the monomers used and the microstructure of the chain. Monomers may be placed in the chain in a variety of ways, including random, block, and graft copolymers as depicted in Fig. 15.18.

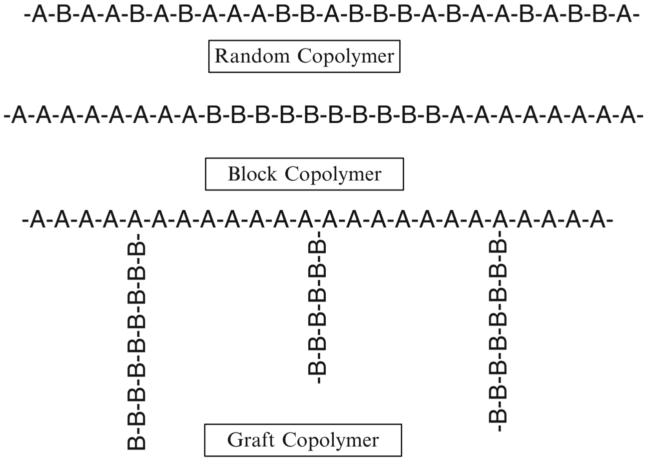


Fig. 15.18 Copolymer structures

Random Copolymerization

Copolymerization can occur through any of the chain-reaction polymerization mechanisms described above; however, the reactivity of a given monomer toward the second monomer can vary. Thus, not all combinations of monomers may be copolymerized. Each active end will exhibit different reactivity toward each monomer, which can be expressed as reactivity ratios, r_1 and r_2 [37]. These reactivity ratios (r_1 in this example) show the tendency of a given active end, for example $M_1{}^*$, to add its own monomer (M_1) over the other monomer (M_2). The copolymer composition at any instant can be determined by the composition of the feedstock and the reactivity ratios by

$$F_1 = \frac{r_1 f_1^2 + f_1 f_2}{r_1 f_1^2 + 2 f_1 f_2 + r_2 f_2^2}$$

where F_1 is the mole fraction of monomer 1 in the copolymer, f_1 and f_2 are the mole fractions of monomer 1 and 2 in the feedstock, and r_1 and r_2 are the reactivity ratios. In general, the copolymer composition will not be the same as the feed composition.

The properties of random copolymers are often a weighted average of the two polymers. For example, the glass transition temperature (T_g) of a single-phase copolymer typically falls somewhere in between the T_g's of the two homopolymers. This can be estimated using [38]

$$a_1 c_1 (T_g - T_{g1}) + a_2 c_2 (T_g - T_{g2}) = 0$$

where T_{g1} and T_{g2} are the glass transition temperatures for the pure homopolymers, a_1 and a_2 depend on the monomer type, and c_1 and c_2 are the weight fractions of monomer 1 and 2. For crystalline polymers, the degree of crystallinity and melting point decrease as the second monomer is added [39].

Block and Graft Copolymers

As mentioned above, the ability to have "living polymerizations" offers the potential to make block copolymers. In the preparation of a block copolymer the sequence of addition can be important to ensure that the second monomer is capable of adding to the living end. An example is the formation of a polystyrene—polymethyl methacrylate block copolymer [40]. In this case polystyrene is polymerized first, followed by addition of the methyl methacrylate. The block copolymer could not be formed if methyl methacrylate were polymerized first, as styrene will not add to the methyl methacrylate living end. An important example of block copolymerization is the synthesis of ABA triblock copolymers from styrene and diene monomers to form thermoplastic elastomers. They may be prepared by sequential polymerization from one end or from the middle using a difunctional initiator. Block copolymers may also be prepared with polymers containing functionalized end groups that are linked together.

There are three basic methods to produce graft copolymers [41]. In the first method a monomer is polymerized in the presence of a polymer and chain transfer provides the branching. Initiator, monomer, and polymer are combined to prepare the copolymer. The initiator may function by polymerizing the monomer, which then reacts with the polymer to form the graft, or by forming a reactive site on the polymer, which then polymerizes the monomer onto the polymer.

The second method of forming a graft copolymer is to polymerize a monomer with a polymer containing a reactive functional group (or sites that may be activated). Irradiation is commonly used to form active sites for graft copolymerization, but other methods may also be used. When irradiation is used the grafting mechanism is free radical in nature. In cases where the monomer is present when the polymer is irradiated homopolymerization may also occur. If homopolymerization is a concern, then the polymer may be irradiated in presence of air (or oxygen) to form hydroperoxide groups. The pretreated polymer is then mixed with monomer and heated to initiate the grafting reaction by decomposition of the peroxide groups.

The third method of preparing a graft copolymer is to combine two polymers with functional groups that can react together. An example is the reaction of oxazoline-substituted polystyrene with polymers containing functional groups such as alcohols, amines, and carboxylic acids.

Unlike the random copolymers, block and graft copolymers separate into two phases, with each phase exhibiting its own T_g (or T_m) [42]. The modulus-temperature behavior of a series of triblock copolymers prepared from styrene and butadiene (SBS) is shown in Fig. 15.19. With increasing styrene content the modulus in the plateau region (between the two glass transition temperatures) increases.

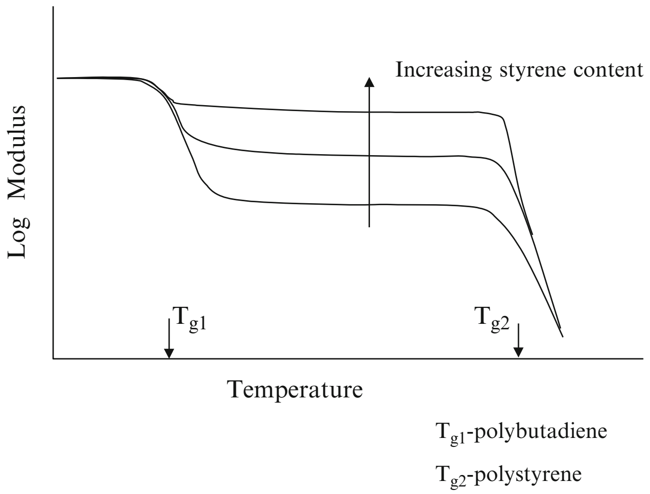


Fig. 15.19 Modulus-temperature behavior of a triblock copolymer

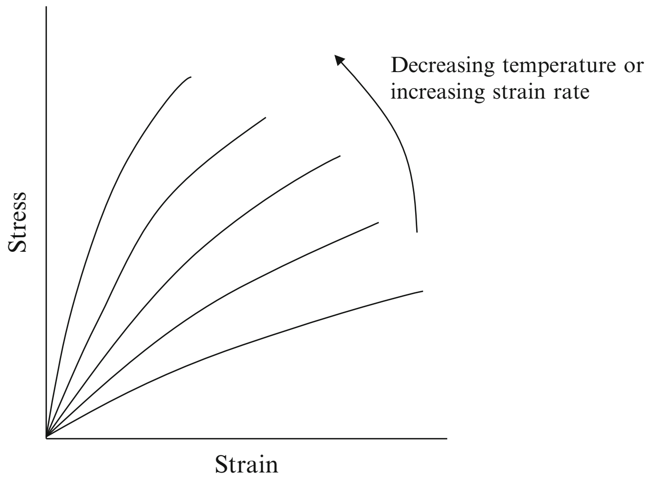


Fig. 15.20 Viscoelastic behavior of polymers as shown by stress–strain curves

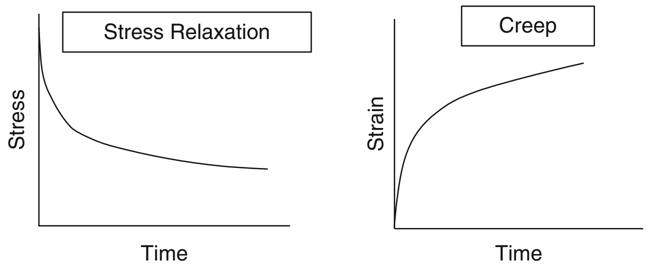


Fig. 15.21 Creep and stress relaxation

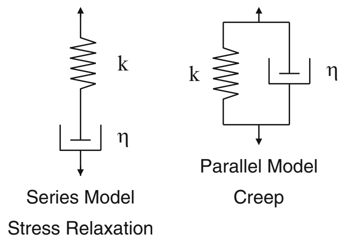


Fig. 15.22 Spring and dashpot models for stress relaxation and creep

Mixing of the two phases may shift the values of the two transition temperatures.

Mechanical Properties

Many factors influence the mechanical behavior of polymers including molecular weight, polymer type, processing conditions, and test procedures. Modulus values may be determined from a standard tensile test with a given rate of crosshead separation. The elastic or Young's modulus, E, is found from the initial part of the curve. Unlike many other materials, polymeric material behavior is affected by factors, such as test temperature and rates. When a product or material is used or tested at temperatures near the glass transition temperature, this becomes particularly important as the properties show dramatic changes (orders of magnitude) in properties (see Fig. 15.8).

Viscoelasticity

Polymer properties may exhibit considerable time-dependent behavior, depending on the polymer type and test conditions. The material appears rigid with an increase in testing rate or decrease in temperature and appears softer with an increase in temperature or decrease in rate. Viscoelastic response of a polymer to changes in testing rate or temperature is shown in Fig. 15.20. Long-term effects such as stress relaxation or creep [43] are the result of this time-dependent behavior. Creep and stress relaxation are shown in Fig. 15.21. A specimen held under a fixed load will continue to elongate with time, a process called creep. A spring and dashpot in parallel is often used as a simple model to predict the creep behavior (see Fig. 15.22). This model predicts the time-dependent strain to be

$$\varepsilon(t) = \varepsilon_o \mathrm{e}^{-t/\tau}$$

where τ is the characteristic relaxation time (η/k).

Stress relaxation is the decrease in load of a material held at a fixed displacement. Figure 15.22 shows the spring and dashpot in series that can be used to model the stress relaxation behavior. Using this model one obtains the following equation for the stress behavior.

$$\sigma(t) = \sigma_o \mathrm{e}^{-t/\tau}$$

These models are useful for understanding the general concept of viscoelasticity, but are typically unable to

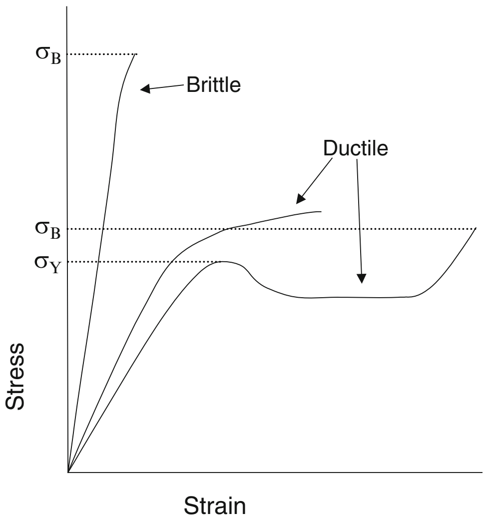


Fig. 15.23 Types of stress–strain behavior

accurately model the time-dependent behavior. More accurate prediction can be obtained by using models with more elements.

If stress relaxation curves are obtained at a number of different temperatures, it is found that these curves can be superimposed by horizontal shifts to produce what is called a “master curve” [44]. This concept of time–temperature equivalence is very important to understanding and predicting polymer behavior. As an example, a polymer at very low temperatures will behave as if it were tested at higher temperatures at much higher testing rates. This principle can be applied to predict material behavior under other rates or temperatures that are difficult to test experimentally by the use of shift factors (a_T).

$$\ln a_T = \ln\left(\frac{t}{t_o}\right) = -\frac{17.44(T - T_g)}{51.6 + T - T_g}$$

where T_g is the glass transition temperature of the polymer. This has practical applications where one may be interested in the mechanical properties of a material at low temperatures under high rate (perhaps impact rate) conditions. The design engineer should recognize that material properties measured at room temperature may not adequately reflect the material behavior at much lower temperatures or higher rates.

Failure Behavior

Part design requires the avoidance of failure without overdesign of the part, which leads to increased part weight and cost. Failure behavior depends on material type, service temperature, and rates. A tensile stress strain test can be used to gather some information on material strength and behavior, for example, stress (or strain) at break, σ_B.

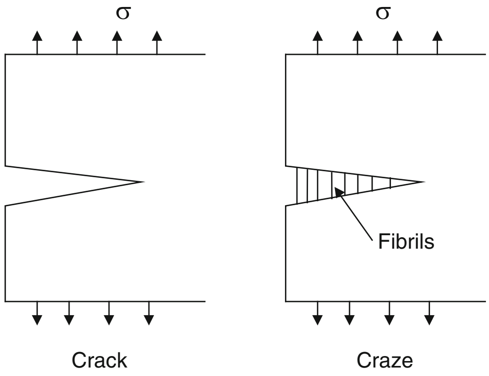


Fig. 15.24 Crazes and cracks

Figure 15.23 shows some different types of failure behavior. Brittle failure is when a materials fails at low elongations (1% strain or less) as exhibited by general purpose styrene and acrylics [45]. Failure initiates at a defect where stresses are concentrated. Once formed, a crack will continue to grow from the stress concentrations at the crack tip. Amorphous polymers can also show will what appear to be cracks, but are load bearing with fibrils of material bridging the two surfaces (Fig. 15.24). These structures are termed “crazes.” Crazing enhances the toughness of a material and is a form of yielding.

Polymers also exhibit ductile failure by yielding of the polymer or slip of the molecular chains past one another. This is found by a maximum in the tensile stress–strain test or the yield point, σ_Y (see Fig. 15.23). With further extension the sample may show lateral contraction, called necking [46]. In the necked regions, molecules are oriented in the direction of deformation, resulting in increased stiffness in the necked region. As a result of this localized stiffness increase, material adjacent to the neck is deformed and the neck region continues to grow, known as cold-drawing (see Fig. 15.25). Elongations of several hundred percent can occur with cold-drawing.

As might be expected, temperature will influence the behavior of the material. The effect of temperature on modulus has been discussed above. Very different behavior may be seen in a single polymer simply by changing the temperature [47]. Beginning at temperatures well below the T_g, the stress–strain behavior will show very low elongations, with no yield point. As the temperature is raised, the material will exhibit a yield point and elongation increases. The yield point is typically seen at temperatures near the glass transition temperature of the polymer. The higher the testing rate, the higher the temperature must be for yielding to occur.

A material may fail at stresses well below the single-cycle failure stress found in a typical tensile test under repeated cyclic stress, a process called fatigue [48]. Fatigue behavior is usually plotted as the maximum stress vs. the number of

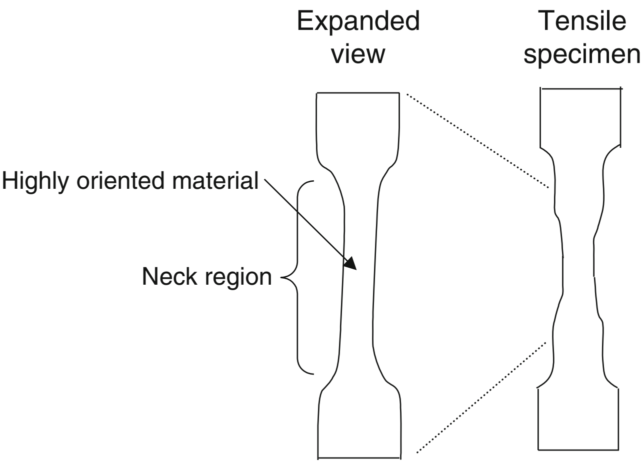


Fig. 15.25 Cold-drawing

cycles to failure. The exact conditions of the fatigue test can be varied and are specified by the service requirements. Type of loading (compression, tension, shear), load or displacement control, frequency, waveform (typically sinusoidal), and ratio of maximum to minimum stress may all be varied. The fatigue life is also influenced by the presence or absence of cracks and thermal effects.

Part II: Commercial Plastic Materials

Classes of Families of Commercial Plastics

Commercial plastic materials may conveniently be divided into five major classes. Commodity thermoplastics are families, which are produced in volumes of a billion pounds or more per year. Engineering thermoplastics are generally designed to withstand use at higher temperatures. Specialty thermoplastics are designed for a greater variety of performance and applications. Thermoplastic elastomers can be processed like thermoplastics, and give many of the properties of conventional thermoset vulcanized rubbers. Thermoset plastics are permanently cross-linked during processing, generally to give higher performance than would be possible with linear thermoplastics.

Commodity Thermoplastics

Seven families of thermoplastics exceed a billion pounds per year in the United States: polyethylenes, polypropylene, PVC, polystyrene and its copolymers, polyethylene terephthalate, acrylonitrile-butadiene-styrene (ABS), and nylon.

Table 15.2 Major members of the polyethylene family: typical properties

Polyethylene	HDPE	LLDPE	LDPE
Density	0.96	0.93	0.92
Melting point (°C)	130	124	108
Modulus (MPa)[a]	1,400	450	180
Tensile strength (MPa)[a]	26	14	12

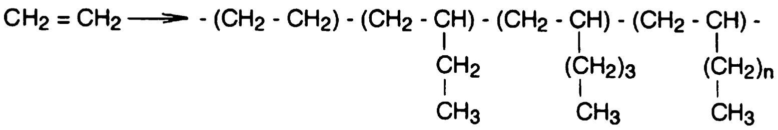
[a] To convert to psi, multiply by 145

Polyethylenes. The major members of the polyethylene family are distinguished primarily by the amount of side-branching, which decreases regularity, crystallinity, density, melting point, modulus (rigidity), and strength (Table 15.2).

High-Density Polyethylene (HDPE). Polymerization of ethylene to polyethylenes is most often carried out at low temperature and pressure, using either the Ziegler aluminum triethyl plus titanium tetrachloride catalyst system, the Phillips chromic oxide plus silica plus alumina system, or more recently the newer metallocene single-site catalyst systems. This produces a regular, linear polymer which is 50–90% crystalline, making it rigid and strong. Its largest use is blow-molded bottles for milk, water, juice, noncarbonated drinks, laundry detergent, antifreeze, and motor oil. Other large uses are large molded containers, pipe and conduit, and shopping bags. For detergent bottles, environmental stress-crack resistance is increased by increasing molecular weight, adding ethylene/propylene rubber (EPR), or copolymerization with a small amount (1–2%) of L-butene to reduce crystallinity. For auto gasoline tanks, impermeability may be increased by dispersing nylon in the polyethylene. For extreme abrasion resistance in materials-handling machinery, and in hip and knee bone replacement, Ultra-High Molecular Weight (UHMWPE: MW 3–6 × 10^6) is outstanding.

$$n\mathrm{CH_2{=}CH_2} \rightarrow -(\mathrm{CH_2{-}CH_2})_n-$$

Low-Density Polyethylene (LDPE). Polyethylene was originally made by polymerization at high temperature and pressure; during the polymerization, side-reactions produced several percent of short-chain branching (ethyl and butyl groups), which reduced regularity/crystallinity (e.g., to 30–40%) and gave a stiffly flexible plastic material; chain transfer reactions also produced some long-chain branching, which gave pseudoplastic melt processability, important in the production of extruded blown film. The largest use is in packaging film; the second largest is for coating and laminating on cardboard, paper, and foil.

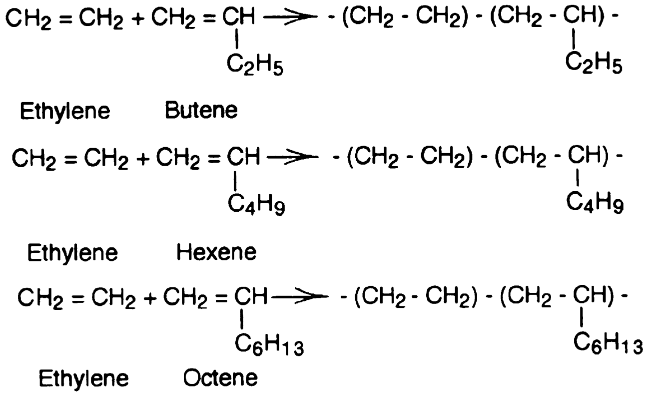


$$\mathrm{CH_2{=}CH_2} \longrightarrow -(\mathrm{CH_2{-}CH_2})-(\mathrm{CH_2}-\underset{\underset{\mathrm{CH_3}}{|}}{\underset{\mathrm{CH_2}}{\underset{|}{\mathrm{CH}}}})-(\mathrm{CH_2}-\underset{\underset{\mathrm{CH_3}}{|}}{\underset{(\mathrm{CH_2})_3}{\underset{|}{\mathrm{CH}}}})-(\mathrm{CH_2}-\underset{\underset{\mathrm{CH_3}}{|}}{\underset{(\mathrm{CH_2})_n}{\underset{|}{\mathrm{CH}}}})-$$

Table 15.3 Typical properties of polypropylenes

Type	General purpose	High impact	40 % Talc	40 % Glass fiber
Modulus (MPa)	1,450	1,100	3,100	7,000
Tensile strength (MPa)	36		32	70
Ultimate elongation (%)	500		10	2
Notched Izod impact strength (J/cm)	1.1	5.6	1.3	3.9
Heat deflection temperature (°C at 1.82 MPa)	57	53	79	149

Linear Low-Density Polyethylene (LLDPE). Ziegler and metallocene polyethylene can be modified by copolymerization with several percent (e.g., 10%) of butene, hexene, or octene, to reduce the regularity/crystallinity and modulus, and thus produce LDPE at low temperature and pressure. The resulting LLDPE is stronger than the original LDPE, and therefore preferred for packaging film and trashcan liners. Lack of long-chain branching makes it difficult for extrusion-blown film, so processors add enough LDPE to improve processability.



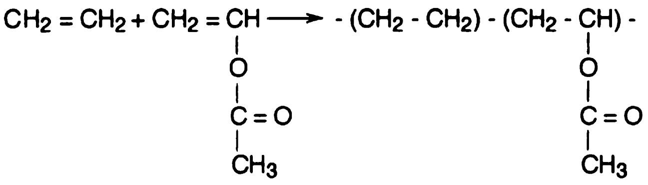
$CH_2=CH_2 + CH_2=\underset{C_2H_5}{CH} \longrightarrow -(CH_2-CH_2)-(CH_2-\underset{C_2H_5}{CH})-$

$CH_2=CH_2 + CH_2=\underset{C_4H_9}{CH} \longrightarrow -(CH_2-CH_2)-(CH_2-\underset{C_4H_9}{CH})-$

$CH_2=CH_2 + CH_2=\underset{C_6H_{13}}{CH} \longrightarrow -(CH_2-CH_2)-(CH_2-\underset{C_6H_{13}}{CH})-$

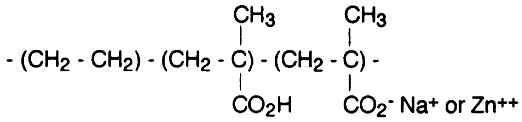
Very Low Density (VLDPE) and Ultra Low Density (ULDPE) Polyethylenes. These are made by copolymerization with increasing amounts of comonomers, especially 1-octene, reducing the regularity/crystallinity (density 0.91–0.86) down toward EPR. These are soft and flexible enough to compete with plasticized polyvinyl chloride and thermoplastic elastomers in some applications.

Ethylene/Vinyl Acetate Copolymers (EVA). LDPE is easily internally plasticized by copolymerization with increasing amounts (2–20%) of vinyl acetate comonomer. This reduces the regularity/crystallinity (melting points 108–70°C), resulting in increasing softness/flexibility (modulus 138–19 MPa), transparency, and polar adhesion, again competing with plasticized PVC and thermoplastic elastomers, particularly in packaging film, coatings, and adhesives.



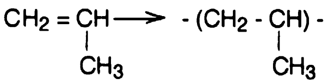
$CH_2=CH_2 + CH_2=\underset{\substack{|\\O\\|\\C=O\\|\\CH_3}}{CH} \longrightarrow -(CH_2-CH_2)-(CH_2-\underset{\substack{|\\O\\|\\C=O\\|\\CH_3}}{CH})-$

Ionomers are generally ethylene copolymers with 5–10% of methacrylic acid, partially-neutralized by sodium or zinc. This reduces the regularity/crystallinity, increasing the softness/flexibility (modulus 250 MPa) and transparency. The ionic groups increase strength (20 MPa) and polar adhesion, again competing with plasticized PVC and thermoplastic elastomers, particularly in packaging films and sporting goods.



$-(CH_2-CH_2)-(CH_2-\overset{CH_3}{\underset{CO_2H}{C}})-(CH_2-\overset{CH_3}{\underset{CO_2^-\ Na^+\ \text{or}\ Zn^{++}}{C}})-$

Polypropylene. Polymerization of propylene to polypropylene (Table 15.3) produces an asymmetric carbon atom, which can be either right-handed or left-handed. Peroxide initiation of free radical polymerization produces a random ("atactic") mixture of right- and left-handed carbon atoms, which is too irregular to crystallize, so the resulting polymer is a soft gummy rubber. Ziegler triethyl aluminum plus titanium chloride catalysis of polymerization pre-orients each monomer unit before inserting it into the polymer chain, and produces an all-right-handed or all-left-handed ("isotactic") polymer, which is very regular and therefore able to crystallize (e.g., 60%). This is a high-melting (165°C), rigid, strong, low-cost commodity plastic, very competitive with more costly polystyrene and ABS. Another unique quality is its ability to form an "integral hinge," so that a rigid packaging container plus lid can be molded as a single part.

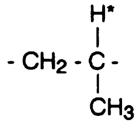


$CH_2=\underset{CH_3}{CH} \longrightarrow -(CH_2-\underset{CH_3}{CH})-$

Table 15.4 Typical properties of rigid PVC

Type	General purpose	High impact	30 % Glass fiber	Chlorinated PVC
Modulus (MPa)	3,300	2,900	8,300	2,800
Tensile strength (MPa)	48	43	90	56
Ultimate elongation (%)	145	130	2	35
Notched Izod impact strength (J/cm)	0.32	8.56	0.70	0.91
Heat deflection temperature (°C at 1.82 MPa)	67	66	76	103

Polypropylene has two major weaknesses: (1) Methyl branching leaves tertiary hydrogen atoms, which are sensitive to thermal oxidative aging.

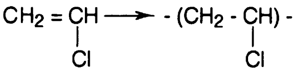


$$-CH_2-\overset{H^*}{\underset{CH_3}{C}}-$$

(2) Bulky methyl groups cause steric hindrance and stiffen the polymer molecule; while this increases modulus, it also causes embrittlement below room temperature. The polypropylene industry has successfully retarded aging by adding phenolic and aliphatic sulfide antioxidants; and it has reduced low-temperature embrittlement by adding or grafting ethylene–propylene–diene (EPDM) rubber. A third common improvement is the addition of fillers to increase rigidity, making it even more competitive with polystyrene.

With these problems solved, polypropylene use has been growing rapidly. Extrusion and stretch orientation produces excellent synthetic fiber for carpeting and synthetic turf, and excellent film for packaging. Other large uses are rigid packaging (replacing polystyrene), automotive parts (replacing ABS), and a great variety of injection-molded consumer products.

Polyvinyl Chloride (Table 15.4). This is the most versatile of the commercial thermoplastic polymers. It is used mainly for rigid and flexible plastics, for rubberlike products, for coatings on steel, cloth, and paper, and in smaller amounts for specialty fibers. It is processed mainly by extrusion and calendering, and in smaller amounts by injection, compression, and blow molding, thermoforming, rotomolding, plastisol/organosol technology, casting, pouring, and foaming. This extreme versatility comes primarily from the use of plasticizers, aided by a host of other types of additives.



$$CH_2=\underset{Cl}{CH} \longrightarrow -(CH_2-\underset{Cl}{CH})-$$

Table 15.5 Typical properties of flexible PVC

Plasticizer	Dioctyl phthalate			Dioctyl adipate		
Parts/hundred of resin	30	50	75	30	50	70
Shore A hardness	95	79	62	90	74	60
Tensile strength (MPa)	31	21	14	27	20	14
Ultimate elongation (%)	220	300	400	270	395	410
Flex temperature (°C)	+20	0	−26	−15	−48	−75

Vinyl chloride is polymerized primarily by peroxide-initiated polymerization in aqueous suspension, producing a fine porous powder which is easy to blend with compounding ingredients. A smaller amount is polymerized in emulsion and spray-dried for plastisols and organosols and an even smaller amount is copolymerized with vinyl acetate in organic solution, to produce a uniform copolymer which precipitates at a constant composition and molecular weight.

Rigid PVC is polymerized at 55–70°C and then compounded with an organotin stabilizer, acrylic processing aids, lubricants, and optionally with a rubbery impact modifier.

The polymer has high modulus, strength, and resistance to creep, weathering, and chemicals; and compounding with rubbery impact modifiers produces high resistance to brittle failure. It is mainly extruded into pipe, house-siding, and window and door frames, while smaller amounts are injection molded into pipe fittings and blow-molded into bottles for detergents and other chemicals.

Flexible PVC (Table 15.5) is made by polymerizing at 40–55°C and then compounding with 20–80 PHR (parts per hundred of resin) of dioctyl phthalate and/or other monomeric liquid plasticizers (e.g., dioctyl adipate for low-temperature flexibility, oligomeric polyesters for permanence, organic phosphates for flame-retardance), plus a synergistic stabilizer system usually composed of barium or calcium soap, zinc soap, epoxidized fatty ester, and organic phosphite.

This flexible-rubbery material is most often calendered into film and sheet for clothing, luggage, raincoats, upholstery, and flooring, or extruded into garden hose, wire and cable insulation, and medical tubing.

For plastisol/organosol technology, high molecular weight PVC powder is slurried in liquid plasticizer, and poured, dip- or spread-coated onto metal, fabric, or paper to produce dishwasher racks, upholstery, and wallpaper, respectively.

The excellent combustion resistance is one of the most important properties of PVC. The good flame resistance results from the releasing of HCl gas during the combustion of PVC, which is denser than air thus restricts air from reaching the flame. Therefore, PVC occupies around 60%

Table 15.6 Typical properties of polystyrenes

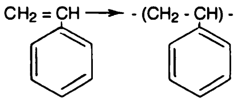
		Impact styrene		
Polymer	Polystyrene	Moderate	Very high	Styrene acrylonitrile
Modulus (MPa)	2,900	1,700	1,100	3,500
Tensile strength (MPa)	46	21	13	73
Notched Izod impact strength (J/cm)	0.17	1.28	3.70	0.26
Heat deflection temperature (°C at 1.82 MPa)	83	81	81	100

of the flame resistant wire and cable plastic market. The good weatherability, easy processing, and high combustion resistance contribute to PVC occupying around 70% of the construction market for plastics. Also, PVC's outstanding capabilities in the plastisol process allow them to serve 25% of the coating market, such as the metal structure coating for dishwasher racks and flooring for wet kitchen and bathroom areas. PVC also has excellent resistance to acids, bases, oils, and other hydrocarbons. This property permits PVC to be widely used in parts for the chemical processing industries. Again, because of the excellent combustion properties, PVC has a large share of the market in electrical enclosures, which requires the materials have the flame resistance with UL 94V-0 ratings [49].

Regardless of their excellent combustion resistance, which can be used to protect people to avoid the loss of life in a fire, the environmentalist groups would like to eliminate the halogen products. Although European market has increasing demand for halogen-free flame-retardant, the fires caused by television set have increased in Europe [50].

There are many suppliers of PVC compounds. The major suppliers include Shin-Etsu, Formosa Plastics Corp., Solvay, Atofina, Georgia Gulf, LG Chem etc.

Polystyrene and Its Copolymers. Polystyrene (Table 15.6) is made by continuous bulk polymerization, initiated by peroxides and heat, with temperature rising continuously to keep the system molten, and optionally adding solvent ("solution polymerization") to facilitate the process. It was the first commodity thermoplastic, combining easy melt processability, rigidity, and glass-like transparency for major applications in packaging, toys, and housewares. When swollen with 10% pentane and heated, it expands to rigid closed-cell foams with densities as low as 0.01 (expanded polystyrene or EPS), which are popular for packaging hot and cold foods, delicate instruments and appliances.

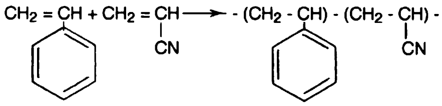


The major weaknesses of polystyrene are brittleness and softening in hot water. Brittleness is remedied by dissolving 2–10% of rubber in styrene monomer before polymerization, producing "high-impact styrene" (HIPS), in which 10-μm rubber particles improve impact strength by an order of magnitude, with some sacrifice of other mechanical properties and transparency; this accounts for more than half of the total "polystyrene" market.

Table 15.7 Typical properties of acrylonitrile-butadiene-styrene (ABS)

Type	Moderate impact	Very high impact
Modulus (MPa)	2,500	2,000
Tensile strength (MPa)	46	38
Notched Izod impact strength (J/cm)	3.5	9.0
Heat deflection temperature (°C at 1.82 MPa)	94	88

Heat resistance is improved by copolymerization with 15–30% of acrylonitrile, producing styrene/acrylonitrile (SAN). Polarity and hydrogen bonding stiffen the polymer molecule, improving mechanical properties and especially resistance to hot water. This improvement is useful in household products, autos, and appliances.



Acrylonitrile/Butadiene/Styrene (ABS). The benefits of impact styrene and SAN are combined in ABS (Table 15.7). Typically, 15–30% acrylonitrile and 45–75% styrene are copolymerized in the presence of 5–30% of polybutadiene rubber, producing some graft terpolymer, and dispersing 1-μm rubber domains in an SAN matrix.

This balance of properties has found particular usefulness in appliances, autos, electronic equipment cabinets, and rain-waste-vent (DWV) pipe.

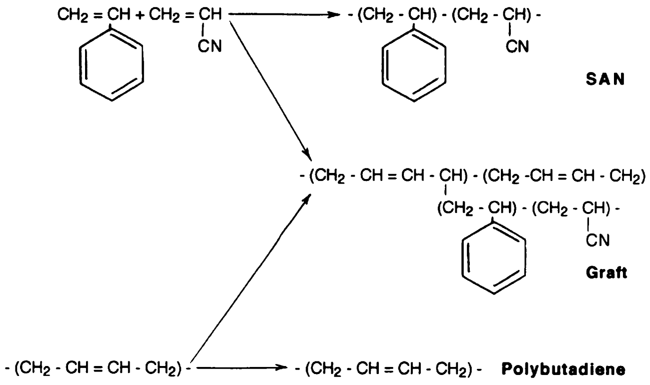


Table 15.8 Typical properties of PET

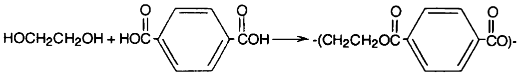
Form	Fiber	Film	Bottles	Reinforced molding 15 % Glass fiber	55 % Glass fiber
Modulus (MPa)		3,600	3,200	5,900	18,000
Tensile strength (MPa)	550	200	120	110	200
Ultimate elongation (%)	29	125		2	1
Notched Izod impact strength (J/cm)				0.64	1.07
Heat deflection temperature (°C at 1.82 MPa)				210	229

Other styrene-based copolymers. The commercial success of ABS has led to the development of other styrene-based copolymers to overcome the disadvantages of ABS, such as poor weathering resistance and lack of transparency. For example, methacrylate-butadiene styrene (MBS) is another typical styrene-based copolymer. In MBS, methyl methacrylate and styrene are grafted onto the polybutadiene backbone. This structure results in two advantages over ABS: higher clarity and better resistance to the discoloration in the presence of ultraviolet light. The disadvantages of MBS over ABS are lower tensile strength and heat deflection temperatures (HDT). The better aging of MBS comes from the inhibition or retardation of methyl methacrylate on the oxidative attack at the double bonds in the polybutadiene backbone. The further improved weathering resistance of styrene polymers has been developed by replacing butadiene rubber with an ethylene-propylene terpolymer rubber of the EPDM type. EPDM rubber has a small amount of a diene monomer in the polymerization formulation. In addition to the higher clarity of MBS, transparent polystyrene polymers with high impact strength are produced by blending polystyrene and styrene-butadiene-styrene (SBS) block copolymers. The SBS block copolymers can be either star or linear shapes. The optical properties depend on the ratio of polystyrene to SBS, the presence of additives, and the processing conditions etc. 1,2-Polybutadiene can also be blended with polystyrene to achieve high impact materials with a degree of transparency. The careful control of processing conditions has to be done to avoid excess cross-linking [51].

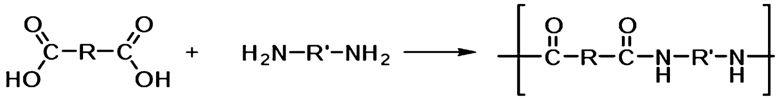
Styrenic poly(phenylene oxides) (PPO) is another important styrene-based engineering material. Its production is about one-tenth that for ABS. This material has useful characteristics similar to polystyrene including good dimensional stability, low water absorption and excellent dielectric properties. In addition, it has higher heat distortion temperatures than PS. Some grades have heat distortion temperatures up to 160°C. More interestingly, it has lower price than heat-resistant materials, such as polycarbonate and polysulphones. Styrenic PPO has applications in the areas, such as autos, electronics, pumps and other plumbing applications [51].

Polyethylene Terephthalate. PET (Table 15.8) is produced by continuous melt condensation polymerization of ethylene glycol plus terephthalic acid, followed by solid-state finish to reach high molecular weight and high purity. It was originally developed for synthetic fibers, replacing cotton as the leading textile fiber worldwide. Later it became popular in packaging film (e.g., boil-in-bag) and magnetic tape because of its high strength, cling, melting point, impermeability, and clarity. Finally, it became the basic material for blow-molding bottles for carbonated beverages, after which it proved completely recyclable into bulk fiber for bedding, furniture, and clothing. More specialized engineering grades for solid molding applications are generally reinforced by short glass fibers.

Although PET has achieved commercial significance for the bottles, films, and fibers etc., it is also being increasingly modified by copolymerizing with isophthalic acid (IPA), cyclohexanedimenthanol (CHDM) or other components to improve the toughness, clarity, water and weathering resistance, and barrier properties. Such commercial products include Kodar, Vestan, and A-Tell.



Nylon. Nylons (Table 15.9) are made of repeating amide units and are frequently referred to as polyamide (PA). The general reaction is the condensation polymerization of diamine and dicarboxylic acid. They can also be polymerized by the self-condensation of an amino acid or ring-opening polymerization of a lactam. A nylon made from diamine and dicarboxylic acid is called Nylon *x*,*y*, where *x* and *y* represent the number of carbons in the diamine chain and dicarboxylic acid chain, respectively. A nylon made from either an amino acid or a lactam is designated Nylon *x*, where *x* is the carbon numbers in the repeating units [52].



$$HO-\overset{O}{\overset{\|}{C}}-R-\overset{O}{\overset{\|}{C}}-OH + H_2N-R'-NH_2 \longrightarrow \left[-\overset{O}{\overset{\|}{C}}-R-\overset{O}{\overset{\|}{C}}-\underset{H}{N}-R'-\underset{H}{N}-\right]$$

Table 15.9 Typical properties of nylon [53]

Type	Nylon 11	Nylon 12	Nylon 46	Nylon 6	Nylon 610	Nylon 612	Nylon 66
Flexural modulus (MPa)	152–1,570	110–1,760	350–11,600	1,560–3,130	1,820–2,080	300–8,490	1,500–3,750
Tensile strength (MPa)	30–70	25–59	30–214	37–98	47–66	26–173	42–91
Tensile elongation (%)	2–56	0.60–200	0.6–53	0.40–25	2.4–100	2.0–32	0.7–19
Rockwell hardness	77–109	79–115	90–100	110–123	112–116	113–118	79–122
Notched Izod impact (J/m)	16–110	16–160	40–100	10–98	35–50	29–89	10–95
DTUL at 1.8 MPa (°C)	39–177	44–167	155–291	44–215	54–68	51–207	56–256

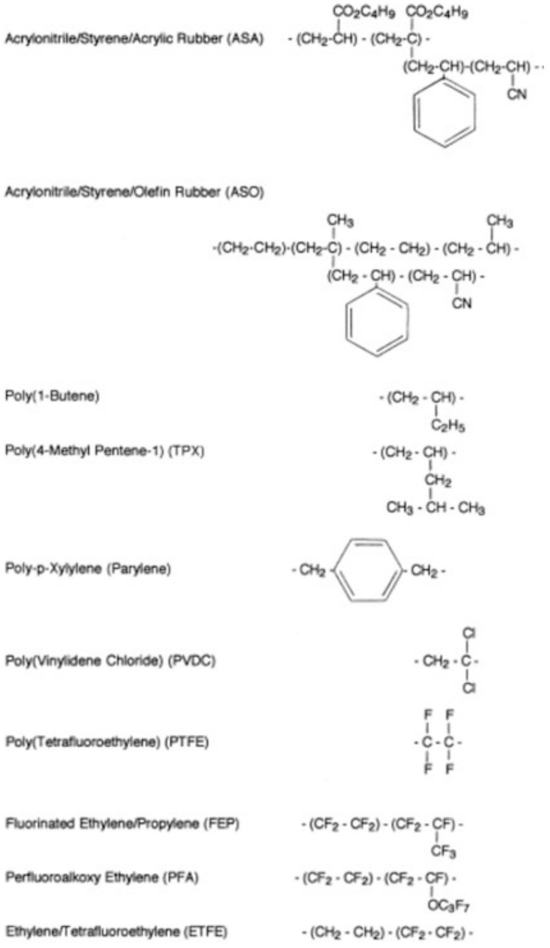
Nylon contains the polar –CONH– groups spaced at regular intervals, which give rise to nylon crystalline structures with a high intermolecular attraction. The aliphatic chain segments give nylon flexibility in the amorphous regions. This combination leads to nylon being characterized by high toughness, modulus, strength, abrasion resistance, flexibility, rigidity, and good hydrocarbon resistance. Because the polar –CONH– group absorbs water, they are generally not suitable for applications in humid conditions but they are quite good electrical insulators at room temperature in dry conditions and low frequencies.

There are a variety of commercial nylons available. In the nylon family, Nylon 6,6 is the most common commercial grade of nylon and Nylon 6 is the most common commercial grade of molded nylon. Nylons are most often used in the form as fibers, which account for almost 90% of the nylon production in the world. Nylon 6,6 and nylon 6 account for almost all of the polyamide in fiber applications. Nylon fibers are used in many applications including fabrics, bridal veils, carpets, hose and belt reinforcement, musical strings, tire cord and rope. The other 10% of nylon production goes into plastics applications. The largest application of solid nylon is for mechanical parts, such as machine screws, gears, cams, bearings, bushes, and other low- to medium-stress components previously cast in metal. Other applications for nylons include sterilizable nylon moldings in medicine and pharmacy, fishing line, package film, and brush bristles. Because of their excellent properties, nylons have found increasing applications in specialty purposes in spite of their relatively higher cost compared to the general purpose materials, such as polystyrene and polyethylene. Glass fiber reinforced nylons occupy the largest market among the glass filled variety of thermoplastics. Glass fiber reinforced nylons have high rigidity, heat deflection temperature and creep resistance, low coefficient of friction, good low-frequency electric insulation properties and nonmagnetic characteristics. All of these excellent properties enable them to be widely used in a range of applications including housings, casings, car components, domestic appliances, and the aerospace and telecommunications fields [54].

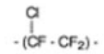
In summary, commodity thermoplastics are manufactured readily at low cost, and offer a combination of processability, mechanical, thermal, optical, and chemical properties that are useful in a wide range of mass markets and products.

Engineering and Specialty Thermoplastics

Many present and future applications of thermoplastics make greater demands for higher properties and especially combinations of properties, than are available from the commodity materials. To satisfy these demands, organic polymer chemists and chemical engineers have developed and commercialized over four dozen major types of polymers, offering many improved properties to meet these demands. They may be listed as follows, and then compared in their abilities to satisfy these requirements.

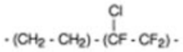


Chlorotrifluoroethylene (CTFE)



$-(\overset{\overset{Cl}{|}}{C}F-CF_2)-$

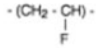
Ethylene/Chlorotrifluoroethylene (ECTFE)



$-(CH_2-CH_2)-(\overset{\overset{Cl}{|}}{C}F-CF_2)-$

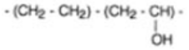
Poly(Vinylidene Fluoride) (PVDF) $-(CH_2-CF_2)-$

Poly(Vinyl Fluoride) (PVF)



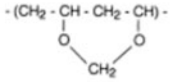
$-(CH_2-\underset{\underset{F}{|}}{C}H)-$

Ethylene/Vinyl Alcohol (EVAL or EVOH)

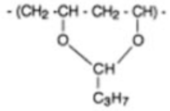


$-(CH_2-CH_2)-(CH_2-\underset{\underset{OH}{|}}{C}H)-$

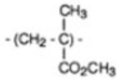
Poly(Vinyl Formal)



Poly(Vinyl Butyral)

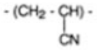


Poly(Methyl Methacrylate) (PMMA)



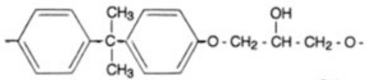
$-(CH_2-\overset{\overset{CH_3}{|}}{\underset{\underset{CO_2CH_3}{|}}{C}})-$

Poly(Acrylonitrile) (PAN)

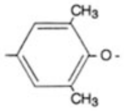


$-(CH_2-\underset{\underset{CN}{|}}{C}H)-$

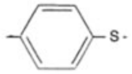
Phenoxy Resin



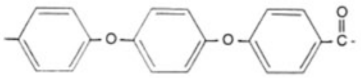
Poly(Phenylene Ether) (PPE or PPO)



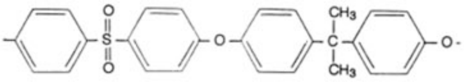
Poly(Phenylene Sulfide) (PPS)



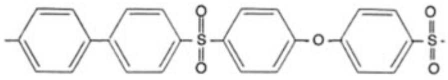
Poly(Ether Ether Ketone) (PEEK)



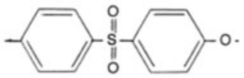
Polysulfone (Udel)



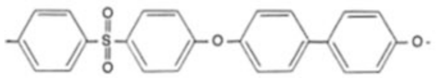
Polyaryl Sulfone



Polyether Sulfone

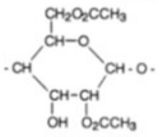


Polyphenyl Sulfone (Radel)

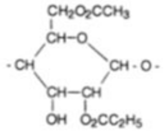


Poly(Oxymethylene) (Polyacetal) $-(CH_2-O)-$

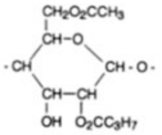
Cellulose Acetate (CA)



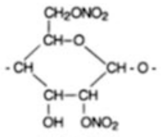
Cellulose Acetate/Propionate (CAP)



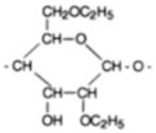
Cellulose Acetate/Butyrate (CAB)



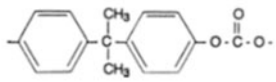
Cellulose Nitrate



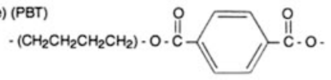
Ethyl Cellulose (EC)



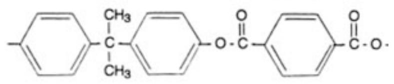
Polycarbonate (PC)



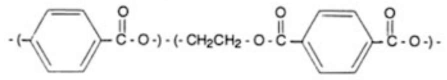
Poly(Butylene Terephthalate) (PBT)



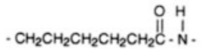
Polyarylate



Liquid Crystal Polyesters (LCP)

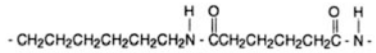


Nylon 6 (PA6)



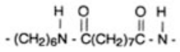
$-CH_2CH_2CH_2CH_2CH_2\overset{\overset{O}{\|}}{C}-\overset{\overset{H}{|}}{N}-$

Nylon 66 (PA66)



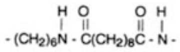
$-CH_2CH_2CH_2CH_2CH_2CH_2\overset{\overset{H}{|}}{N}-\overset{\overset{O}{\|}}{C}CH_2CH_2CH_2CH_2\overset{\overset{O}{\|}}{C}-\overset{\overset{H}{|}}{N}-$

Nylon 69 (PA69)



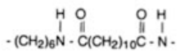
$-(CH_2)_6\overset{\overset{H}{|}}{N}-\overset{\overset{O}{\|}}{C}(CH_2)_7\overset{\overset{O}{\|}}{C}-\overset{\overset{H}{|}}{N}-$

Nylon 610 (PA610)



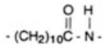
$-(CH_2)_6\overset{\overset{H}{|}}{N}-\overset{\overset{O}{\|}}{C}(CH_2)_8\overset{\overset{O}{\|}}{C}-\overset{\overset{H}{|}}{N}-$

Nylon 612 (PA612)



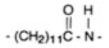
$-(CH_2)_6\overset{\overset{H}{|}}{N}-\overset{\overset{O}{\|}}{C}(CH_2)_{10}\overset{\overset{O}{\|}}{C}-\overset{\overset{H}{|}}{N}-$

Nylon 11 (PA11)



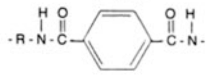
$-(CH_2)_{10}\overset{\overset{O}{\|}}{C}-\overset{\overset{H}{|}}{N}-$

Nylon 12 (PA12)

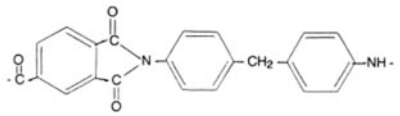


$-(CH_2)_{11}\overset{\overset{O}{\|}}{C}-\overset{\overset{H}{|}}{N}-$

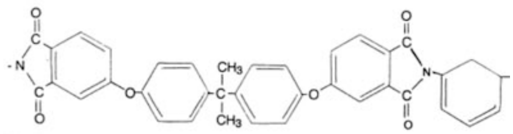
Poly(Phthalamide) (PPA)



Poly(Amide Imide) (PAI)



Poly(Ether Imide) (PEI)



Poly(Ether Imide) (PEI)

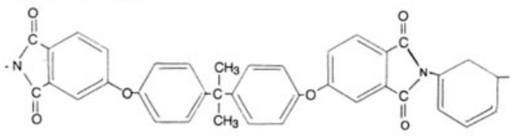


Table 15.10 Typical temperature resistance of engineering thermoplastics

Polymer	Reinforcing fiber	HDT (°C)	CST (°C)
Liquid crystal polyester		347	355
Polyetheretherketone	30 % Glass	316	250
Polyamideimide	Filled	280	80
Polyphthalamide	15% Glass	277	
Polyphenylene sulfide	30% Glass	265	220
Perfluoroalkoxy ethylene	20% Carbon	260	260
Nylon 6,6	33% Glass	250	130
Nylon 6,10	30% Glass	215	110
Polyaryl sulfone	30% Glass	213	
Nylon 6,12	33% Glass	210	110
Nylon 6	33% Glass	210	121
Polyetherimide		210	170
Polyethylene terephthalate	30% Glass	210	140
Ethylene/tetrafluoroethylene	25% Glass	210	177
Polyether sulfone		204	179
Polybutylene terephthalate	33% Glass	204	
Polyphenyl sulfone		190	
Polysulfone	10% Glass	183	160
Nylon 11	23% Glass	176	90
Fluorinated ethylene/propylene	20% Glass	176	204
Polyvinylidene fluoride	20% Carbon fiber	173	121
Polyarylate		172	
Polyoxymethylene	30% Glass	160	105
Nylon 12	23% Glass	160	100
Polyphenylene ether		146	105
Ethylene/chlorotrifluoroethylene	20% Glass	136	150
Polycarbonate		130	

The leading materials, in terms of market volume, are: nylon, polycarbonate, polybutylene terephthalate, polyphenylene ether, polyoxymethylene, and polyethylene terephthalate.

Perhaps even more important than their structures, most of these polymers are frequently reinforced by glass fibers or even carbon fibers, which contribute tremendously to their properties, and must be considered in any comparison of their practical performance. Reinforcing fibers generally raise modulus two to fourfold and usually increase breaking strength somewhat. In crystalline plastics, they often raise maximum use temperatures dramatically.

Maximum Use Temperature. The most frequent requirement for higher engineering performance is retention of properties at higher temperatures. Whereas most commodity thermoplastics soften and distort in boiling water, engineering thermoplastics are most often characterized by their ability to stand much higher temperatures. This is measured most often by the short-term HDT under a load of 1.82 MPa, less often but perhaps more practically by an estimated Continuous Service Temperature (CST) in long-term use (Table 15.10).

Major applications are primarily in electrical and electronic products, auto parts, industrial products, and appliances.

Impact Strength. When plastics are compared with metals and wood, they often fail under high-speed impact. Flexible molecules, such as nylons have some inherent impact strength. Of the rigid molecules, only polycarbonate and polyphenyl sulfone combine inherent rigidity and high impact strength. Most plastics can be reinforced with long fibers to increase their impact strength, but processing becomes much more difficult. Some plastics have been toughened by dispersing tiny rubber domains in them, and this technique is currently expanding with the development of compatibilization technology. One unique material is poly (1-butene), whose high creep- and puncture-resistance make it particularly desirable in pipe and tubing.

Transparency. Some applications of plastics require transparency. Amorphous plastics should be able to transmit light. Some factors which prevent transparency include unsaturation/light absorption, crystallinity, fillers and reinforcing fibers, and use of rubber particles to increase impact strength. The plastics most often used for their transparency are poly(4-methylpentene-1) (TPX), poly(methyl methacrylate) (almost equal to glass), cellulose acetate, propionate, and butyrate, polycarbonate, and polysulfones (slightly yellow). The advent of nanoscale fillers offers the potential for transparent filled materials, provided the fillers are adequately dispersed.

Processability. Engineering performance generally requires rigid molecules to give maximum rigidity, strength, and high temperature performance. Melt processability, on the other hand, generally requires flexible molecules to give a fluid melt. Several polymers that manage to combine easy melt processability with high rigidity/strength/heat resistance, include nylons, polyoxymethylene (polyacetal), and liquid crystal polymers. The first two combine flexible molecules in the melt, plus high crystallinity in the solid form. The molecular rigidity of liquid crystal polymers explains their high modulus and strength, but they also have easy melt processability.

Lubricity. Plastic gears and bearings are less polar than metals and therefore are relatively self-lubricating, without the need for lubricating oil. Ultra-high molecular weight polyethylene, fluoropolymers, polyoxymethylene (polyacetal), and nylons are the best. To improve other plastics even further, adding powdered polytetrafluoroethylene to them can produce an even greater increase in lubricity, and simultaneously also an increase in abrasion resistance.

Barrier Plastics. When plastics replace metals and glass in packaging, their permeability is often a limiting property. Barrier performance generally increases with density and crystallinity. The most promising barrier plastics include ethylene/vinyl alcohol, polyvinylidene chloride, polyacrylonitrile, and polyethylene naphthoate. These are used most efficiently by laminating them to commodity plastics such as polyethylene and polyethylene terephthalate. The use of nanoclay fillers has also shown improved barrier materials.

Weather-Resistance. For long-lived outdoor products, most plastics can be stabilized somewhat by opaque UV reflectors or at least dissolved UV stabilizers. For inherent resistance to sunlight, rain, and other components of weather, some preferred plastics include acrylonitrile/styrene/acrylic rubber, acrylonitrile/styrene/ethylene-propylene rubber, polyvinyl chloride, fluoropolymers, and polymethyl methacrylate.

Chemical Resistance. Plastics are generally superior to metals in resistance to aqueous inorganic environments. Their resistance to organic solvents depends mainly on crystallinity and difference in polarity: nonpolar polymers are more resistant to polar organics, while polar polymers are more resistant to nonpolar organics.

Table 15.11 Major end use market, 2008

Market	Millions of pounds
Transportation	2,751
Packaging	24,097
Building and construction	12,313
Electrical/electronic	1,755
Furniture and furnishings	2,671
Consumer and institutional	15,461
Industrial/machinery	834
Adhesives/inks/coatings	937
All other	1,375
Exports	11,962
Total	*74,156*

Individual Specialties. There are a number of special plastics, which are used for their individual special properties and applications. TPX combines rigidity, impact resistance, heat resistance, transparency, and chemical resistance, making it a unique replacement for glass in chemical equipment. Poly-*p*-xylylene is an extreme specialty coating, applied by vapor deposition to produce thin uniform films for electrical insulation. Polyvinyl formal is used in specialty wire coating. Polyvinyl butyral offers high toughness, adhesion, and clarity, which make it the critical component in safety glass. Phenoxy resin was one of the first engineering plastics, offering a combination of processability, rigidity, strength, toughness, transparency, adhesion, and chemical resistance; its present use is primarily in coatings and adhesives. Cellulose nitrate was the first commercial plastic; its present uses are primarily as high quality coatings on wood furniture and leather goods. Ethyl cellulose is a very tough transparent adhesive material, mainly used in coating bowling pins and specialty papers. Many other commercial polymers are more important in nonplastic applications such as rubber, textiles, paper, coatings, and adhesives (see Table 15.11).

Biopolymers

An alternative source for plastics is animal and vegetable matter. These biobased feedstocks are viable because plastics, chemicals, and petrochemicals currently account for only 2.1% of fossil fuel usage. Therefore, unlike biofuels, they will not affect food prices. Moreover, using biodegradable, biobased plastics (biopolymers)—along with improved recycling systems—can reduce the reviled plastics trash, especially in packaging. In the long term, the potential production of plastics from algae can also reduce green house gases.

Biopolymers, however, create new issues for the scientists and engineers who will use them. First, not all biobased polymers are biodegradable. Second, these bioplastics are far behind petroleum-based plastics in terms of performance and processibility. Third, the properties of biobased polymers tend to vary with the source matter, process, location, and year. Fourth, most biobased polymers are currently more expensive than their fossil fuel-based counterparts; so, noneconomic factors must be considered when specifying these materials for a product.

Depending on the synthesis process and sources, biobased and biodegradable polymers can be classified into four different categories [55]: (1) polymers from biomass such as the agro-polymers from agro-resources; (a) polysaccharides, e.g., starches (wheat, potatoes, maize) [56], ligno-cellulosic products (wood, straws, . . .) [57] and others (pectins, chitosan/chitin, gums) [58]; (b) protein and lipids, e.g., animals (casein, whey, collagen/gelatin) [59–63], and plants (zein, soya and gluten) [64]; (2) polymers obtained by microbial production, e.g., polyhydroxyalkanoates (PHA) such as poly(hydroxybutyrate) (PHB) and poly(hydroxybutyrate cohydroxyvalerate) (PHBv); (3) polymers chemically synthesized using monomers obtained from agro-resources, e.g., poly(lactic acid) (PLA); (4) polymers whose monomers and polymers are both obtained by chemical synthesis from fossil resources, e.g., polycaprolactones (PCL), polyesteramides (PEA), aliphatic co-polyesters (e.g., PBSA), and aromatic co-polyesters (e.g., PBAT). While all of the categories are biodegradable only the first three categories are obtained from renewable resources.

Until recently, biobased were the product of a biosynthetic process. That is, either from natural sources such as plants, animals or microorganisms (e.g., algae) by an extraction process. Today, chemical synthesis can be used on a biopolymer as a starting material leading to a monomer which is then again transformed to a biopolymer by a chemical polymerization process. PLA, a thermoplastic that is derived from corn, as produced by Cargill-Dow Inc. is a well-known example [65]. The plastic coffee mug shown in Fig. 15.26 is produced from PLA.

Companies such as Natureworks LLC and Metabolix have manufacturing facilities capable of producing 300 million pounds (140,000 metric tons) of PLA. Natureworks is the first company to offer a family of commercially available low carbon footprint polymers derived from 100% annually renewable resources with cost and performance that compete with oil-based plastics and fibers. Although both synthetic and biobased routes are known to produce resins of fairly consistent properties, the variability in properties with biobased sources is much higher than resins made from petroleum-based sources. As such, these biobased polymers are not fully competitive with conventional thermoplastics since their properties maybe too weak and/or are difficult to process. Therefore to extend the range of their applications, many biobased polymers are reformulated to include nanoclay fillers for the development of biorenewable nanocomposites with improved properties (e.g., stiffness and permeability) and processability [66].

Fig. 15.26 Poly(lactic acid) (PLA), a thermoplastic derived from corn

In summary, biopolymers have moved beyond the limitations of first generation products and technologies with improved processability and compostability, better cost positions, and an expanding customer base. Today's biobased and biodegradable polymers are proving to be economically feasible and environmentally advantageous. They have already succeeded in packaging and fibers, and are entering the automotive, medical, and consumer electronics markets. Global demand for these polymers is expected to reach $6 billion by 2015.

Thermoplastic Elastomers

Soft flexible-rubbery behavior depends on long flexible polymer molecules in the form of random coils. Strength, heat, and chemical resistance depend on attachment between the coils. Conventional rubber chemistry uses vulcanization, permanent thermoset primary covalent cross-links, usually by sulfur or peroxide, to hold the chains together; but this makes processing and recycling more difficult. In the past 40 years, this technology has been supplemented by the development of thermoplastic elastomers. These are based mainly on block copolymers, in which long flexible blocks form the continuous rubbery matrix, and short glassy or crystalline blocks form the thermoplastic "cross-links," secondary attractions that give strength, heat, and chemical resistance. Once heated these blocks (domains) melt and can be processed. Upon cooling the materials phase separate and

Table 15.12 Typical properties ranges of thermoplastic elastomers [67]

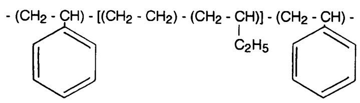
TPE type	Shore hardness	Tensile strength (MPa)	Ultimate elongation (%)	Low temperature limit (°C)	High temperature limit (°C)
Styrenics	5A-60D	1.7–12	400–1,840	−60 to −90	95
Rubber-polyolefin blends	50A-65D	5.4–15	80–940	−60	165
Thermoplastic vulcanizates (TPV)	35A-80D	4–12	375–650	−60	165
Thermoplastic polyurethanes	60A-85D	14–52	250–870	−40 to −60	190
Thermoplastic co-polyesters	90A-72D	14–46	250–550	−40 to −60	185–220
Thermoplastic polyamides	25D-63D	29–49	300–680	−40 to −60	220–275

the glassy or crystalline domains solidify and act as tie points or physical cross-links. These are much more attractive to the plastics processing industry, offering easy thermoplastic processing and good recyclability. They are the fastest growing sector of the elastomers market. There are six major families of commercial thermoplastic elastomers (Table 15.12).

The classes are divided by the type of hard block (segment) used to act as a physical cross-link. In each of the families, the ratio of rubbery soft block to glassy or crystalline hard block, can be varied to give a wide range in properties. The lower temperature range is dependent on the type of elastomer block used, while the high temperature properties are controlled by the hard block type [68].

Styrenic Block Copolymers. These ("styrenic") thermoplastic elastomers are block copolymers of styrene with butadiene (SBS) or isoprene (SIS) in about 30/70 monomer ratio.

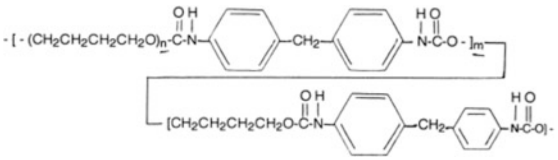
The elastomer segment give the rubbery properties and the glassy polystyrene domains hold the rubber chains together up to the softening point of polystyrene. Since the unsaturation in these elastomers is sensitive to oxygen and ozone aging, SBS is often saturated by hydrogenation (SEBS) to improve age resistance. They are typically compounded with other additives and used mainly in adhesives, shoe soles, wire and cable insulation, kitchen utensils, and auto parts.



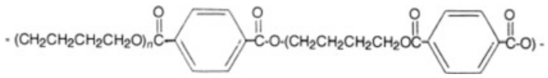
Rubber-Polyolefin Blends. These thermoplastic elastomers are blends of olefins (semicrystalline) with elastomers. The most common blend is polypropylene (PP) with ethylene-propylene rubber (EPR), but other combinations can be used. They have the elastomeric nature of the EPR rubber matrix and the crystalline polypropylene (PP) domains hold them together. The morphology of the blend is important to properties, typically an interpenetrating polymer network is required. The saturated elastomers will have natural resistance to oxygen and ozone aging. They are used in the automotive and wire and cable sectors.

Thermoplastic Vulcanizates. These materials are composed of a blend of polyolefin and an elastomer that is cross-linked during the mixing process (dynamic vulcanization). The most common combination is polypropylene and ethylene/propylene/diene (EPDM) rubber. In this process very small cross-linked rubber particles are dispersed in the thermoplastic matrix. The vulcanized rubber has higher heat and chemical resistance than ordinary EPR, and the PP matrix provides thermoplastic processability. In more specialized grades, the vulcanized rubber may be nitrile rubber for greater oil resistance, or butyl rubber for impermeability. They are used where greater heat and/or chemical resistance are required, for example, in oil wells, mechanical goods, and building and construction.

Thermoplastic Polyurethanes. These elastomers are multi-block, composed of hard polyurethane hard segments and soft polyether or polyester elastomer blocks. The polyether grades offer better stability to hydrolysis, while the polyester grades give better abrasion resistance. They are used in applications ranging from auto parts, wire and cable, medical products, mining screens, shoe soles and caster wheels.



Thermoplastic Co-polyesters. These thermoplastic elastomers are typically block copolymers of polyoxybutylene rubber with polybutylene terephthalate crystalline domains. They are creep resistant, chemically resistant and can sustain high loads. They are used in automotive and other mechanical parts requiring dynamic properties, flexibility, and strength.

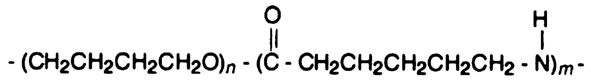


Thermoplastic Polyamides. These thermoplastic elastomers are multi-block copolymers of and elastomer with nylon

Table 15.13 Thermoset plastics market

Family	Percent of thermoset market
Polyurethane	36
Phenol–formaldehyde (phenolic)	29
Urea–formaldehyde	19
Polyester	10
Epoxy	4
Melamine–formaldehyde	2

crystalline domains. They combine the soft flexible properties of polyether or polyester elastomers with the high strength, heat-, oil-, and gas-resistance of nylons, and thus find use in auto parts, wire and cable, and sporting goods.

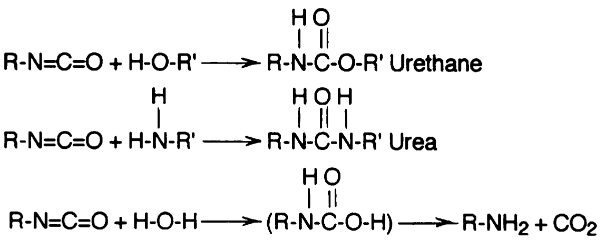


$$-(CH_2CH_2CH_2CH_2O)_n-(\overset{O}{\overset{\|}{C}}-CH_2CH_2CH_2CH_2CH_2-\overset{H}{\overset{|}{N}})_m-$$

Thermoset Plastics

Whereas difunctional monomers produce linear thermoplastic polymers, monomers with higher functionality can react further during processing, cross-linking up to infinite molecular weight. Such thermosetting processing may be more difficult, but infinite cross-linking produces extreme increases in rigidity, creep resistance, dimensional stability, heat resistance, and chemical resistance, which are valuable in many demanding engineering applications. Thus, thermoset plastics account for about 15% of the total plastics market. The major thermoset plastics families may be ranked in order of market size as shown in Table 15.13.

Polyurethanes. Most polyurethane chemistry may be simplified down to three basic reactions:

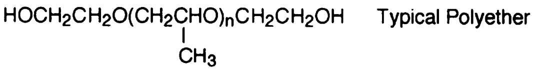


$$R\text{-}N{=}C{=}O + H\text{-}O\text{-}R' \longrightarrow R\text{-}\overset{H}{\overset{|}{N}}\text{-}\overset{O}{\overset{\|}{C}}\text{-}O\text{-}R' \text{ Urethane}$$

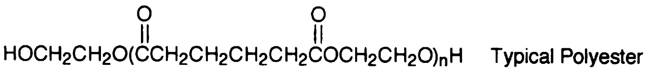
$$R\text{-}N{=}C{=}O + H\text{-}\overset{H}{\overset{|}{N}}\text{-}R' \longrightarrow R\text{-}\overset{H}{\overset{|}{N}}\text{-}\overset{O}{\overset{\|}{C}}\text{-}\overset{H}{\overset{|}{N}}\text{-}R' \text{ Urea}$$

$$R\text{-}N{=}C{=}O + H\text{-}O\text{-}H \longrightarrow (R\text{-}\overset{H}{\overset{|}{N}}\text{-}\overset{O}{\overset{\|}{C}}\text{-}O\text{-}H) \longrightarrow R\text{-}NH_2 + CO_2$$

Polyols are usually aliphatic polyethers or polyesters.



$$HOCH_2CH_2O(CH_2\underset{CH_3}{\underset{|}{C}}HO)_nCH_2CH_2OH \quad \text{Typical Polyether}$$

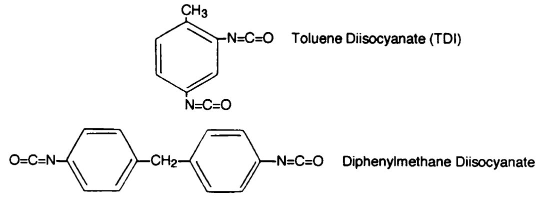


$$HOCH_2CH_2O(\overset{O}{\overset{\|}{C}}CH_2CH_2CH_2CH_2\overset{O}{\overset{\|}{C}}OCH_2CH_2O)_nH \quad \text{Typical Polyester}$$

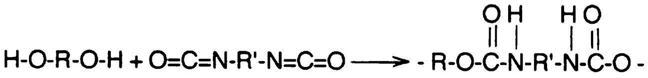
Polyisocyanates are usually toluene diisocyanate (TDI) or diphenylmethane diisocyanate.

Table 15.14 Polyurethane markets

Form	Percent of polyurethane market
Flexible foam	48
Rigid foam	28
Reaction injection molding (RIM)	6
Rubber, spandex, sealants, adhesives, coatings	17

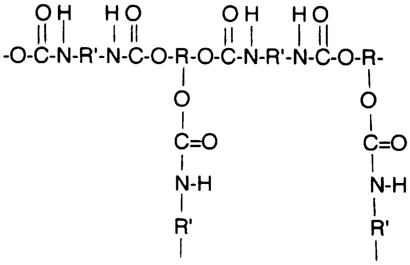


When the polyols are linear (difunctional) they form thermoplastic polyurethanes.



$$H\text{-}O\text{-}R\text{-}O\text{-}H + O{=}C{=}N\text{-}R'\text{-}N{=}C{=}O \longrightarrow -R\text{-}O\text{-}\overset{O}{\overset{\|}{C}}\text{-}\overset{H}{\overset{|}{N}}\text{-}R'\text{-}\overset{H}{\overset{|}{N}}\text{-}\overset{O}{\overset{\|}{C}}\text{-}O-$$

When the polyols are trifunctional or higher, they form thermoset polyurethanes.



In most processes, the reactive liquids are mixed and poured, and polymerized and cured rapidly to the final products (Table 15.14).

Most polyurethane is foamed during the polymerization/cure reaction.

Flexible Foam. This is made by mixing long trifunctional polyol with isocyanate to form the polyurethane, and adding a little excess isocyanate and water to the reaction to produce carbon dioxide, which produces the foam. The largest use is in furniture, with smaller amounts in auto seating, mattresses, rug underlay, textiles, and packaging.

Rigid Foam. This is made by mixing a short polyfunctional polyol with di- or higher polyisocyanate, and foaming either with volatile liquid or with isocyanate and water. The largest use is in building insulation, with smaller amounts in refrigeration, industrial insulation, packaging, autos, and marine flotation.

Table 15.15 Typical properties of molded phenolic plastics

Grade	General purpose	Engineering
Fibrous filler	Wood flour	Glass
Modulus (MPa)	9,000	18,000
Flexural strength (MPa)	76	275
Notched Izod impact strength (J/cm)	0.20	5.3
Heat deflection temperature (°C at 1.82 MPa)	168	240
Maximum continuous use temperature (°C)	149	194

Table 15.16 Phenolics as engineering plastics

Property	Phenolic	Engineering thermoplastics
Creep modulus, 100 h/14 MPa (MPa)	28	5.5–10
Compressive creep, 14 MPa/50 °C (%)	0.02	0.1–1.4
Heat resistance (°C)	315	121–260

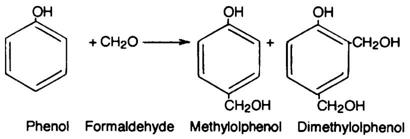
Reaction Injection Molding (RIM). RIM mixes polyol, polyamine, polyisocyanate, and a strong catalyst, and injects the mixture rapidly into a mold, where it cures rapidly to form large parts very economically. It is used primarily for producing front ends and other large parts of autos.

Phenol–Formaldehyde. Phenolic plastics were the first commercial synthetic plastics in 1908, and were the leading commodity plastic for 40 years, until the growth of vinyl and styrenic thermoplastics. Now quite mature, they remain the second largest family of thermoset plastics.

While their largest use is as adhesives for outdoor plywood and glass wool insulation, they provide a group of compression molding fiber-reinforced plastics, which meet high engineering performance requirements (Table 15.15).

For performance under severe conditions, they compare very favorably with more expensive engineering thermoplastics as shown in Table 15.16.

The chemistry of phenolic molding powders begins with the reaction of phenol with formaldehyde.



The high reactivity of the methylol groups makes it easy to polymerize and cure phenolic polymers, and also to copolymerize them with melamine and furan plastics, and with cellulosic reinforcements such as wood, cotton, and paper.

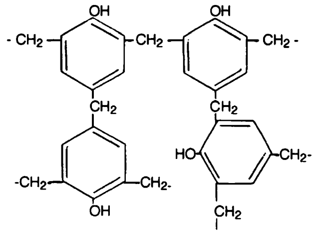
$$\mathrm{P{-}CH_2OH + HOCH_2{-}P \rightarrow P{-}CH_2OCH_2{-}P \rightarrow P{-}CH_2{-}P}$$

$$\mathrm{P{-}CH_2OH + H{-}P \rightarrow P{-}CH_2{-}P}$$

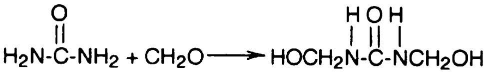
$$\mathrm{P{-}CH_2OH + HOCH_2{-}M \rightarrow P{-}CH_2{-}M}$$

$$\mathrm{P{-}CH_2OH + HO{-}Cellulose \rightarrow P{-}CH_2{-}O{-}Cellulose}$$

It is generally a three-stage process: (1) initial reaction to A-stage resin produces low molecular weight oligomers, which are still soluble, fusible, and reactive. (2) Melt compounding advances them to B-stage resins, which are fairly fusible to doughy melts, and still reactive. (3) Molding them into finished products advances them to C-stage resins, which are fully cross-linked to stable thermoset plastics of three-dimensional infinite molecular weight, high modulus and strength, and very resistant to heat and chemical environments.

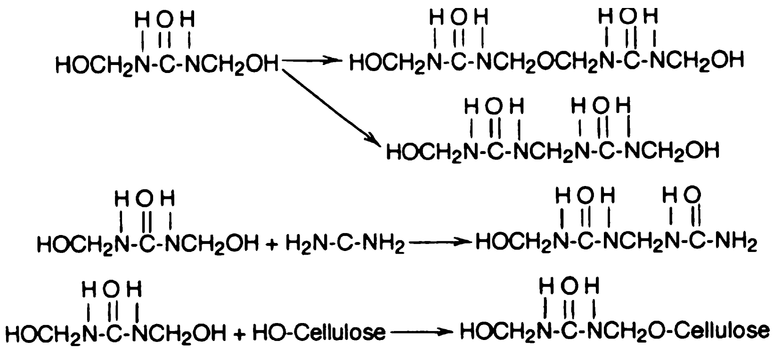


Urea–Formaldehyde. These plastics became commercial about 1929. Urea and formaldehyde react very readily to form methylol compounds, mainly dimethylol urea.



$$\mathrm{H_2N{-}\overset{\displaystyle O}{\overset{\|}{C}}{-}NH_2 + CH_2O \longrightarrow HOCH_2\overset{\displaystyle H}{\overset{|}{N}}{-}\overset{\displaystyle O}{\overset{\|}{C}}{-}\overset{\displaystyle H}{\overset{|}{N}}CH_2OH}$$

The methylol groups are very reactive, condensing with each other, with the N–H groups in urea, and with the –OH groups in cellulose.



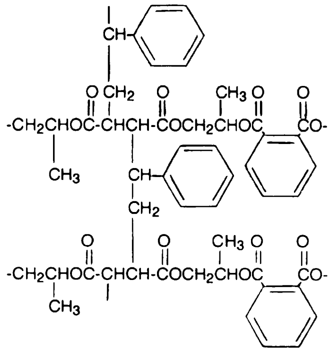
Thermosetting cure produces highly cross-linked three-dimensional molecules of infinite size. The largest use is for binding fibrous and granulated wood into indoor composition board. Smaller uses are for wet-strength paper and permanent-crease textiles. About 4% of urea–formaldehyde resin is combined with alpha-cellulose to make molding powders (Table 15.17). These find use mainly in electrical parts such as switches, wall plates and receptacles, circuit breakers, electric blankets, handles and knobs.

Unsaturated Polyesters. The chemistry of unsaturated polyesters was developed in the 1930s, and manufacture of glass–fiber-reinforced polyesters began in the early 1940s. They matured early, and form about 10% of the present thermosetting plastics market.

Their chemistry is fairly complex. The most common material is made from propylene glycol plus maleic anhydride plus phthalic anhydride. When these are cooked together, maleic anhydride isomerizes to fumaric acid, and they condense to form low molecular weight propylene fumarate phthalate copolyester oligomers. These are mixed with styrene monomer, reinforced by glass fibers, usually extended with low-cost fillers, and cured by peroxide to form rigid strong products which are very resistant to impact and heat (Table 15.18).

Table 15.17 Typical properties of urea–formaldehyde-alpha-cellulose moldings

Flexural modulus (MPa)	10,000
Flexural strength (MPa)	100
Notched Izod impact strength (J/cm)	0.16
Heat deflection temperature (°C at 1.82 MPa)	132



The mechanics of processing are carried out by a number of techniques. *Hand lay-up or spray-up* produces large shapes such as boat hulls, recreational vehicles, mobile homes, truck cabs, and tub-shower units. *Continuous panel processing* produces room dividers and skylights. *Compression molding* of sheet molding compound (SMC) and bulk molding compound (BMC) produces auto body parts, bathtubs, septic tanks, trays, tote boxes, and equipment housings. *Pultrusion* produces flagpoles, archery bows, and park benches. *Filament winding* is highly engineered to produce maximum strength in pipes and storage tanks.

Epoxy Resins. These were developed in the 1940s and offered a unique combination of engineering performance, which made them a popular family of thermoset plastics: fast low-temperature cure with low pressure and low shrinkage; high adhesion to polar surfaces; hardness, heat- and chemical resistance (Table 15.19). Their largest use is in coatings for corrosion protection and electronic equipment. The second largest is in printed circuit boards. Three other important uses are adhesives, flooring, and high performance fiber-reinforced plastics.

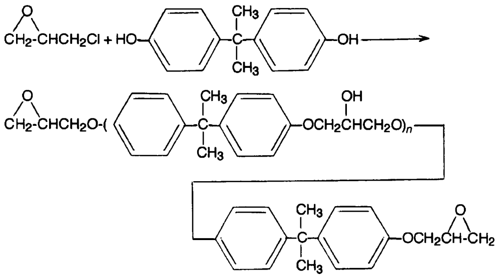
Table 15.18 Typical properties of glass–fiber-reinforced thermoset polyesters

Process	BMC	SMC	Woven cloth	Filament wound
Modulus (MPa)	14,000	11,000	14,000	6,000–24,000
Strength (MPa)	121	159	414	283–586
Notched Izod impact strength (J/cm)	4.0	7.8	9.4	
Heat deflection temperature (°C at 1.82 MPa)	182	225	205+	

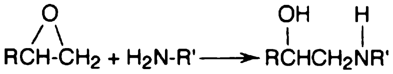
Table 15.19 Property ranges for cured epoxy resins

Flexural modulus (MPa)	14,000–34,000
Flexural strength (MPa)	55–655
Notched Izod impact strength (J/cm)	0.16–21.0
Heat deflection temperature (°C at 1.82 MPa)	93–288

Epoxy chemistry is complex. Most epoxy resins are made by reaction of epichlorohydrin with bisphenol A.

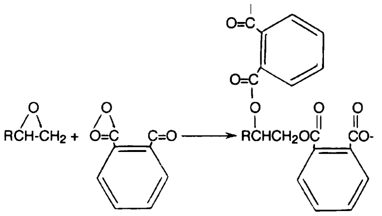


While they are called "resins," they are really monomers to low molecular weight oligomers, liquids to soluble fusible solids, with high reactivity in the epoxy rings, and fair reactivity in the internal hydroxyl groups. They are cured most often by room-temperature reaction with polyamines or polyamide amines.



$$R\overset{\overset{O}{/\ \backslash}}{CH-CH_2} + H_2N\text{-}R' \longrightarrow R\overset{\overset{OH}{|}}{C}HCH_2\overset{\overset{H}{|}}{N}R'$$

For higher temperature cure and heat resistance, they are cured most often by acid anhydrides.



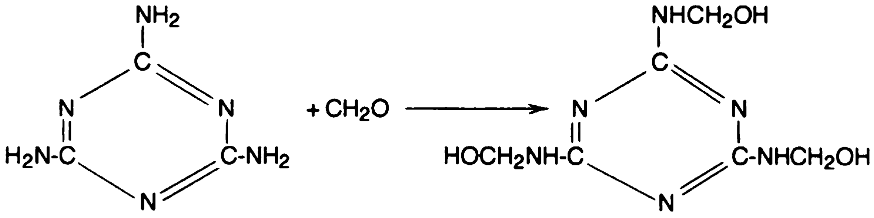
There are also many more types of epoxy resins and curing agents for more specialized applications.

Melamine–Formaldehyde. These resins became commercial in the 1930s. Their combination of high thermosetting reactivity, cured hardness, and resistance to heat, weather, and chemical environments made them particularly valuable for their good appearance and durability (Table 15.20). Their largest use is in coatings, where they are used to cure acrylic automotive coatings and polyester appliance coatings. Their second largest use is in countertops, where they protect the decorative surfaces against abrasion, heat, and chemical attack. Their third use is in dinnerware, where their lightweight, impact resistance, and attractive appearance are very competitive with china.

Table 15.20 Typical properties of cellulose-filled melamine–formaldehyde

Flexural modulus (MPa)	9,000–11,000
Flexural strength (MPa)	70–124
Notched Izod impact strength (J/cm)	0.13–0.21
Heat deflection temperature (°C at 1.82 MPa)	127–143

Melamine chemistry begins with the addition of 2–3 mol of formaldehyde to form methylol melamines.



These are very reactive with each other, with the remaining N–H bonds on melamine, with the hydroxyl groups in acrylic and polyester coatings, and with the hydroxyl groups in paper for countertops and in alpha-cellulose for molded dinnerware.

$$RNHCH_2OH + HOCH_2NHR' \rightarrow RNHCH_2OCH_2NHR' \rightarrow RNHCH_2NHR'$$

$$RNHCH_2OH + R'NH_2 \rightarrow RNHCH_2NHR'$$

$$RNHCH_2OH + HOR' \rightarrow RNHCH_2OR'$$

The combination of resonance stabilization in the melamine heterocycle, and the high cross-linking between methylol melamines and with the other polymers, all produce the outstanding properties which make it a valuable specialty member of the thermoset plastics spectrum.

General Considerations

Structure–Property Relationships

When plastics engineers want to improve properties in an existing product, or when they want to select the optimum material for a new product, a routine search of existing tables of properties may sometimes be sufficient. For more professional judgment and problem-solving, however, and for planning development of new materials, they need to understand the basic relationships between polymer structure and practical properties. Here are some starting guidelines.

Molecular Weight. Low molecular weight gives lower melt viscosity for injection molding, and easier solution processing in general. High molecular weight is preferred for extrusion, and particularly for blow molding, thermoforming, stretch orientation, and thermoplastic foaming. In finished products, high molecular weight generally gives higher mechanical strength and chemical resistance.

Molecular Flexibility/Rigidity. Flexible molecules generally give lower melt viscosity for easier processing; softer, more flexible, more extensible products; and higher impact strength, friction, and acoustic absorption. Rigid molecules generally give higher rigidity, strength, creep resistance, heat deflection temperature, and impermeability; and lower coefficient of thermal expansion, and dielectric constant and loss.

Crystallinity. When polymers crystallize, their melting points are much higher and sharper than the softening points of amorphous plastics. This requires higher processing temperatures, but gives higher maximum use temperatures in the final products. Increasing crystallinity generally increases rigidity, strength, creep resistance, dimensional stability, impermeability, and chemical resistance; but decreases impact strength and transparency.

Orientation. Stretch orientation of extruded fibers and films greatly increases modulus, strength, transparency, and impermeability. It is also useful for producing shrink-packaging; conversely, a disadvantage is thermal dimensional instability. When orientation occurs accidentally in injection molding, calendering, thermoforming, and other processes, it generally produces undesirable anisotropy of final structure and properties.

Polarity and Hydrogen Bonding. Whereas hydrocarbon polymers are nonpolar and have weak intermolecular attraction, introduction of negative atoms into the polymer molecule—oxygen, nitrogen, chlorine—produces permanent polarity, giving strong intermolecular attractions. Increasing polarity generally requires higher processing temperatures, and gives higher modulus, strength, creep resistance, heat deflection temperature, crystalline melting point, dielectric constant and loss, and gasoline and oil resistance.

When the polymer contains oxygen and especially nitrogen, polarity also produces hydrogen bonding, which is an even stronger intermolecular attraction, and produces all the same effects to an even greater extent. Another effect of hydrogen bonding is water absorption. In fabrics this produces greater comfort; but in plastics it decreases modulus, strength, and dimensional and chemical stability. These effects are most noticeable in nylons and cellulosics.

Cross-linking. Thermoplastics are stable linear molecules which are softened by heat and soluble in solvents of similar polarity; this makes for easy processability. Primary covalent cross-linking in thermosets converts them into three-dimensional molecules of infinite size, with tremendous changes in properties: insolubility and infusibility; higher modulus, creep resistance, maximum use temperature, and chemical resistance; and lower extensibility, impact strength, thermal expansion, dielectric constant and loss, solvent swelling, and permeability. Cross-linking produces shrinkage strains, which embrittle the polymer, so most thermoset plastics must be reinforced with fibers; the result is synergistic improvement of modulus, strength, impact resistance, and dimensional stability, producing enhanced engineering performance.

Additives

Polymers are rarely used in pure form. They are almost always improved by use of additives to enhance specific properties. The major classes of additives may be briefly summarized as follows.

Stabilizers. Organic polymers are not perfectly stable. Specific polymers and specific products require additives to improve their stability during processing and/or long-term use of finished products.

Antioxidants are added to polyolefins and rubber-modified impact plastics to protect against atmospheric oxygen; these are primarily hindered phenols and polyphenols, sometimes synergized by aliphatic sulfides or organic phosphites, used in fractions of a percent up to several percent.

Thermal Stabilizers must be added to PVC to prevent dehydrochlorination, discoloration, and cross-linking during melt processing: organotin esters are strongest, used as a fraction of a percent for processing rigid PVC; barium/zinc soaps plus epoxidized fatty esters plus organic phosphites, total concentration several percent, form a synergistic stabilizer system for plasticized flexible PVC; and basic lead oxide compounds, several percent, are best for wire and cable insulation.

Ultraviolet light Stabilizers are needed in products for outdoor applications: cyclic hindered amines and *o*-hydroxy benzophenones and benzotriazoles are used at a fraction of a percent in clear products, while carbon black and especially aluminum flake are extremely effective in opaque products.

Biostabilizers are used to protect natural polymers and monomeric additives against attack by microorganisms; these are chemicals which require a delicate balance between toxicity to microorganisms vs. safety for macroorganisms like ourselves.

Fillers are inorganic powders added in large amounts to increase modulus, dimensional stability, and opacity, and often to reduce cost.

Reinforcing Fibers are mostly glass, occasionally carbon or organic fibers, typically added in concentrations of 10–40%, to increase modulus, strength, impact strength, creep resistance, and dimensional stability; long and continuous fibers give the greatest improvement in properties, while short chopped fibers (1/16–2 in.) permit fairly conventional melt processing.

Coupling Agents are chemical surface treatments applied to fibers, and sometimes to fillers, to strengthen the interface between inorganic fillers and fibers and organic polymer matrixes, to improve dispersion and stress transfer across the interface. Most common are organosilicon compounds of the type $(RO)_{2-3}SiX_{2-1}$, where RO is typically methoxy or ethoxy to react with the silanol surface of glass fibers, and X is an organic group designed to react with a thermosetting polymer matrix, or at least to be attracted toward a thermoplastic polymer matrix.

Plasticizers are typically organic liquids of very low volatility, which are miscible with a polymer, and are added to it to improve processability and, in larger amounts (20–80 PHR), to make it soft and flexible, or even rubbery and/or adhesive. The major portion of the plasticizer market (80%) is aliphatic and aromatic esters, which go to convert rigid PVC into flexible PVC. The remainder goes to improve the processability of cellulosics, and for a variety of specialized uses in other polymers.

Lubricants are a variety of proprietary additives, which are used either to improve melt flow, release from steel process equipment, or self-lubricity in final products such as gears and bearings.

Flame-Retardants may be built into the polymer during polymerization or cure, or they may be physical additives to the finished composition. Organic phosphorus compounds are the most effective, typically requiring about 2% of phosphorus to prevent burning. Organobromine (10–20%) and chlorine (10–40%) are effective when used in larger amounts. Antimony oxide, and some other metal oxides, can synergize the action of bromine and chlorine, reducing total flame-retardant concentrations to 5–10%. Alumina trihydrate and magnesium hydroxide, which release water when heated, are becoming increasing popular when used in large enough amounts to be effective (40–60%).

Colorants. One third of plastic materials are used in their natural color. The other two thirds are colored for esthetic and/or functional reasons, typically using about 1% of colorant. Inorganic minerals and synthetic colors give greater stability and opacity; while organic colors are available in greater variety, miscibility, and efficiency, and less likely to raise questions of toxicity. Thus use of inorganics is decreasing, while use of organics is increasing.

Organic Peroxides are used to initiate free radical polymerization of ethylene, butadiene, styrene, vinyl chloride, vinyl acetate, and methyl methacrylate. They are also used to cure unsaturated polyesters, occasionally to cross-link thermoplastics such as polyethylene and polyacrylates, and increasingly for grafting and compatibilization of polymer blends. A variety of organic peroxides offer useful reactivity over a temperature range from 0 to 130°C or more, for different polymers and different processes.

Polymer Blends. Blending of polymers with each other accounts for approximately 40% of the present plastics market, and the practice is continually growing, because it permits the development of improved properties without the cost of inventing new polymers. When polymers are fairly miscible, as in the case polyethylenes and in polyphenylene ether plus polystyrene, blending can be used to produce a balance of properties. Most polymer blends are not miscible, and separate into microphases; if these can be strongly bonded at the interface, it is often possible to produce synergistic improvement of properties, particularly the balance between rigidity and heat deflection temperature on the one hand, plus ductility and impact strength on the other; or between soft, flexible, rubbery properties on the one hand, plus the strength of “thermoplastic cross-linking” on the other. Some of the newer blends are used to produce barrier properties and other valuable improvements.

Critical Properties: Challenges to the Plastics Industry

There have been numerous accomplishments of the plastics industry, but there still remain areas where major breakthroughs could open major new markets and uses for plastics in the future. It may be stimulating to explore some of these here.

One-Step Conversion of Monomer to Finished Product could reduce processing steps and costs dramatically. Epoxy cure and polyurethane RIM are examples of very fast reactions producing finished products. Monomer casting of acrylics and nylon 6 are commercial one-step processes. Polymer chemists have many more polymerization reactions which can rapidly convert monomers into high molecular weight or even thermoset polymers. What is needed are strong cooperative programs between polymerization chemists and plastics process engineers to develop these possibilities into commercial realities.

Continuous-Fiber Reinforcement gives plastic products which are superior to most present commercial practice. Most plastic processing is limited to conventional melt flow of short-fiber reinforcements, which sacrifices much of the potential benefits of reinforcement. There are a few processes for incorporating continuous-fiber reinforcement—filament-winding, pultrusion, and mixed fabrics of reinforcing fibers and plastic fibers. More vigorous development of such techniques could rapidly produce plastics products with far superior properties.

Modulus and Creep Resistance of plastics are still inferior to metals, ceramics, and glass, which means there are areas where they cannot compete. We know that molecular rigidity, crystallinity, polarity, and reinforcing fibers can all go a long way toward closing the gap.

Abrasion-Resistance of transparent plastics still cannot equal that of glass. We have ways of improving or coating the polymer, but there is still a long way to go.

Coefficient of Thermal Expansion of organic polymers is 1–2 orders of magnitude higher than metals, ceramics, and glass, which gives serious difficulties in product design, and especially in mating plastics parts with inorganic parts in an assembled product. Molecular rigidity, crystallinity, fillers, and especially fibrous reinforcement can go a long way toward bridging the gap, but there remains room for improvement.

Thermal Conductivity of plastics is very low, which makes them excellent insulators against heat and cold. On the other hand, there are times when high thermal conductivity is preferred, for example, in processing, cooking, and heating equipment. Inorganic fillers can help, in proportion to their volume concentration. New developments in nanoscale fillers may help improve these properties.

Heat Resistance of organic polymers is far lower than that of metals, ceramics, and glass. There have been major improvements, based on aromatic and heterocyclic resonance, ladder structures, and other mechanisms, and we may see further improvements in the future. One of the more serious limitations is the high cost of synthesis and the difficulty of processing these polymers into the desired final products. This is an area where the polymer chemist could use more help from the plastics engineer.

Electrical Conductivity of polymers is very low, making them very useful as insulation. On the other hand, there are products where conductivity would be very desirable. Semiconductivity is fairly easily achieved by adding semicompatible hydrophilic organic compounds, and fairly high conductivity can be achieved by metallic fillers, especially fibers. Several polymers are available, that are inherently conductive due to conjugated unsaturation, plus doping with inorganic electron donors.

Dielectric Breakdown occurs when high voltage drop across an insulator causes some current to leak through, turning to heat, and ultimately decomposing the polymer and burning a conductive carbon track right through the insulator. Practically, some polymers are more resistant than others; but more theoretical understanding is needed in order to design polymer structures which will offer superior resistance to dielectric breakdown.

Outdoor Weathering of plastic products has been the subject of both theoretical and practical study. Some plastics can last for many years, others for more limited times, and a few can actually be designed to self-destruct rapidly. Mechanisms involve ultraviolet light, atmospheric oxygen, water, transition metals, acid rain, wind-blown dust, and microbiological action. Polymer structure and additives respond to these mechanisms in various ways. Whereas there presently exist a fair theoretical understanding and practical control measures, there remains much to learn in order to achieve the ultimate goals of long-term weather stability and efficient control of solid waste.

Solvent-Resistance of organic polymers varies with polarity, crystallinity, and cross-linking, so it is usually possible to solve solvent-attack problems by proper choice of polymer; but it is important to remember that they do not easily compete with metals, ceramics, and glass.

Barrier Properties of organic polymers cannot equal metals and glass. Plastics offer so many other advantages that we often try to compromise or laminate to optimize overall balance of properties. It is known that molecular rigidity and crystallinity improve barrier performance, and from practical experience we can identify several very high-barrier polymers; new developments in nanoclay filled materials are helping in this arena.

Cost of engineering performance is a major factor retarding growth. Organic polymer chemists can easily design and synthesize polymer structures with higher properties, but the cost of synthesis and difficulty of processing often inhibit their use for many years. Closer cooperation between chemists and engineers should be the optimum route to more efficient development.

Fire Performance

Wood burns. Most of the fires throughout history, causing death and destruction, have been caused by wood. It is rare that anyone will say “Ban wood.” Plastics are also carbon compounds, and they also burn. When plastics are involved in a fire, even to the slightest degree, there is often an outcry, “Ban plastics.” Thus the use of plastics in building and transportation is seriously restricted by this prejudice.

Plastics are not all equally flammable. Some burn as readily as fuels. Some do not burn spontaneously; but when exposed to a severe fire, they can be burned. Some require enriched or pure oxygen environment before they will burn. Most plastics can be made more resistant to burning, by incorporating flame-retardant elements—phosphorus, bromine, chlorine, antimony, even water—either in the polymer molecule or in physical additives.

Related problems must be considered in individual products. Bromine, chlorine, and antimony add to the smoke of a fire, while phosphorus and water do not, and some metal oxides can actually reduce it. Toxicity of combustion gases is a major concern; but oxidation of carbon compounds in an enclosed space—indoors—produces carbon monoxide, no matter whether the carbon compounds are wood or plastics. Other problems include concern over the effect of brominated flame-retardants on the environment and legislation against their use as an additive, the cost of flame-retardants, difficulties in processing, and loss of mechanical or thermal properties. The designer must balance all of these in each product, and choose the optimum solution to the problem.

Health and Environment

Whenever new chemistry is introduced into the environment, there arises the question of its effect on our health. On the positive side, plastics packaging of food prevents contamination and spoilage, and prolongs its useful life; and use of plastics in medicine has made major contributions to health and longevity. On the negative side, there have been several occasions where plastics chemistry has caused health problems; whenever these have been identified, they have been solved successfully and quickly.

Recycling

Modern science and industry have provided a growing supply of material products. When they reach the end of their useful life, they become solid waste, and disposing of it has become a growing problem. Over-packaging to stimulate sales has been a major contributor to this solid waste. It would be desirable to remove them from solid waste by recycling.

Since 85% of the plastics market is thermoplastic, it is theoretically recyclable. In industry, individual thermoplastic materials are almost always recycled immediately for purely economic reasons. Postconsumer waste, on the other hand, presents serious problems. Voluntary recycling of PET soda bottles has been most successful, because they are easily separated, cleaned, and converted into bulk fiber products. Voluntary recycling of HDPE milk and water bottles has been fairly successful, because they are easily separated and cleaned, but development of markets has been slower in coming. Most other plastics occur in smaller amounts and more diversity, so voluntary efforts have been very limited. The greatest problem is that, when several materials are combined in a single product, separation is difficult. Ultimately, the success of recycling will depend very much on greater cooperative efforts by government, industry, and consumers.

Recycling rarely reproduces virgin plastic materials. With repeated recycling, quality decreases, and potential uses decrease. At some point, the final step should be incineration. This produces useful energy, returns carbon dioxide and water to the ecocycle, and reduces final solid waste to an absolute minimum. Present incineration technology suffers from old inefficient equipment and non optimal operation. New methods to recycle or incinerate waste will help solve this problem.

Part III: Plastic Processing

Rheology

Fundamental Concepts

Plastic processing involves the he flow and shaping of viscous liquids. The scientific study of this flow is called *rheology*. Assuming laminar shear flow, viscosity is defined as the ratio of shear stress to shear rate.

$$\eta = \frac{\sigma}{\gamma}$$

If this ratio is constant, it indicates a simple Newtonian fluid. For most plastic materials, however, increasing shear rate disentangles polymer molecules and aligns them in the direction of flow, so increasing shear rate decreases resistance to flow (viscosity) (Fig. 15.27), and this non-Newtonian behavior is defined as *pseudoplastic*.

With increasing temperature, viscosity decreases in a manner approximately described by the Arrhenius equation,

$$\eta = K\mathrm{e}^{-E/RT}$$

where K is a constant characteristic of the polymer and the test method, E is activation energy, R is the gas constant, and T is absolute temperature (K). With increasing pressure, free volume between polymer molecules decreases, flow becomes more difficult, and viscosity increases. A major factor in polymer viscosity is molecular weight, M, as shown by

$$\eta = KM^a$$

Beyond a critical molecular weight, M_c (typically 5,000–40,000), molecules become entangled and the viscosity increases more rapidly (see Fig. 15.28).

Instrumental Measurement of Flow Properties

Capillary Rheometers measure the effect of pressure on volumetric flow through a cylindrical capillary. They are popular in practical work because shear rate and flow

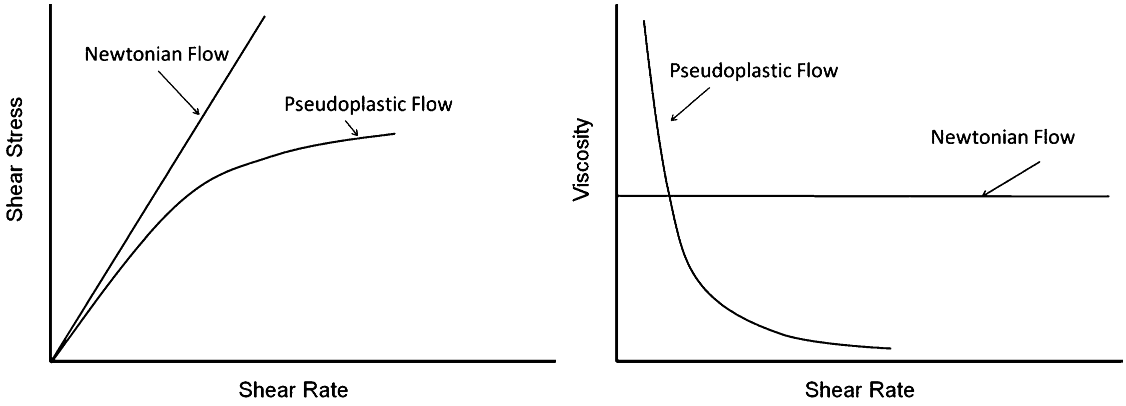


Fig. 15.27 Polymer rheology

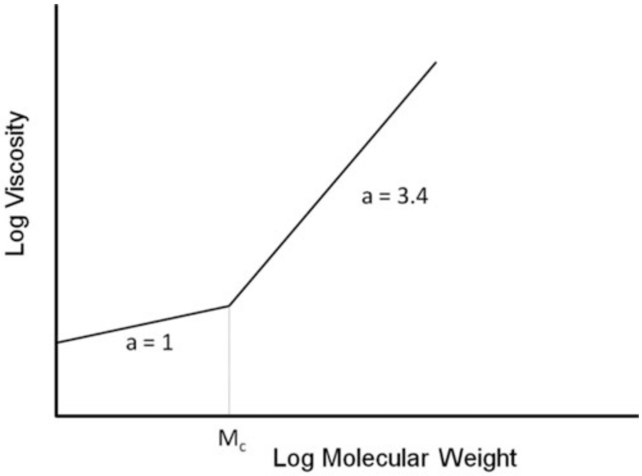


Fig. 15.28 Effect of molecular weight on viscosity

geometry are similar to conditions in extrusion and injection molding. They cover a wide range of shear rates, and they give practical information on die swell, melt instability, and extrudate defects. Their main disadvantage is that they require a number of mathematical corrections to convert to true viscosity.

Cone and Plate Rheometers provide a constant shear rate. They can also be designed to measure torque, dynamic properties, normal stresses, and forces in other directions. A disadvantage is that they are limited to low shear rates.

Parallel Plate Viscometers are used for very high viscosities at low shear rates. Measurement of shear rate is difficult.

Coaxial (Concentric) Cylinder Viscometers provide nearly constant shear rate. A disadvantage is that they are limited to liquids of low viscosity.

Extensional Viscometers are useful to measure tensile viscosity in processes such as stretch orientation.

Dynamic or Oscillatory Rheometers measure viscous and elastic modulus in shear or tension. Energy dissipation produces a phase difference, so stress, strain, and phase angle can be used to characterize complex viscosity behavior.

Practical Effects of Flow Properties

Melt Viscosity is the most critical practical property for the process engineer. When stiff molecules give high viscosity and slow flow rates, higher temperature and/or shear sensitivity (pseudoplasticity) can often be used to decrease the melt viscosity and thus increase processability.

Extrusion Defects are primarily due to melt elasticity. When shear rate occurs faster than polymer molecules can disentangle from each other, they stretch elastically and store potential energy. As the molecules exit from the die, they release this energy and recover elastically. The resulting extrudate diameter is then greater than the die diameter. This is called *die swell*. The effect is aggravated by high molecular weight, high pressure, low temperature, high shear rate, and short *L*/*D* ratio. Two practical ways to compensate for die swell are to: (1) reduce the die diameter, or (2) pull the extrudate away from the die at a velocity at least as great as it went through the die.

Melt Fracture is the occurrence of distorted extrudate coming from the extruder. It is caused by flow disturbance at the point where flow cross-sectional area is rapidly reduced from the large diameter of the melt feed to the much smaller cross-sectional area of the die orifice. Here again it is aggravated by melt elasticity.

Land Fracture is a fine surface roughness on the extruded product. It is caused by friction between the melt and the wall (land) of the die. It is solved by addition of lubricants to reduce melt/die adhesion (see Table 15.21).

Extrusion

Extrusion is the process of forming a material continuously through an opening, forming two-dimensional shapes.

Most extruders do this by rotating a screw inside a stationary heated cylindrical barrel, to melt the polymer and pump the melt through a suitably shaped orifice (Fig. 15.29). This is used for direct manufacture of finished products such as film or pipe. It may also be used to feed a second process such as injection molding, blow molding, coating, laminating, or thermoforming. It is also important in compounding—blending polymers with additives to improve the overall balance of properties—and in recycling plastics out of solid waste.

Basic Functions

The extruder generally performs six successive functions: solids conveying, melting ("plasticating"), melt conveying (pumping), mixing, devolatilization (venting), and forming. While a complete process may involve all six, some processes may omit one or more of these steps.

Solids Conveying is carried out in two steps: the feed hopper and the back (entrance) portion of the screw. The feed hopper is an inverted cone or pyramid, in which solid pellets or powder flow downward from the force of gravity. If they flow poorly ("arcing" or "bridging"), the problem may be solved by installing a vibrator or a stirrer ("crammer/feeder") in the hopper, or by coating the particles with a small amount of a chemical that acts as an "external lubricant."

Table 15.21 Plastics processes

Process	Use in industry (%)
Extrusion	34
Injection molding	31
Blow molding	13
Calendering	6
Thermoforming	6
Coating	5
Compression molding	3
Rotomolding	2

Once the solid particles enter the back end of the screw, they are carried forward and compressed by the rotation of the screw and the friction between the solids and the barrel. Friction can be increased by roughening the surface of the barrel, and particularly by grooving the internal surface of the barrel.

Melting of the solid polymer is the result of heat from two sources: conduction from the heated barrel, and friction between the barrel and the polymer. Most of the heat comes from friction between the barrel, the viscous melt film, and the remaining solid polymer. This frictional heating is most efficient when the melt viscosity is high, and when the melt film is thin. Thinness of the melt film depends on designing and maintaining a very small clearance between the screw flights and the barrel surface.

Melt Conveying is the forward motion of the molten polymer through the extruder, due to the pumping action of the rotating screw. This simple "drag flow" is proportional to melt density, down-channel velocity, and cross-sectional area of the screw channel. In most cases, however, there is also a pressure gradient as the melt moves downstream, either positive pressure approaching the exit die, or negative pressure following the solid/melting zones; so this "pressure flow" must be subtracted from the drag flow to calculate the net output of the extruder. Since pressure flow is inverse to melt viscosity, which is non-Newtonian and temperature sensitive, this complicates the calculation considerably.

Mixing in the extruder is important for homogenization of temperature and pressure, and especially for uniform blending with additives. In a simple single-screw extruder, melt flow is fairly linear and provides little mixing. A variety of modifications of screw design have been developed to build in mixing elements. Some improve distributive mixing of the liquid melt, to homogenize temperature and pressure

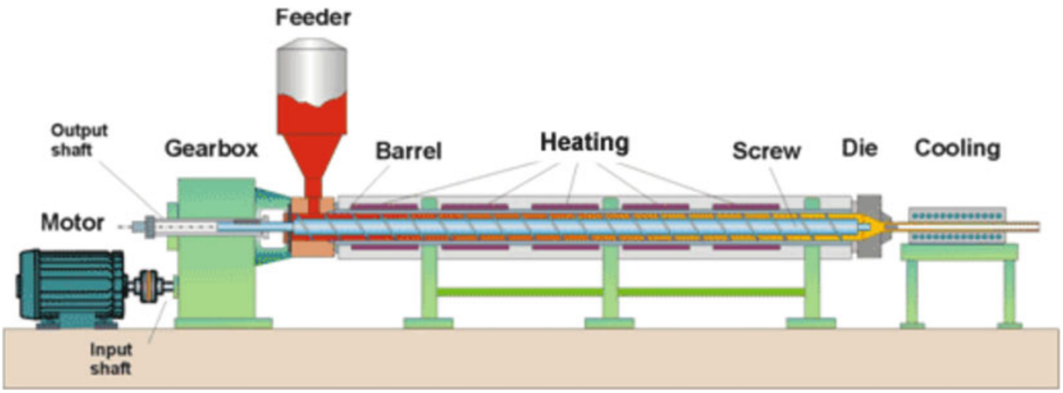


Fig. 15.29 Extruder. (http://www.theadvancedteam.com/AdvancedTeamCapabilitiesStatement.pdf, Accessed 4 Nov 2011)

fluctuations. Others build in higher shear (dispersive mixing), to break down particles of additives and blend them uniformly into the molten polymer, while still others combine the two types of action.

Twin-screw extruders permit much wider variation in design and performance. The screws may be non-intermeshing, just touching ("tangential"), partially or fully intermeshing. They may be corotating or counter-rotating. They are assembled ("programmed") of different sections ("elements"), designed for feed, melting, conveying, distributive melt mixing, dispersive shear for additives mixing, and sealing pressure or vacuum at the vent or the die. The barrels are also programmed of elements to provide functions such as feed ports, venting, and abrasion resistance. Their higher mixing efficiency makes them particularly useful in compounding with additives, processing polyvinyl chloride, reactive extrusion, and devolatilization.

Devolatilization can be used to remove up to 5% of volatile impurities from the plastic melt. The first melt conveying (metering) zone builds up melt pressure. Then channel depth is increased abruptly in the vent zone, the melt is decompressed, and volatiles escape through the vent. After devolatilization, the melt enters a second metering zone, which builds up melt pressure again, and feeds it to the die.

Die Forming forces the melt into the shape and dimensions desired in the final product. Temperature, pressure, viscosity, die design, flow rate, and flow patterns must all be optimized and controlled closely to make a suitable product. Typical guidelines include small approach angles, land length = 10 × land clearance, avoid abrupt changes in geometry, no dead spots, generous radii, and thin uniform wall sections.

Calibration is a technique for maintaining the shape and dimensions of the product from the time the melt exits from the die until it can be cooled enough to solidify and stabilize it. Depending on the size and shape of the product, calibrators can use water-cooled plates, internal mandrel, internal air pressure, external vacuum, or a post-extrusion die to change the shape of the molten extrudate before it solidifies.

Major Processes and Products

Blown Film is produced from a single-screw extruder by extruding a tube, cooling it with external and/or internal air streams, stretching it in the machine direction by pulling it away from the die ("draw-down"), stretching it in the transverse direction (typically 2–4× "blow-up" ratio) by internal air pressure up to 34 kPa (5 psi), flattening it by passing through nip (pinch) rolls, and winding it onto a cylindrical roll. Optional post-stretching operations may include flame or corona surface treatment for wettability/adhesion, sealing, slitting, and bag-making.

Flat Film, Sheet, and Coating are produced from a single-screw extruder with a high *L*/*D* ratio (27–33/1), which feeds the molten polymer through a flat die. The die opening is adjusted to control the thickness of the film, and the film is solidified in a cold water bath, or preferably, for transparency and gloss, over two or more water-cooled steel chill rolls. Compared to air-cooled blown film, water-cooled flat film generally has higher clarity. The take-off and wind-up line may include automatic feedback thickness control, surface treatment, and/or slitting.

For extrusion-coating, the substrate—paper, plastic, or metal foil—is preheated, and may be pretreated, before the extruder deposits a layer of molten polymer onto its surface. Low molecular weight and high temperature help the polymer to flow into a uniform adhesive coating. The laminated layers pass between pressure and chill rolls, and optionally through surface treatment, printing, and slitting before collecting on the final wind-up rolls.

Coextrusion produces multilayer laminates in a single process step. Two or more extruders feed different molten polymers into a multi-manifold die, which layers them directly, or into a modular feedblock which layers them before feeding them into the die. This is used primarily in the packaging field, to sandwich an impermeable barrier layer between two commodity outer-film layers, and often includes adhesive tie layers to bond the barrier layer to the outer layers.

Pipe, Hose, and Tubing are extruded through an annular (ring) die. *L*/*D* ratio is typically 24/1 or greater. Rigid vinyl is the leading material, and often requires a conical twin-screw extruder. The molten pipe is solidified by water-cooling. Pipe dimensions are controlled initially by the die, but then finally calibrated by pull-off rate, internal mandrel, vacuum, or compressed air. Flexible tubing is collected on a wind-up unit; while rigid pipe is hauled off by a caterpillar puller to a cut-off saw and stacker.

Profile Extrusion of siding, window frames, gasketing, and other shapes is complicated by the effects of their asymmetry on heating, viscosity, cooling, and dimensional control. This generally requires modification of the die, and vacuum-driven calibration of the extruded product.

Wire Coating extrudes plastic insulation around electrical wire and cable as it passes through a T-shaped cross-head die. The entire process line includes pay-off and capstan to feed and preheat the wire, extrusion-coating, water-cooling, spark testing, diameter and eccentricity controls, and take up capstan and wind-up. The most common "pressure-coating" die applies the plastic coating inside the die; whereas, for larger wire and cable, the "tubing (tool) die" applies the coating as wire and plastic exit from the die, using internal vacuum to pull the molten coating onto the wire. A fine wire coating line can run at speeds up to 10,000 ft/min.

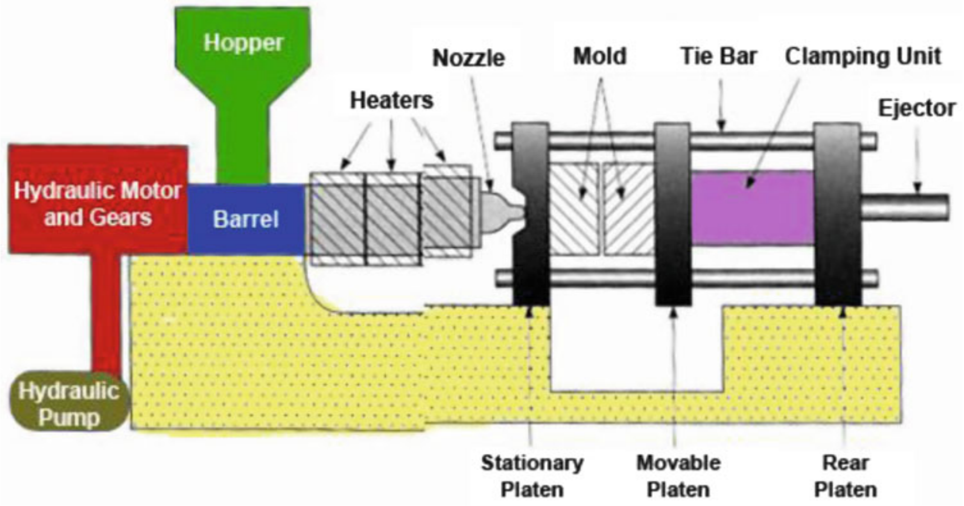


Fig. 15.30 Injection molder. (http://hicksplastics.com/Injection_Molding_Hicks_Plastics_MI.html, Accessed 4 Nov 2011)

Injection Molding

Introduction

Extrusion and injection molding are the two leading methods of converting plastic materials into solid products. Extrusion produces continuous products with a fixed cross-section; injection molding produces discrete products with more complex shapes. Modern injection molding is a very fast, automated process for large-scale manufacture of complex products at minimum cost. An injection molding machine can be used for many different jobs and for many years, so the amortization per job can be very low. On the other hand, it requires a different mold to make each product, so the cost of the mold must be amortized over the life of its individual production run.

Injection Molding Cycle (Fig. 15.30)

The injection molding cycle involves a sequence of events:

- Closing the mold
- Melting the plastic material
- Injecting the melt into the mold
- Filling the mold cavity to form the product
- Cooling the molten product to solidify it in the mold
- Opening the mold and ejecting the solid product

The entire cycle is usually completed within a minute or less.

Closing the Mold. One half of the mold is attached firmly to the melting/feeding end of the injection molding machine. The other half slides horizontally to open or close the mold. The moving half is driven mechanically or hydraulically. When it closes the mold, it is held firmly by clamping action to withstand the injection pressure.

Melting the Plastic Material. The plastic pellets are fed through a hopper into an extruder screw, typically with an *L*/*D* ratio of 20/1. The first half of the screw compresses the pellets and squeezes out air. The third quarter of the screw melts the pellets, 70% by friction, 30% by conduction from heater bands on the extruder barrel. This is called "preplastication." The fourth quarter of the screw pumps the melt forward to the front of the extruder, where it goes through a one-way valve to prevent backflow. As the molten plastic accumulates at the front of the screw, this pushes the screw back to form a reservoir of molten material.

Injecting the Melt into the Mold. When the reservoir contains the desired (preset) volume of molten material, hydraulic force pushes the screw forward rapidly, injecting the melt into the mold cavity. For this reason it is called a "reciprocating screw." In order to fill the mold completely, and make up for shrinkage during cooling with or without crystallization, the screw and the one-way valve maintain high "boost" ("hold") pressure during this part of the cycle. An important auxiliary function is "venting," tiny openings to allow air to escape from the mold, so that the molten material can fill the mold completely and flawlessly.

Filling the Mold Cavity to Form the Product. The mold cavity is designed and machined to form the shape of the finished product. Some major considerations are fast, uniform flow, avoidance of degradation, minimization of orientation/anisotropy, fast cooling/solidification, shrinkage and dimensional tolerances, and of course, final properties of the product.

Cooling the Molten Product to Solidify it in the Mold. The molten product must be solidified before it can be removed from the mold. This is accomplished by flowing cold water through channels machined into the mold. Computerized design of the cooling channels and refrigeration of the cooling water to speed the cooling process, are major considerations. Since heat transfer through organic polymers is slow, the design favors thin-wall products, generally under 1/4 in. Polymers which can crystallize rapidly offer promise of faster, more economical molding cycles. Polymers which crystallize slowly, and amorphous polymers which stiffen gradually, often require longer molding cycles.

Opening the Mold and Ejecting the Solid Product. When the plastic product has cooled sufficiently to be solid and retain its shape, the moving half of the mold is opened automatically. Some products will eject from the mold spontaneously; this depends on design of the mold and the product, low polarity of the plastics vs. the high polarity of the metal mold, and the use of external lubricants to help release the product from the mold. In most cases, however, the moving half of the mold is fitted with "knockout pins." When the mold is opened, the knockout pins automatically project into the mold cavity, and press gently on the product to force it away from the mold surface. Then, as the mold closes for its next cycle, the knockout pins automatically retract again.

Variations and Details

Drying. Although most plastics are quite resistant to water, the powder or pellets may still absorb or adsorb small amounts of moisture during storage and handling. In the heat of molding, this moisture evaporates, causing microscopic voids ("blushing") or macroscopic voids (bubbles), or even hydrolysis to lower molecular weight and less desirable properties. Thus many plastics require pre-drying before molding. They may be dried on shallow trays in a circulating-air oven, or by passing hot dry air through the hopper as they are fed into the extruder. They may be dried by opening a vent midway down the extruder screw to allow the steam to escape.

"Two-Stage" Injection Molding Machine. Instead of using a reciprocating screw to melt the plastic and to push the melt into the mold, the extruder screw may feed the molten plastic into a separate reservoir, and then a separate plunger can feed the melt into the mold.

Electric Drive. This may replace the hydraulic drive in some injection molding machines.

Multicavity Molds. While the mold can be machined with a single cavity to produce a single product, in most cases it is more efficient to machine multiple cavities into the mold and fill them all with a single shot of molten material. This requires a system of runners (tunnels) to distribute the melt to all the cavities, which in turn requires much more sophisticated engineering design to balance them all equally.

Hot Runners. When the molten plastic is pumped into the water-cooled mold, the cooling system solidifies both the plastic product in the mold cavities and also the plastic material in the runners. Later the solid runners must be separated, reground, and reused. This is an extra burden on the process. An alternative is to avoid cooling the runners, and keep them hot, so that the molten polymer in runners remains fluid for the next shot into the mold.

Thermosetting Plastics. While injection molding is used primarily for thermoplastic polymers, it is so efficient and economical that processors working with thermosetting plastics may find it attractive. This has led some of them to adapt the injection molding process to thermosets. This may involve some modification of the reactivity of the thermosetting plastics and/or the injection molding machinery, gentle heating to melt the reactive materials without curing them prematurely, and use of warm or hot molds to finish the curing process quickly and thus speed the molding cycle.

In-Plant Recycling. Injection molding produces a certain amount of scrap: runners, gates, flash, and rejects of imperfect products. In theory, thermoplastic polymers should be perfectly recyclable. In practice, in-plant scrap can be kept clean, reground, and recycled. While 100% recycle is theoretically conceivable, most processors simply determine how high a percent of recycle they can blend with their next batch of virgin material without any harm to their process or product. In most cases they reuse all of their scrap internally, without any contribution to solid waste. This is both good economics and good environmental practice.

Reaction Injection Molding

Overview

RIM is a fast, low-pressure, low-temperature, low-cost process for one-step conversion of reactive liquids into large finished solid plastic products. Liquid polyol and liquid diisocyanate are mixed by impingement, pumped instantly to fill a large mold cavity, and polymerized/cured rapidly to form a thermoset polyurethane product. The cured polymer may be a stiffly flexible product such as automotive bumper covers, front ends, and trim; or a rigid foamed product such as furniture and housings (cabinets) for computers, business machines, TV, and radio.

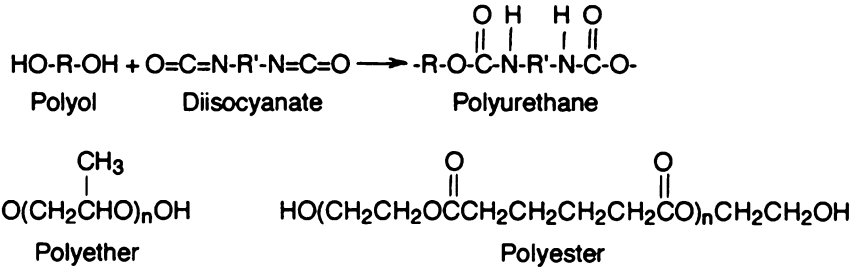
Ingredients

The basic reaction to form polyurethanes is the *polyol*, which may be a polyether or a polyester: Long polyols (high *n*) give flexible polyurethanes; short polyols give stiff, rigid, and/or crystalline polyurethanes. Branched polyols, with three or more –OH groups, give cross-linked thermoset polyurethanes;

short multi-branched polyols give rigid products. Replacing part or all of the polyol by a *Polyamine*

$$H_2N{-}R{-}NH_2$$

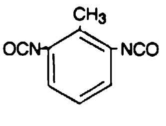
gives faster reaction and therefore shorter cure cycles; and also contributes increased hydrogen bonding, which produces higher rigidity and strength.



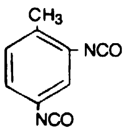
$$\underset{\text{Polyol}}{HO{-}R{-}OH} + \underset{\text{Diisocyanate}}{O{=}C{=}N{-}R'{-}N{=}C{=}O} \longrightarrow \underset{\text{Polyurethane}}{{-}R{-}O{-}\overset{O}{\overset{\|}{C}}{-}\overset{H}{\overset{|}{N}}{-}R'{-}\overset{H}{\overset{|}{N}}{-}\overset{O}{\overset{\|}{C}}{-}O{-}}$$

$$\underset{\text{Polyether}}{HO(CH_2\overset{CH_3}{\overset{|}{C}}HO)_nOH} \qquad \underset{\text{Polyester}}{HO(CH_2CH_2O\overset{O}{\overset{\|}{C}}CH_2CH_2CH_2CH_2\overset{O}{\overset{\|}{C}}O)_nCH_2CH_2OH}$$

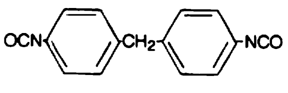
The *diisocyanate* is generally toluene diisocyanate (TDI) or methylene diphenyl isocyanate (MDI) or oligomers based on them.



2,6-TDI



2,4-TDI



MDI

Foaming Agents are either volatile liquids or a trace of water. For flexible products, only a trace of foaming is needed to optimize the product. For rigid products, a higher degree of foaming is used to produce light-weight products. Other ingredients include catalysts to control/balance the polymerization/cure reactions, surfactants for foam uniformity, fillers for stiffening, pigments for coloration, flame-retardants where needed, and short glass reinforcing fibers to increase strength and dimensional stability in Reinforced RIM (RRIM).

Equipment and Process

Raw material holding tanks are warmed to 30–38°C (85–100°F) to keep them at optimum viscosity. The raw materials are measured and pumped by piston pumps, to deliver exactly equivalent amounts of the two reactants, bring them up to 14–21 MPa (2,000–3,000 psi), and shoot them into the mixing head. Here they are mixed by high-speed impingement. The mixed liquid system is very reactive, so it must be pumped into the mold cavity to fill it as rapidly as possible. Molding itself is a low-pressure process, rarely more than 0.7 MPa (100 psi), so the mold can be rather light construction; but since molds and products are generally large, the total force needed to clamp the mold is still considerable. The reaction is exothermic, and the temperature rise could over-cure the product, so mold cooling is required. The entire process is rapid, and can be accelerated further by use of polyamines and catalysts to speed the polymerization/cure reactions. It has proved particularly attractive for mass production of large polyurethane parts.

Structural Foam

Definition

When a solid plastic is expanded to moderately lower its density, it is called a "structural foam." Bone and wood are natural structural foams, which benefit from this moderate degree of expansion. Most plastic materials can similarly be expanded during melt processing, and also gain many benefits as a result.

General Description

A blowing agent is added to the plastic material, to compress gas into the melt during processing. It can be compressed nitrogen gas, which is inexpensive but difficult to disperse; or it can be a "chemical blowing agent," usually an organic azo compound RN=NR, which gives much more uniform dispersion in the melt, and reacts to produce compressed gases in the hot melt.

In extrusion, when the hot melt exits from the die, the compressed gas expands and foams the extrudate. In injection molding, the hot melt is injected into the mold, but the quantity is reduced ("short shot") so that there is not enough melt to fill the mold; the compressed gas expands, and the melt foams and fills the mold. Either way, the foamed melt cools and solidifies, producing a product with a solid skin and a somewhat expanded internal foam structure. It appears to be a solid product, but it is lighter in weight and offers a number of advantages.

Degree of Expansion

Plastics with solid densities of 0.9–1.2 are generally expanded 13–25%, giving foamed densities of 0.75–0.9. At this modest degree of expansion, they look and perform like solid products.

Benefits

Use of 13–25% less material is an obvious economic advantage. Low pressure in the mold (1.4–3.4 MPa, 200–500 psi) permits construction of a lighter mold and operation at lower pressure, both which provide further economic advantage. A low degree of foaming is enough to expand the molding against the walls of the mold, eliminating any accidental "sink marks." Higher degree of foaming increases rigidity 3–4× over a solid part of equal weight. Low-pressure molding also eliminates molded-in stresses, which would weaken the product and could also cause post-molding warpage. This in turn permits design and production of more complex parts in a single molding step, eliminating need for later machining and assembling.

Problems in Structural Foaming

Foamed products tend to be thicker, and contain gas bubbles. Both of these factors reduce heat transfer, taking longer to cool and solidify the product, and thus lengthening the manufacturing cycle.

Tiny foam bubbles in the skin of the product tend to form swirl patterns which are unsightly in consumer products. Several modified injection molding processes offer promise for reducing or eliminating this problem: (1) The *Farrel/USM Process* injects the melt into an expandable mold at high pressure, then expands the mold to lower the pressure and permit foaming. (2) The *Coinjection Process* uses two reciprocating screws to feed the mold; the first feeds a nonfoamable short shot, which will form the solid skin, and then the second feeds a foamable second short shot which expands, pushing the skin ahead of it and filling the interior with foam. (3) The *Gas Counterpressure Process* seals the mold, compresses nitrogen gas into it, injects a short shot of melt containing blowing agent, and then uses gradual controlled venting to release the nitrogen gas and allow the foamable melt to expand, forming a solid skin and a foamed core.

Low-Density Foams

When plastics are foamed to low densities, containing more air than polymer, they have unique properties and applications. Major uses are in crash padding and thermal insulation. Closed-cell foams are outstanding for flotation, rigidity, and insulation; while open-cell foams are outstanding for softness, resilience, and comfort.

While foam production and properties may all belong to one unified theoretical basis, commercial practice is different for each of the major polymer families; so it is best to review them individually.

Polyurethane

The largest quantity of low-density foam is polyurethane, made by mixing liquid polyols with liquid diisocyanates, pouring the mixed liquid and allowing it to foam, polymerize, and cure rapidly into its finished form. The auxiliary ingredients—foaming agent, tertiary amine and organotin catalysts, surfactant, and optionally flame-retardant—are mixed with the polyol prior to mixing with the diisocyanate. The two liquid streams are metered by piston or gear pumps, and fed at high pressure into an impingement chamber where they mix by turbulence. This mixture is very reactive, so it is poured rapidly into the desired form or location, and allowed to polymerize, foam, and cure into the finished product. The largest amount is open-cell soft flexible foam; a smaller amount is closed-cell rigid foam for thermal insulation.

Flexible Foam is made from long-chain diols with a small amount of triols for cross-linking to give strength and resilience. Foaming is produced by reaction between measured amounts of isocyanate and water to liberate carbon dioxide. Molded products are made by pouring the reactive liquid mixture into a mold cavity at 50°C. They foam and cure in 2–10 min, after which they are temporarily crushed or vacuum-shocked to open the cell walls and insure softness. Then they are allowed to condition for several hours to finish the cure. The major products are auto seating and headrests, and furniture cushions.

Large slabstock is made by pouring the reactive liquid mixture into a moving paper form, up to 8 × 5 × 50 ft long, which takes up to 5 min. Polymerization, foam rise, and cure are exothermic, and the interior of the slab can reach 140–170°C. The rough surface is trimmed off, and the slab is conditioned 12–24 h to finish the cure. The finished slab is sliced to the desired thickness for furniture, mattresses, public transportation seating, textile backing (sportswear), carpet backing, and packaging.

Rigid Foam is made from short-chain polyols with a typical average functionality of four hydroxyls per molecule, to give high cross-linking for rigidity. Foaming is produced by volatile organic solvents, which are boiled by the exothermic polymerization/cure reaction. Sheet, slab, and block are made by pouring the reactive liquid mixture onto a moving paper form, or directly onto facing panels for laminated structures. The Pour-in-Place technique is used to fill the wall cavities of refrigerators and freezers, holding them at about 40°C to control the reaction. Spray coating insulation is applied to buildings, tanks, and pipes in the field, using a highly catalyzed mixture that will foam and gel in less than 10 s, so that it will stay in place without running down; the coating can be built up to 2 in. thick in a single pass, more by repeated passes.

Rigid polyurethane foam is used primarily as thermal insulation for buildings, trucks, rail cars, shipping containers, tanks, pipelines, cold-storage warehouses, and frozen food display cases.

Polystyrene

Polystyrene is foamed by swelling with pentane, heating to soften the polymer and vaporize the pentane, and allowing it to expand at atmospheric pressure. Extrusion produces foamed sheet and board stock, which are used mainly as thermal insulation in commercial refrigerators and freezers, and also in food packaging, roof and wall insulation, and pipe insulation.

Moldable beads are produced by suspension polymerization, swelling with pentane, warming to soften and "pre-expand" the beads, pouring them into a mold, and steaming to expand them fully, soften them, and fuse them together into a finished product. This produces drinking cups, molded packaging, board stock, and display and novelty products.

Loose fill for packaging is produced by extruding and chopping polystyrene into various shapes, swelling with pentane, and steaming to expand them into "peanuts."

Polyvinyl Chloride

Flexible PVC foams are generally laminated to layers of cloth and/or solid vinyl, and used for leather-like clothing and luggage, upholstery in autos and furniture, and resilient flooring. They are made by polymerizing vinyl chloride in emulsion, spray-drying to a fine powder, mixing into liquid plasticizer at room temperature to form a viscous paste, compounding with stabilizers/catalysts and azodicarbonamide foaming agent, roller-coating onto a moving belt of textile or paper or solid vinyl film, oven-heating to dissolve the resin in the plasticizer and activate the foaming agent, and cooling to room temperature to gel the plasticized vinyl and stabilize the flexible foam. An alternative for large-scale production is to compound general purpose PVC with plasticizer, stabilizer/catalyst, and foaming agent, calender and/or laminate sheet below the activation temperature of the azodicarbonamide, and then pass it through a hotter oven to foam the sheet.

Polyethylene

Polyethylene foam sheet is used mainly in package cushioning, and also in roof insulation. Extruded profile is also used as construction sealants and pipe insulation. It is made mainly by tandem extrusion. The first extruder melts the polyethylene and blends in liquid volatile organic foaming agent. The second extruder cools it to the optimum foaming temperature, and pumps it through a die into room-temperature air, where it expands into foamed sheet.

For higher performance, cross-linked foamed sheet can be produced by radiation or chemical cross-linking. For radiation cross-linking, polyethylene is melted in an extruder and mixed with powdered chemical foaming agent, extruded into sheet, cross-linked by electron beam radiation, and foamed in an oven. For chemical cross-linking, polyethylene is melted in an extruder and mixed with peroxide cross-linking agent and chemical foaming agent, extruded into sheet, and passed through a two-stage oven for cross-linking and foaming.

Molded foam for package cushioning, flotation devices, and sports equipment is made from foamed pellets. Polyethylene pellets are foamed with volatile organic liquid and cross-linked by peroxide or electron beam. The foamed pellets are poured into a mold, the mold is closed, and the pellets are softened and expanded further to fuse into a molded foam product.

Blow Molding

Blow molding uses compressed air to blow and expand a hot plastic tube ("parison") in a female mold cavity, until the plastic conforms to the walls of the mold. This has developed into a major way to produce plastic bottles, and also more specialized hollow shapes such as fuel tanks, seat backs, tricycles, surfboards, etc. The leading material is HDPE for milk, water, and household chemicals. Second largest is PET for carbonated beverages. A number of other thermoplastics are blow-molded in smaller amounts for more specialized uses.

There are two ways to produce the plastic tube ("parison"): injection molding and extrusion. For injection blow molding, a test-tube shape is first injection molded, then transferred into a bottle mold and blown. For extrusion blow molding, the extruder produces the tube continuously, and a rotating or alternating series of bottle molds take turns clamping around the tube and blowing it. Injection blow molding is used for bottles up to 500 mL in size, while extrusion blow molding is used for bottles 250 mL and larger.

Stretch blow molding holds the parison above its glass transition temperature (T_g) and stretch-orients it to increase modulus, strength, impact resistance, transparency, and impermeability. This is most important for PET, and is also used for PVC, polypropylene, and polyacrylonitrile.

Multilayer blow molding uses a parison containing two or more polymers in concentric layers, and produces a multilayer laminated bottle or fuel tank. In general this can combine the best properties of each layer. Sandwich structure food packaging, with a mid-layer of ethylene/vinyl alcohol copolymer (EVOH), provides impermeability to oxygen. Sandwich structure fuel tanks of HDPE/nylon/HDPE gain impermeability to gasoline.

To speed the cooling portion of the blow-molding cycle, chilled air can help. For more extreme action, liquid carbon dioxide or nitrogen can give even faster cooling.

Surface fluorination has been claimed to make polyethylene less permeable to gasoline and nonpolar solvents in general. Fluorine gas can be used to blow the container, which treats only the inside surface or the finished container can be fluorinated both inside and outside in a single post-treatment. This produces a fluorinated layer 20–40 nm thick.

Thermoforming

Thermoplastic sheet can be softened by warming, placed in or on a mold, pressed gently but swiftly to conform against the walls of the mold, cooled to solidify, and trimmed to separate the product from the surrounding unused sheet.

Sheet can be produced by extrusion or calendering. For use in thermoforming, the sheet must be extremely uniform. For improved properties, the sheet may be biaxially stretch oriented up to 300%. For large automatic production runs, it is fed continuously off a roll; for short runs, large, and/or specialty jobs, it may be cut into individual sheets and fed one at a time.

The sheet must be preheated to make it soft and pliable ("rubbery melt"). For fastest processing and best final uniformity and properties, it should be as hot as possible, without losing melt strength or beginning thermal degradation.

There are three methods of preheating. (1) Convection oven is slow, but very uniform. (2) Conduction heating is done with electrically heated Teflon-coated aluminum plates. (3) Radiant heating is most efficient, especially when the infrared wavelengths are chosen to match the maximum absorption frequencies of the plastic material. There are three ways to press the warm sheet gently and swiftly against the surface of the mold. (1) Vacuum forming is the most popular, pulling the soft sheet against the mold surface, with atmospheric pressure applying the driving force. The vacuum must be at least 25 in. Hg to give the fastest possible cycle. (2) Pressure forming (compressed air), at pressures of 140–860 kPa (20–125 psi), is faster and gives better final properties. (3) Mechanical force (tensile, flexural, or compressive) can also drive the warm sheet to conform to the mold surface. In some cases, two or all three of these methods may be combined for optimum performance.

The best molds are made of aluminum with water-cooling channels cut into them. Tooling costs are low, and heat transfer is high. The formed plastic sheet must be held against the mold surface until it is cooled to the solid state. This is often judged by the HDT at 455 kPa (66 psi).

Whether the process uses continuous sheet or individual sheets, the product is formed from the center of the sheet, and the edges must be trimmed off to give the final product. Die cutting knives are common, but a variety of other mechanical and thermal techniques are also in use. The trim may be 10–70% of the original sheet. It is reground, blended up to 50% with virgin resin, and recycled directly into the same process. Considering the variability of recycle and blending, this requires careful control to maintain quality.

Overall, thermoforming is a very useful method of fast low-cost production for a great variety of plastic products, from packaging to building to automotive parts.

Rotational Molding ("Rotomolding")

Hollow products can be made by placing powdered plastic in a closed mold, tumbling and heating it until the plastic has coated the walls of the mold, cooling it to solidify the product, and opening the mold to remove the product.

Polyethylene is by far the most popular material. Others used occasionally include vinyl plastisol (liquid rather than powder), nylon, polypropylene, and polycarbonate. The powder is usually ground to 35 mesh.

The most popular machine is a carousel design with 3–4 arms operating independently, providing separate stations for loading, heating, cooling, and unloading. Molds are generally made of aluminum, two-piece plus clamps.

Slow rotation gives the most uniform products. The pattern of biaxial rotation must be determined by trial and error. Heating is best done in a gas-fired oven with a fan for circulation.

Cooling is a two-stage process, first a fan and then a cold water spray or mist. Advantages are low capital investment, design freedom, strain-free products, and little or no post-molding secondary operations. For initial costing, multiply the cost of the raw material by 5. Blow molding is faster and uses a wider variety of plastics.

Typical products are tanks (20–85,000 L, 5–22,500 gal), containers for packaging and materials handling, portable outhouses, battery cases, light globes, vacuum cleaners, garbage cans, surfboards, toys, traffic barricades, display cases, and ducting.

Powder Coating

A solid plastic may be ground to a solid powder and then used to apply a plastic coating to a metal product, either for decorative reasons (color) or for functional performance such as insulation or corrosion resistance. Plastics most often used in this way include nylon, vinyl, acrylic, polyethylene, polypropylene, and epoxy. Coating without solvents is beneficial both economically and environmentally, and 100% utilization of material eliminates waste. After the powder is bonded to the metal surface, it is often reheated to flow into a more uniform coating and, in the case of thermosetting resins, to complete the cure reaction.

There are three techniques for applying the powder coating to the surface of the metal product: fluid bed, electrostatic fluid bed, and electrostatic spray.

Fluid Bed Coating

The equipment is simply a horizontal box with a finely-porous shelf near the bottom. Powdered plastic, ground to 20–200 μm, is poured above the shelf. Compressed air is fed in below the shelf, percolating up through the pores, and percolating the powder so that it rises and flows much like a liquid. Sometimes the box is also vibrated to produce greater uniformity.

The metal product to be coated is preheated to a temperature which will melt the powder. Then the product is dipped into the fluid bed. The powder particles melt and flow onto the metal surface. Coatings up to 2.5 mm (0.1 in.) thick can be applied in a single dip.

Typical products are electric motors, electronics, transformers, valves, pumps, refinery equipment, and appliances.

Electrostatic Fluid Bed

In a modification of the fluid bed technique, the product to be coated, instead of being preheated, is passed over the bed, and the powder is attracted to it by a static charge. Then it is passed through a heating oven to fuse it into a finished coating. This process can be used for either discrete or continuous products.

Electrostatic Spray

Conventional electrostatic spray coating pumps a liquid coating formulation through a spray gun, which puts an electrostatic charge on the liquid, and sprays charged droplets toward a grounded metal product. The droplets are attracted to the product, where they discharge and adhere to the metal surface.

Plastic powders can be applied in a similar way. The solid plastic is ground to a 30–200 μm powder, fluidized, and conveyed by compressed air through a spray gun which uses high voltage to apply an electrostatic charge to the particles, and then sprays them at the grounded metal product. Spraying is normally done in a hood, to protect the worker and to collect and recycle the excess powder. The charged powder is attracted to the metal product, attaching a coating 50–75 μm thick on a cold product, or up to 250 μm on a preheated product. A post-fusion step melts and flows the powder into a uniform coating and, in the case of thermosetting resins, also completes the cure reaction.

Calendering

In calendaring a viscous bank of excess thermoplastic or rubber is pressure-formed between a pair of parallel corotating rolls to form a thin "film" or a thicker "sheet." Most calenders consist of four rolls in L- or Z-shape, plus additional feed rolls and post-calender laminating, patterning, and/or wind-up rolls.

PVC is a commonly calendered plastic, especially in plasticized flexible form. Other materials often calendered are rubber, ABS, polyurethane, and thermoplastic elastomers.

The plastic material delivered to the calender must be molten, homogeneous in composition and temperature, and at optimum viscosity for the calendering operation. For rigid vinyl, optimum temperature may be 180–190°C; for flexible vinyl, 10–20°C lower. Nip pressures between the rolls are typically 160–1,050 kN/m (900–6,000 lb/in.) of roll face.

A surface pattern can be applied by the calender roll or by a post-embossing operation. Feeding hot sheet to a cold embossing roll works well. The pattern is measured by a profilometer. Calendering can also be used to apply a plastic surface to a substrate web such as metal, cloth, or paper.

Capital investment for a calender line may be about $5 million. Operating cost may be about $500/h. Calendering is a very high-speed method of producing plastic film and sheet, but high-speed economy must be balanced against quality of the finished product.

Vinyl Plastisol Processing

Plastisol processing is a specialized technique for making flexible vinyl products. It is also sometimes called "paste" or "dispersion" technology. In this method powdered PVC is stirred into liquid plasticizer at room temperature to form a viscous liquid or spreadable paste. This can be poured or spread into the shape of the final product, heated to fuse, and cooled to gel into the final solid product. Simple pouring or gentle spreading requires no heavy equipment, and does not damage delicate substrates, such as paper and cloth. Consequently, this accounts for about 10% of all vinyl production.

About 90% of PVC is made by suspension polymerization and processed as viscous melt at high pressure in heavy-duty equipment. Some 10% of PVC is made by emulsion polymerization and spray-drying, for use in plastisol processing. The resin is stirred into liquid plasticizer at room temperature, along with stabilizer and other optional additives, to form a viscous liquid or easily spreadable paste. This is poured or spread into the form of the desired final product. It is then heated, passing through several phases. At first the resin particles absorb the liquid plasticizer, swell, and form a gel. On further heating, the resin melts and dissolves in the hot plasticizer, forming a hot solution. On cooling, the resin crystallizes somewhat, turning the hot solution into a hot gel. On cooling to room temperature, the gel solidifies to a firm rubbery plastic.

If the plastisol is too viscous for pouring at room temperature, it may be thinned with a volatile solvent; on heating, the solvent evaporates, leaving a firmer final product. This is called "organosol" technology. Alternatively, a polymerizable plasticizer may be added to thin the plastisol; on heating, it polymerizes to a solid ingredient, giving a firmer final product. This is sometimes called a "rigisol."

If molds are required to shape the plastisol into a finished product, they are generally light-weight low-cost molds of aluminum, electroformed nickel, or ceramic. Oven-heating is slow, but common. The major handicap is slow heat transfer, so overall production cycles may take 4–20 min. Several different methods are described below for converting plastisols into finished products.

Dip Coating. A product is preheated and dipped into the plastisol. The plastisol gels onto the surface of the product. It is withdrawn from the liquid, allowed to drain off excess liquid, and then placed in a 191–204°C oven to heat it in a few seconds to about 177°C and fuse the gel to a homogeneous solution. It is then cooled in air or water to solidify the material into the final rubbery coating. Typical products are tool handles, kitchen implements, and electrical insulation.

Dip Molding. Instead of a heated product, a heated male mold is dip-coated in the same way. In this case, the final flexible vinyl product is stripped off of the mold. A typical product is medical gloves.

Slush Molding. This is the reverse of dip molding. A female mold is used to give any desired surface finish on the product. The mold is preheated, plastisol is poured into it, and gels onto the surface of the mold. The excess liquid plastisol is poured out, and the internally coated mold is heated to gel, cooled to solidify the rubbery product, and the product is stripped out of the mold. Typical products are arm rests, head rests, road safety cones, anatomical models, dolls, toys, and auto parts.

Hot–Melt Molding. Hot molten plastisol is injected into a cooled mold to make products like fishing baits and novelties.

Rotational Casting. Like rotomolding, the plastisol is poured into a cold mold, which is rotated and heated to gel the plastisol onto the walls of the mold. It is then heated to fuse, cooled to solidify, opened, and the product removed. Typical products are volley balls, basketballs, dolls, and auto parts.

Open Molding. The plastisol is poured into an open mold, heated to gel and fuse, cooled to solidify, and stripped out from the mold. Typical products are auto air filters, oil filters, truck flaps, and place mats.

Closed Molding. In this process, the plastisol is filled into a closed mold, heated to gel and fuse, then cooled to solidify. A typical product is switch mats for automatic door openers.

Spray Coating. Liquid plastisol can be spray coated up to 1.25 mm (50 mils) in a single pass. The plastisol is formulated to be liquid enough to spray, then firm enough to hold on a vertical surface without running down. It is heated to gel and fuse, then cooled to form the final rubbery coating. A typical product is tank linings.

Continuous Coating. Doctor blade or roll coating applies plastisol continuously to a moving web of metal or cloth, or two successive coatings apply a solid vinyl skin and a foamable vinyl core to a fabric backing. Typical products are house-siding, conveyor belting, and resilient flooring.

Silk-Screen Inks. Plastisol coatings can be applied a few mils thick on cloth to produce T-shirts and athletic uniforms.

Organosol Coatings. The low viscosity of organosol formulations permits coatings under 10 mils thick, particularly for chemical resistance.

Liquid Casting Processes

Reactive liquids may be mixed and poured to form a solid plastic product. They may be two co-reactive monomers or prepolymers, or one monomer/prepolymer, plus a catalyst or curing agent. They are primarily thermosetting plastics and elastomers. Rigid plastics are primarily epoxy, polyester, and acrylic. Elastomers are generally room-temperature vulcanizing (RTV) polymers: polysulfide, polyurethane, and silicone. They are used mostly to provide electrical insulation and mechanical and environmental protection for delicate electrical/electronic products, such as resistors, coils, solenoids, capacitors, transformers, printed circuit boards, optoelectronics, and light-emitting diodes. There are a number of such liquid casting processes, generally carried out at room temperature with or without a later heating step for complete cure.

Casting. The reactive liquid mixture is poured into a mold, and allowed to polymerize and cure. This is used to make solid tires for industrial vehicles, and to make decorative items, such as simulated wood trim, furniture, picture frames, and lamp bases.

Potting. Delicate electronic devices are placed inside an empty shell, and the space is filled by pouring the reactive liquid mixture into it and curing it.

Encapsulation/Embedment. After potting is completed, the outer shell is removed, and the solid plastic is the container.

Conformal Coating. Using a thixotropic reactive liquid, the product is dipped into the liquid and removed with a coating of the liquid on it. This is then heat-cured to solidify the coating and make it permanent.

Impregnation. A porous product can be dipped into a fluid reactive liquid, which soaks into the pores and fills them completely. The liquid is then cured to leave the product completely filled with solid polymer in the pores.

Syntactic Foam. Hollow glass, ceramic, or plastic spheres are dispersed in the reactive liquid system before it is cast. When the liquid is polymerized and cured, the hollow spheres make it a unicellular foam. The air "bubbles" in the cells make it low-density, low dielectric constant and loss, and very resistant to compressive forces such as hydrostatic head in deep-sea equipment.

Aside from simply casting at atmospheric pressure, vacuum is often used to remove air bubbles and volatiles, which would give an imperfect casting, or to help promote porous impregnation. Small specialty runs can be done manually, with no capital investment but high labor cost; larger production runs can be automated with only modest capital investment. Since many of these liquid systems may be volatile and unpleasant to work with, worker protection often includes ventilation and/or protective clothing for eyes, skin, and lungs.

Compression Molding and Transfer Molding

Thermosetting plastics and rubber are heated to soften ("plasticate") them, and then pressed at 14–28 MPa (2,000–4,000 psi) to cross-link ("cure") them. Then the mold is opened and the finished product is removed.

Properties. Compared to engineering thermoplastics, thermosets can offer equal or superior properties at lower cost. Major advantages include rigidity, creep resistance, dimensional stability, impact strength, heat resistance, and chemical resistance.

Typical Applications. Common compression molded thermoset products include electrical equipment, appliance handles and knobs, dinnerware, distributor caps, under-the-hood parts, automatic transmissions, brakes, and pumps.

Compression Molding. This press has two horizontal platens facing each other. The upper one is generally stationary; the lower one moves vertically, driven by mechanical, hydraulic, or pneumatic power. The upper mold half is fastened to the upper platen, the lower mold half to the lower platen. Most molds are electrically heated. In semiautomatic operation, the operator puts the granular molding powder in the lower half of the mold, and presses a button which closes the press, holds it till cure is complete, and opens the press again; then the operator removes the molding from the mold, and repeats the cycle. In fully automatic molding, the entire cycle is automatic and does not require an operator.

Transfer Molding. This preheats the molding powder in the upper part of the mold. Then a plunger pushes the fluid material down into the (lower) mold cavity and holds it there until cure is complete. This gives more fluid flow than compression molding, and avoids damage to delicate molds or molded parts.

Additional Considerations. Compared to injection molding of thermoplastics, compression molding is less capital intensive, more labor-intensive, and takes a longer molding cycle. Transfer molding can equal injection molding cycle times. Compression and transfer molding do less damage to reinforcing fibers, so they can give more impact-resistant products.

The molding cycle can be shortened by preforming and preheating. The granular molding powder is cold-pressed into a pill of the desired weight, and then preheating brings it to near molding temperature before it is inserted into the mold. This gives better flow and faster cure.

"Venting" and "breathing" must often be included in the molding cycle. Venting allows trapped air to escape through small grooves as the molding powder is compressed and flows into a solid part. Breathing opens the mold about 3 mm (in.) for a second or two, early in the molding cycle, to allow escape of water and other volatiles from the condensation cure reaction. This permits production of bubble-free parts. Ejector pins help to separate the finished molding from the mold cavity. These are driven mechanically during the mold-opening stage of the cycle. Post-cure ovens may be used to help complete the cure cycle, especially to control gradual cooling of the molded part. Deflashing is often required to remove excess trim ("flash") from the edges of the molded product. This may be done by tumbling in a rotating drum, or by an air-blast grit.

Reinforced Plastics Processing

Adding short-fiber reinforcements to thermoplastics can produce major increases in modulus, strength, and dimensional stability; it makes melt processing more difficult. Adding fibrous reinforcement to thermosetting plastics produces outstanding modulus, strength, impact strength, dimensional stability, and heat and chemical resistance; but requires entirely new methods of processing to convert them into finished products.

The polymers most often used are unsaturated polyesters and epoxy resins; for ultra-high performance, polyimides and other specialty resins are also used in small amounts. The reinforcing fibers are primarily glass; for ultra-high performance, aramid, carbon/graphite, and metallic fibers are also used.

The polymers are generally polymerized to low molecular weight "A-stage" resins, which are still liquids or fusible solids, and potentially very reactive. In pure form they can

Table 15.22 Typical mix for a bulk molding compound (BMC)

Resin	22%
Fiber	25%
Filler	53%

be stored for months to years. The liquids are easily mixed with reinforcing fibers and other ingredients, and shaped into the form of the ultimate products. Addition of peroxide initiators to polyesters, or amine or anhydride curing agents to epoxies, liberates their high reactivity to permit fast cure cycles to completely cure products. Some processes use heat and/or pressure to complete the shaping and cure cycle, while others are carried out simply at ambient conditions. These variables offer a wide variety of processes which can be used to manufacture reinforced thermoset plastic products.

Matched Die Molding Processes

By analogy with compression molding, thermosetting resins, plus fibrous reinforcement can be pressed between matching mold halves, with heat and moderate pressure, and cured directly into finished products. A typical cure cycle is about a minute. Half of all reinforced thermoset products are made in this way. Major applications are in the automotive, electrical/electronics, appliance, and business machine industries.

Bulk Molding Compound (BMC). Liquid resin, 1/4–1/2 in. long glass fiber, and a simple filler (typically calcium carbonate or clay) are mixed in a dough mixer or cold extruder. A group II metal oxide is added to gel the liquid resin (Table 15.22). This doughy mix ("premix") is placed in a hot mold, pressed to flow and fill the mold cavity, and held until cured.

Sheet Molding Compound (SMC). The SMC machine pours liquid resin and 1-in. glass fiber onto a moving belt, passes it through calender rolls to make a good sheet, and then through an oven to begin cross-linking and gel the resin. The sheet is placed in a hot mold and pressed to shape and cure it. While the short fiber in BMC is easier to mold, the longer fiber in SMC gives greater strength and impact resistance.

Cold Press Molding. Occasionally the mix may be placed in a cold mold and pressed at about 345 kPa (50 psi). The cure cycle is considerably longer.

Preform. A metal screen is made in the shape of the final product. Glass fiber is chopped 2 in. long and sprayed uniformly over the shaped screen, using vacuum on the back side of the screen to assist the process. A small amount of binder, typically 5% of polymer in latex form, is sprayed onto the fiber to hold its shape. It is then removed from the screen, placed in the mold, saturated with an equal weight of liquid resin, and the mold is pressed at 1,380 kPa (200 psi) and heated until cured, typically 3–15 min. This early process has been largely replaced by SMC.

Prepreg. Fabric is impregnated with 25% of liquid resin and laid in the mold. To insure isotropic properties, or to maximize properties in a specific direction, successive layers of impregnated fabric are carefully oriented in different directions. The mold is closed, pressed, and heated till cured. Products made from such impregnated fabrics have much higher strength than simple random fiber reinforcements.

Resin Transfer Molding (RTM). Reinforcing fibers are distributed uniformly in the mold and the mold is closed. Liquid resin is injected into the mold until the excess comes out of the vents. The mold is pressed and heated, similarly to preform molding, until cure is complete.

Open Molding

Instead of applying pressure in a closed mold, the mix of liquid resin and reinforcing fiber may be laid into an open mold, and optionally pressed gently at room temperature until cured. To accelerate low-temperature cure, more active catalyst systems are added. Alternatively, the assembly may be UV or oven-cured. This requires less capital investment, but more skilled labor, so it is useful for prototype and small production runs. It permits unlimited size, so the largest reinforced thermoset products are made in this way, for example, large tanks and whole boat hulls.

Hand Lay-Up (Contact Molding). A layer of liquid resin is applied to the surface of the mold. A layer of glass fiber mat (low strength) or fabric (high strength) is hand laid over it. Liquid resin is poured over it, and brushed or rolled (squeegeed) into it. The process is repeated to build up the desired thickness of the product. The assembly is allowed to stand until cured.

Spray-Up. Instead of hand lay-up, continuous glass roving and liquid resin are fed into a gun, which chops the glass fiber, mixes it with the resin, and sprays it into the mold. This can be automated for lower labor cost and greater uniformity. Spraying is often followed by hand-rolling to expel air and densify the assembly. Then it is allowed to stand until cured. Products are similar to those from simple hand lay-up.

Vacuum-Bag Molding. After hand lay-up, the assembly is covered with an air-tight film, typically polyvinyl alcohol, occasionally nylon or other material. Then a vacuum is pulled on the underside of the film, to squeeze out air and excess resin. Use of a hand-held paddle may help. The assembly is allowed to stand until cured. Products are void-free, and quality is better than simple hand lay-up.

Pressure-Bag Molding. This is similar to vacuum-bag molding, but 345 kPa (50 psi) air pressure is applied to the outside of the cover film. A rubber bag may be used to facilitate the process.

Autoclave Molding. This is similar to pressure-bag molding, but it is carried out in an autoclave to apply the pressure to the outside of the cover film.

Rubber Plug Molding. Liquid silicone rubber is cast into a female mold cavity and cured. The plug is removed, and layers of mat or fabric are built onto the plug. This assembly is then inserted into the mold cavity, pressed, and heated until cured.

Special Processes

Filament Winding. This requires a mandrel to shape the desired finished product. Continuous filament or woven tape is fed through a liquid resin bath to impregnate it, and then wound onto the mandrel in a calculated pattern to optimize the final properties. The assembly is oven-cured. A collapsible mandrel can then be removed from the plastic product; or the mandrel can be left as a part of the finished product. Typical products are pipes, tanks, and pressure bottles. Other suggested products include rocket motor cases, railroad hopper cars, automotive springs, drive shafts, ship hulls, housing modules, helicopter rotor blades, and helicopter tail sections.

Pultrusion. Filaments, woven tapes, or fabrics are fed continuously through a liquid resin bath, through a shaping die, through a curing oven, and then cut to any desired length. This makes continuous products of any desired cross-section. Typical products include fishing rods, flag poles, tool handles, ladder rails, tubing, and other electrical, corrosion resistant, construction, and transportation applications.

Continuous Laminating. Reinforcing fiber and liquid resin are deposited on a moving belt, densified between squeeze rolls, passed through a curing oven, and cut to length. This process is used for mass production of glazing, paneling, and roofing.

Centrifugal Casting. Reinforcing fibers are distributed inside a circular mold. The mold is rotated, and liquid resin is distributed inside it to impregnate the fibers. Rotation of the assembly is continued inside an oven until it is cured. This process is used for making pipes, tanks, and hoops.

Foam Reservoir Molding. Flexible open-cell polyurethane foam is impregnated with liquid resin, faced with glass fiber mat, and gently compression molded. This squeezes the liquid resin into the glass fiber surface mat. Heat curing produces a sandwich structure of low density, high flexural and impact strength.

References

1. http://www.merriam-webster.com/dictionary/plastic. Accessed 23 Jan 2011
2. http://www.americanchemistry.com/2009-year-in-review
3. http://www.americanchemistry.com/production-and-sales-data-by-resin-2009-vs-2008
4. Sperling LH (2001) Introduction to polymer science, 3rd edn. Wiley, New York, p 2
5. Rudin A (1999) The elements of polymer science and engineering, 2nd edn. Academic, San Diego, pp 132–141
6. Rudin A (1999) The elements of polymer science and engineering, 2nd edn. Academic, San Diego, p 150
7. Billmeyer FW (1971) Textbook of polymer science, 2nd edn. Wiley, New York, p 16
8. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, pp 70–74
9. Nielsen LE, Landel RF (1994) Mechanical properties of polymers and composites. Marcel Dekker, New York, pp 18–23
10. Nielsen LE, Landel RF (1994) Mechanical properties of polymers and composites. Marcel Dekker, New York, p 21
11. Nielsen LE, Landel RF (1994) Mechanical properties of polymers and composites. Marcel Dekker, New York, pp 17–18
12. Painter PC, Coleman MM (1997) Fundamentals of polymer science, 2nd edn. Technomic Publishing, Lancaster, pp 284–290
13. Nielsen LE, Landel RF (1994) Mechanical properties of polymers and composites. Marcel Dekker, New York, pp 50–51
14. Sperling LH (2001) Introduction to polymer science, 3rd edn. Wiley, New York, pp 230–233
15. Billmeyer FW (1971) Textbook of polymer science, 2nd edn. Wiley, New York
16. Billmeyer FW (1971) Textbook of polymer science, 2nd edn. Wiley, New York, pp 270–271
17. Carraher CE Jr (1996) Polymer chemistry, 4th edn. Marcel Dekker, New York, p 265
18. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, p 206
19. Odian G (1981) Principles of polymerization, 2nd edn. Wiley, New York, p 182
20. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, pp 208–209
21. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, p 213
22. Billmeyer FW (1971) Textbook of polymer science, 2nd edn. Wiley, New York, pp 317–318
23. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, p 222
24. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, pp 223–225
25. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, pp 234–235
26. Carraher CE Jr (1996) Polymer chemistry, 4th edn. Marcel Dekker, New York, p 277
27. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, pp 237–245
28. Carraher CE Jr (1996) Polymer chemistry, 4th edn. Marcel Dekker, New York, p 282
29. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, pp 246–249
30. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, p 173

31. Billmeyer FW (1971) Textbook of polymer science, 2nd edn. Wiley, New York, pp 355–357
32. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, p 174
33. Shaffer KA, DeSimone JM (1995) Trends Polym Sci 3:146
34. Billmeyer FW (1971) Textbook of polymer science, 2nd edn. Wiley, New York, pp 358–359
35. Billmeyer FW (1971) Textbook of polymer science, 2nd edn. Wiley, New York, pp 359–361
36. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, p 175
37. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, pp 194–195
38. Billmeyer FW (1971) Textbook of polymer science, 2nd edn. Wiley, New York, p 231
39. Billmeyer FW (1971) Textbook of polymer science, 2nd edn. Wiley, New York, p 227
40. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, p 227
41. Stevens MP (1999) Polymer chemistry, 3rd edn. Oxford University Press, New York, pp 273–275
42. Sperling LH (2001) Introduction to polymer science, 3rd edn. Wiley, New York, pp 342–343
43. Birley AW, Haworth B, Batchelor J (1992) Physics of plastics. Carl Hanser, Munich
44. Williams ML, Landel RF, Ferry JD (1955) The Temperature Dependence of Relaxation Mechanisms in Amorphous Polymers and Other Glass-forming Liquids, *J Am Chem Soc* 77:3701
45. Powell PC (1983) Engineering with polymers. Chapman and Hall, London
46. Birley AW, Haworth B, Batchelor J (1992) Physics of plastics. Carl Hanser, Munich, pp 283–284
47. Nielsen LE, Landel RF (1994) Mechanical properties of polymers and composites. Marcel Dekker, New York, p 253
48. Nielsen LE, Landel RF (1994) Mechanical properties of polymers and composites. Marcel Dekker, New York, pp 342–352
49. Patric SG (2005) Practical guide to polyvinyl chloride. Rapra Technology Limited, Shropshire
50. Wilkes CE, Summers JW, Daniels CA, Berard MT (1995) PVC handbook. Hanser Gardener, Cincinnati
51. Brydson JA (1999) Plastics materials, 7th edn. Butterworth-Heinemann, Oxford
52. Aharoni SM (1997) n-Nylons, their synthesis, structure, and properties. Wiley, New York
53. http://www.ides.com/generics/Nylon/Nylon_typical_properties.htm
54. Kohan MI (1995) Nylon plastics handbook. Hanser Publishers, New York
55. Averous L, Boquillon N (2004) Biocomposites based on plasticized starch: thermal and mechanical behaviours, *Carbohydr Polym* 56:111
56. Garcia MA, Martino MN, Zaritzki NE (2000) Lipid Addition to Improve Barrier Properties of Edible Starch-based Films and Coatings, *J Food Sci* 65(6):941–947
57. Shaikh HM, Pandare KV, Nair G, Varma A (2009) Utilization of sugarcane bagasse cellulose for producing cellulose acetates: Novel use of residual hemicellulose as plasticizer, *Carbohydr Polym* 76 (1–2):23–29
58. Garcia-Gonzalez CA, Alnaief M, Smirnova I (2011) Polysaccharide-based aerogels—Promising biodegradable carriers for drug delivery systems, *Carbohydr Polym* 86:1425–1438
59. Karnnet S, Potiyaraj P, Pimpan V (2005) Preparation and properties of biodegradable stearic acid-modified gelatin films, *Polym Degrad Stab* 90(1):106–110
60. Bourtoom T (2009) Edible protein films: properties enhancement, *Int. Food Res. J.* 16(1):1–9
61. Sobral PJA, Santos JS, García FT (2005) Effect of protein and plasticizer concentrations in film forming solutions on physical properties of edible films based on muscle proteins of a Thai Tilapia, *J Food Eng* 70(1):93–100
62. Pérez-Gago MB, Krochta JM (2001) Lipid Particle Size Effect on Water Vapor Permeability and Mechanical Properties of Whey Protein/Beeswax Emulsion Films, *J Agric Food Chem* 49 (2):996–1002
63. Sothornvit R, Krochta JM (2001) Plasticizer effect on mechanical properties of -lactoglobulin films, *J Food Eng* 50(3):149–155
64. Bertan LC, Tanada-Palmu PS, Siani AC, Grosso CRF (2005) Effect of fatty acids and ‘Brazilian elemi’ on composite films based on gelatin, *Food Hydrocoll* 19(1):73–82
65. Stienbuchel A, Marchessault RH (eds) (2005) Biopolymers for medical and pharmaceutical applications, vol 1. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim
66. Averous L, Pollet E (2011) Biorenewable Nanocomposites *MRS Bull* 36(9):703–710
67. Holden G (1996) Applications of thermoplastic elastomers (Chapter 16). In: Holden G, Legge NR, Quirk R, Shroeder HE (eds) Thermoplastic elastomers, 2nd edn. Carl Hanser, Munich
68. Walker BM, Rader CP (eds) (1988) Handbook of thermoplastic elastomers, 2nd edn. Van Nostrand Reinhold, New York

Additional References for Part II: American Plastics Council

Billmeyer FW Jr (1984) Textbook of polymer science. Wiley, New York

Brydson JA (1982) Plastics materials. Butterworth Scientific, London

Craver CD, Carraher CE Jr (2000) Applied polymer science, 21st century. Elsevier, Amsterdam

Deanin RD (1972) Polymer structure, properties, and applications. Cahners, Boston

Goodman SH (1998) Handbook of thermoset plastics. Noyes, Westwood

International Plastics Selector (1995) Plastics digest: thermoplastics and thermosets. D. A. T. A, Englewood

Harper CA (2000) Modern plastics handbook. McGraw-Hill, New York

Lubin G (1982) Handbook of composites. Van Nostrand Reinhold, New York

Modern Plastics Magazine (1994) Modern plastics encyclopedia handbook. McGraw-Hill, New York

Plastics News

Rubin II (1990) Handbook of plastic materials and technology. Wiley, New York

Sears JK, Darby JR (1982) The technology of plasticizers. Wiley, New York

Additional References for Part III

Harper CA (2000) Modern plastics handbook. McGraw-Hill, New York

Klempner D, Frisch KC (1991) Handbook of polymeric foams and foam technology. Hanser, Munich

Plastics News

Rubin II (1990) Handbook of plastic materials and technology. Wiley, New York

Rubber

16

D.F. Graves

Introduction

The word "rubber" immediately brings to mind materials that are highly flexible and will snap back to their original shape after being stretched. In this chapter a variety of materials are discussed that possess this odd characteristics. There will also be a discussion on the mechanism of this "elastic retractive force." Originally, rubber meant the gum collected from a tree growing in Brazil. The term "rubber" was coined for this material by the English chemist Joseph Priestley, who noted that it was effective for removing pencil marks from paper. Today, in addition to Priestley's natural product, many synthetic materials are made that possess these characteristics and many other properties. The common features of these materials are that they are made up of long-chain molecules that are amorphous (not crystalline), and the chains are above their glass transition temperature at room temperature.

Rubber products appear everywhere in modern society from tires to biomedical products. The development of synthetic rubber began out of the need for countries to establish independence from natural products that grew only in tropical climates. In times of conflict the natural product might not be available, and its loss would seriously threaten national security. Synthetic rubber, then, became a strategic concern during World Wars I and II [1]. Beyond the security issue, the need for materials with better performance also provided a strong impetus for the development of new rubbery materials. In particular, improvements in oil resistance, high-temperature stability, and oxidation and ozone resistance were needed [2]. Research today is driven to develop materials with even better performance in these areas. In the 1980s and 1990s tires with lower rolling resistance were demanded by car manufacturers to improve fuel economy. This was accomplished, in part, by developing functional tread polymers which chemically bond to fillers resulting in a dramatic reduction in the hysteresis (energy loss) of the tire tread. These new functional polymers will be discussed in this chapter.

The worldwide demand for rubber was estimated to be 23.0 million metric tons for 2011 [3], excluding latex materials. Of this demand, 36% is synthetic rubber of various kinds. Rubber has been classified by use into general purpose and specialty.

The major general purpose rubbers are natural rubber, styrene–butadiene rubber (SBR), butadiene rubber, isoprene rubber, and ethylene–propylene rubber. These rubbers are used in tires, mechanical goods, and similar applications. Specialty elastomers provide unique properties such as oil resistance or extreme heat stability. Although this differentiation is rather arbitrary, it tends also to classify the polymers according to volumes used. SBR, butadiene rubber, and ethylene–propylene rubber account for 78% of all synthetic rubber consumed.

The 2009 synthetic rubber capacity by type is presented in Table 16.1, and Table 16.2 lists the worldwide rubber consumption forecast by type for 2011.

Rubber Concepts

Several key principles (outlined below) greatly help in understanding the performance of elastomeric materials. This outline should present these concepts well enough for use in the following discussions of specific polymer types. However, the reader should consult the references for a more complete understanding of these principles. The concepts are classified as those that relate to polymer structure, those that relate directly to physical properties, and those that relate to the use of the material, as listed below.

1. Polymer structure:
   (a) Macrostructure
      - Molecular weight
      - Molecular weight distribution
      - Branching

D.F. Graves (✉)
Firestone Polymers, Division of Bridgestone/Firestone, Inc,
Firestone Polymers LCC, Akron, OH, USA

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_16,

Table 16.1 Synthetic rubber capacities by type, 2009

Type	Description	Metric Tons
SBR-solid	Styrene/butadiene rubber	5,257,000
BR	Butadiene rubber	3,078,000
IR	Isoprene rubber	707,000
CR	Chloroprene (or neoprene)	385,000
EPDM	Ethylene/propylene terpolymer	1,039,000
NBR	Nitrile or acrylonitrile/butadiene rubber	713,000
SBR	Styrenic Block Polymers	1,834,000
Total		13,013,000

Source: Worldwide Rubber Statistics 2009, International Institute of Synthetic Rubber Producers, Inc., by permission

Table 16.2 Worldwide rubber consumption forecast, 2011 (metric tons)

Total new rubber	22,970,000
Natural	10,100,000
Total synthetic	12,910,000
SBR	3,800,000
Nitrile	473,000
Polybutadiene	2,254,000
EPDM	1,011,000
Other synthetics	5,370,000
% Natural	43.8
% Synthetic	56.2

Source: Worldwide Rubber Statistics 2009, International Institute of Synthetic Rubber Producers, Inc., by permission (http:\crwww.therubbereconomist.com)

    (b) Microstructure
    (c) Network structure
2. Rubber properties:
    (a) Elasticity—the retractive force
    (b) Glass transition temperature
    (c) Crystallinity
3. Rubber use:
    (a) Compounding
    (b) Processing

Polymer Structure

Macrostructure

Molecular weight. The single most important property of any polymer is the size or length of the molecule. The polymerization process consists of combining many of the simple monomer molecules into the polymer chain. Most of the monomers used to produce rubbers are either gases or low-viscosity liquids under normal conditions; upon polymerization they form liquids whose viscosity increases to extremely high values as the chain length is increased. At very low chain length, this increase is linear with molecular weight until the chains are long enough to become entangled. Above the entanglement molecular weight, the viscosity increases to the 3.4–3.5th power of molecular weight increase [4]. In addition to viscosity, a great many other physical properties of any polymer depend upon the molecular weight [5].

Molecular weight distribution. A given polymer sample is composed of many polymer chains, which in most cases are not of the same length. This variability can be a result of the synthesis process or of possible random scission and cross-linking that can occur upon processing. For economic reasons, it is not possible to separate the various polymer chains by length prior to use; so it is important to characterize this distribution in order to describe the polymer and understand its performance. As with any distribution, no single number is a totally satisfactory descriptor.

The commonly used molecular weight parameters are the number, weight, and z average molecular weight, which are defined, respectively, as: [6]

$$M_n = \frac{\Sigma N_i M_i}{\Sigma N_i}$$

$$M_w = \frac{\Sigma N_i M_i^2}{\Sigma N_i M_i}$$

$$M_z = \frac{\Sigma N_i M_i^3}{\Sigma N_i M_i^2}$$

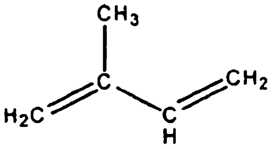
where N_i is the number of moles of species i, and M_i is the molecular weight of species i.

Although there are many different statistical ways to describe any population, the above parameters have been widely used because they are readily understood in physical terms, and they can be measured directly in the laboratory [7]. A fourth parameter, the dispersion index, frequently is used to characterize the breadth of the distribution. This parameter is simply the ratio of the weight to the number average molecular weight, with 1.0 being the lowest possible number (i.e., all chains of exactly the same length). Typical values for commercial polymers are in the 2–5 range, with those under two considered relatively narrow and those over 2.5 considered broad in distribution. The measurement of these molecular weight averages once was a time-consuming task, but with the development of gel permeation chromatography (GPC), also referred to as size exclusion chromatography, the measurement of these distributions has become commonplace [8]. Units are even available that automatically sample polymerization reactors, process the sample, and perform all necessary calculations to provide data for process control. These units can have multiple detectors, thereby providing compositional distribution as a function of molecular weight.

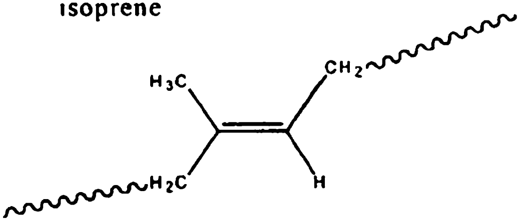
Branching. The concept of a polymer chain implies two ends per chain. However, because of the nature of the process used to form the polymer, the chain may contain one or more branch points, resulting in multiple ends per chain. These chain ends can have an adverse effect on polymer performance. Branching, molecular weight, and molecular weight distribution have been shown to affect process ability as well [9]. The optimum macrostructure often represents a compromise between processing and ultimate performance. Branching can also be measured using the GPC technique with special detectors.

Microstructure

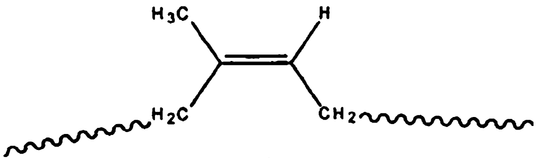
In the formation of elastomers from diolefin monomers such as butadiene or isoprene, there are a number of possible structures. Since the control of these structures is critical in obtaining optimum properties, this area has received great attention from the synthesis chemist. The possible polyisoprene structures are:



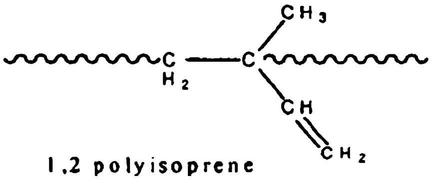
isoprene



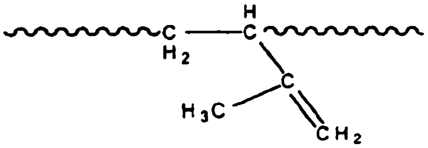
trans-1,4 polyisoprene



cis-1,4 polyisoprene



1,2 polyisoprene



3,4 polyisoprene

For butadiene (no methyl group) the 3,4 form does not exist. The 1,2 addition is referred to as *vinyl addition*.

The polymer in natural rubber (from the *Hevea brasiliensis* tree) is pure *cis* polyisoprene; gutta percha and balata are composed of the *trans* isomer.

Many of the commercial synthetic elastomers are synthesized from more than one monomer, such as styrene–butadiene and ethylene–propylene rubbers. The properties of the resultant polymer depend on the ratio of the two monomers in the polymer and upon the distribution of the monomers within the chain.

If the monomers are uniformly distributed within the polymer chain, the ratio of monomers will define the flexibility of the polymer chain [10]. Because many properties depend on this chain mobility, polymer composition is carefully controlled. In addition to chain mobility, polymer composition also defines the solubility parameter of the polymer, which is a critical property relative to the type of solvents in which the polymer is soluble, the ability of the polymer to accept and hold oil, and the relative compatibility of the polymer with other polymers [11]. Basically, these properties all involve polymer–solvent interaction, with the difference being the increasing molecular weight of the solvent.

In addition to the relative ratio of the monomers, the arrangement of the units in the chain is important. This arrangement is referred to as the copolymer *sequence distribution*. In the previous discussion, the assumption was made that the comonomer units were well mixed in the polymer chain. If this is not the case, parts of the chain can reflect properties of the corresponding homopolymer. It is thus possible to produce polymers that have significantly different properties in different parts of the polymer chain. A most dramatic example of this can be found in styrene–butadiene–styrene (SBS) or styrene–isoprene–styrene thermoplastic elastomers (TPE). The properties of these unique materials will be discussed in the section "Thermoplastic Elastomers."

Network Structure

A critical requirement for obtaining engineering properties from a rubbery material is its existence in a network structure. Charles Goodyear's discovery of vulcanization changed natural rubber from a material that became sticky when hot and brittle when cold into a material that could be used over a wide range of conditions. Basically, he had found a way to chemically connect the individual polymer chains into a three-dimensional network. Chains that previously could flow past one another under stress now had only limited extensibility, which allowed for the support of considerable stress and retraction upon release of the stress. The terms "vulcanization," "rubber cure," and "cross-linking" all refer to the same general phenomenon.

For most rubber systems the network is formed after the polymer is compounded and molded into the desired final shape. Once cross-linked, the material no longer can be processed. If cross-linking occurs prior to compounding or molding, the material is referred to as *gelled*, and it cannot be used. Most rubber is used in a compounded and cured form. There is an optimum *cross-link density* for many failure properties such as tensile strength and tear which will be discussed in the next section.

The one general class of polymers that fall outside this concept is the thermoplastic elastomers, which will also be discussed later.

Rubber Properties

Elasticity: The Retractive Force

The fact that cross-linked rubber materials can be extended to several times their original length and return to that original length when released is certainly their most striking feature. This is in contrast to crystalline solids and glasses, which cannot normally be extended to more than a fraction of their original length and also to ductile metals which can be extended to large deformations but do not return to the original length after the stress is removed.

There have been both statistical and thermodynamic approaches to solving the problem of rubber elasticity leading to a phenomenological treatment; however, these methods are beyond the scope of this chapter. The important, and most interesting, result of these theories is that rubber elasticity arises from changes in entropy of the network. Rubber molecules are capable of geometric isomerization, examples of which are *cis* and *trans*, just like any other unsaturated organic molecule. Upon stretching the network of chains, no "configurational" changes take place. That is, *cis* is not changed to *trans* and thus there is no configurational contribution to entropy. Elasticity *does not* arise from stretching or deforming covalent bonds either. What does take place are rotations about single bonds in the chain backbone during stretching. It is these "conformational" changes which give rise to the entropy decrease upon stretching. The chains do not like being in a stretched state because there are many conformational states not available to them until the stress is released. The chain ends are held apart at a statistically unfavorable distance, which gives rise to the entropy decrease. The following expression, for extensions $>10\%$, shows that the elastic force, f, is directly proportional to the absolute temperature, or the elastic response of the rubber is entirely governed by the decrease in entropy (S) which it undergoes upon extension [12]. The term ∂L is the change in length of the sample:

$$f = -T\left[\frac{\partial S}{\partial L}\right]_{T,V}$$

This expression was derived for constant temperature and volume experiments.

The shear modulus of the rubber network is related to the molecular weight between cross-link points or M_c. The lower the molecular weight of chains between cross-links (network chains), the higher the cross-link density and the higher the modulus. This is shown in the following expression:

$$G_0 = \frac{\rho RT}{M_c}$$

where G_0 is the elastic shear modulus, ρ is the density, R is the gas constant, and T is absolute temperature. There is an optimum cross-link density for ultimate strength properties, above which the highly cross-linked network no longer can dissipate strain energy in the form of heat (hysteresis) so all the energy goes into breaking network chains and the material becomes brittle. Below this optimum cross-link level the material simply has too much viscous flow and pulls apart easily. Therefore, the rubber chemist must optimize the state of cure if high strength is desired.

Glass Transition Temperature

In order for a polymer to behave as a rubbery material, it is necessary for the chain to have great mobility. As the temperature is lowered, the ability of the chain segments to move decreases until a temperature is reached where any large-scale motion is prevented. This temperature is referred to as the glass transition temperature (T_g). Below this temperature the rubber becomes a glassy material—hard and brittle. Above this temperature amorphous plastics, such as polystyrene, can exhibit a rubbery character if the molecular weight is sufficiently high. All rubbery materials, then, must have glass transition temperatures below room temperature. For good low-temperature properties, it obviously follows that a low glass transition temperature polymer is required. The control of the glass transition temperature of the polymer is critical for many properties other than low-temperature use. For example, the wet traction and wear of a passenger tire have been shown to greatly depend directly on the polymer glass transition temperature [13].

Control of this property is possible by controlling the structure of the polymer chain. Monomers with bulky side groups restrict chain mobility and thus raise the glass transition temperature. The composition of copolymers and the ratio of polymer blends often are determined by the desired glass transition temperature of the final product.

The glass transition temperature is usually measured using thermal methods such as differential scanning calorimetry (DSC) which looks at the change in heat capacity when a material goes through its glass transition.

Another very useful method is to use dynamic mechanical properties where the polymer is subjected to a temperature sweep in a dynamic mechanical spectrometer from very low temperature ($-120°\mathrm{C}$) to well above room temperature ($+100°\mathrm{C}$). If the modulus is plotted vs. temperature, there will be a sharp decrease as the polymer warms to above its T_g. There will also be a peak in the energy loss property known as tan delta. The tan delta vs. temperature plot is very useful for predicting tire properties such as rolling resistance and wet traction.

Crystallinity

Polymer chains can exist in spatial arrangements that are orderly enough to allow the chains to form crystalline structures. The existence of strong interchain interactions via polar forces, hydrogen bonding, or ionic groups can facilitate crystallization. The existence of crystallization is very important for many plastics and fibers, but crystallinity cannot be appreciably present in rubbery materials, as the corresponding restriction in chain mobility could preclude the very chain mobility needed for rubbery properties. Rubbery materials must have both their melting temperature (if they have a crystalline point) and their glass transition temperature below the use temperature.

Crystallinity can be reduced by disruption of the order in the chain by copolymerization [14]. For example, both polyethylene and polypropylene are crystalline plastics, whereas ethylene–propylene rubber produced at about a 50:50 ratio is an amorphous elastomer. Compositional excursions much outside this range lead to crystalline materials [15]. For some materials, such as natural rubber, that are close to crystallizing, stretching the chains can align them sufficiently for crystallization to occur. Such polymers can exhibit excellent gum properties and improved strength in the uncured state that greatly facilitate processing.

Attempts have been made with some success to produce other polymers that exhibit this property of natural rubber. Although the melting temperature can be matched by appropriately disrupting the crystallizable structure through controlled introduction of another monomer, an exact match is not possible because the extent of crystallinity and the kinetics of crystallization will differ.

Crystallinity can be measured using the same thermal and dynamic mechanical methods described for measuring T_g; however, the melting transition is much sharper than the T_g because it is a first-order transition compared with second-order for the T_g.

Rubber Use

Compounding

The rubber industry began when Charles Goodyear developed the first useful rubber compound: natural rubber plus sulfur. The concept of mixing materials into rubber to improve performance is still of primary importance today. Without compounding, few rubbers would be of any commercial value. Any given rubber application will have a long list of necessary criteria in addition to cost, encompassing appearance, processing, mechanical, electrical, chemical, and thermal properties. Developing such compounds requires a broad knowledge of material science and chemistry combined with experience. The use of designed experiments can greatly facilitate selecting the optimum compound formulation.

The major components in a compound are curatives, reinforcing agents, fillers, plasticizers, and antidegradants.

Curatives. The function of curatives is to cross-link the polymer chains into a network; the most common ones are the sulfur type for unsaturated rubber and peroxides for saturated polymers. Chemicals called accelerators may be added to control the cure rate in the sulfur system; these materials generally are complex organic chemicals containing sulfur and nitrogen atoms. Stearic acid and zinc oxide usually are added to activate these accelerators. Metal oxides are used to cure halogenated polymers such as polychloroprene or chlorosulfonated polyethylene.

Reinforcing agents. Carbon black and silica are the most common reinforcing agents. These materials improve properties such as tensile strength and tear strength; also, they increase hardness, stiffness, and density and reduce cost. Almost all rubbers require reinforcement to obtain acceptable use properties. The size of the particles, how they may be interconnected (structure), and the chemical activity of the surface are all critical properties for reinforcing agents. In tire applications, new polymers are currently being developed which contain functional groups that directly interact with carbon black and silica, improving many properties.

Fillers. Fillers are added to reduce cost, increase hardness, and color the compound. Generally they do not provide the dramatic improvement in properties seen with reinforcing agents, but they may have some reinforcing capability. Typical fillers are clays, calcium carbonate, and titanium dioxide.

Plasticizers. These materials are added to reduce the hardness of the compound and can reduce the viscosity of the uncured compound to facilitate processes such as mixing and extruding. The most common materials are petroleum-based oils, esters, and fatty acids. Critical properties of these materials are their compatibility with the rubber and their viscosity. Failure to obtain sufficient compatibility will cause the plasticizer to diffuse out of the compound. The oils are classified as aromatic, naphthenic, or paraffinic according to their components. Aromatic oils will be more compatible with SBR than paraffinic oils, whereas the inverse will be true for butyl rubber. The aromatic oils are dark colored and thus cannot be used where color is critical, as in the white sidewall of a tire. The naphthenic and paraffinic oils can be colorless and are referred to as nonstaining.

Antidegradents. This group of chemicals is added to prevent undesirable chemical reactions with the polymer network. The most important are the antioxidants, which trap free radicals and prevent chain scission and cross-linking. Antiozonants are added to prevent ozone attack on the rubber, which can lead to the formation and growth of cracks. Antiozonants function by diffusion of the material to the surface of the rubber, thereby providing a protective film. Certain anti-oxidants have this characteristic, and waxes also are used for this purpose.

Processing

A wide range of processes are used to convert a bale of rubber into a rubber product such as a tire. The first process generally will be compounding. Typical compounding ingredients were discussed previously. In many compounds more than one rubber may be needed to obtain the performance required. Uncured rubber can be considered as a very high-viscosity liquid; it really is a *viscoelastic* material possessing both liquid and elastic properties. Mixing materials into rubber requires high shear, and the simplest method is a double roll mill in which the rubber is shear-mixed along with the other compounding ingredients in the bite of the mill. Large-scale mixing is most commonly done with a high-shear internal mixer called a Banbury. This mixing is a batch process, although continuous internal mixers also are used.

The compounded rubber stock will be further processed for use. The process could be injection or transfer molding into a hot mold where it is cured. Tire curing bladders are made in this fashion. Extrusion of the rubber stock is used to make hose or tire treads and sidewalls. Another common process is calendaring, in which a fabric is passed through rolls where rubber is squeezed into the fabric to make fabric-reinforced rubber sheets for roofing membranes or body plies for tires. The actual construction of the final product can be quite complex. For example, a tire contains many different rubber components some of which are cord or fabric reinforced. All of the components must be assembled with high precision so that the final cured product can operate smoothly at high speeds and last over 50,000 miles.

Natural Rubber

More than 500 years ago, the people of Central and South America were using a product that they collected from certain trees to make balls and to coat fabric to make it waterproof. This material they called cauchuc, which means “weeping wood.” Today we know the tree as the *H. brasiliensis* and the material as natural rubber. Although a number of plants produce rubber, the only significant commercial source is the *Hevea* tree. Natural rubber initially had only limited applications because it flowed when hot and had poor strength. In 1839 Charles Goodyear found that when combined with sulfur and heated, the material changed into cured rubber with properties much as we know them today. The development of the pneumatic tire in 1845 combined with the dramatic growth of the automotive industry led to a rapid increase in the demand for natural rubber.

Although the tree was indigenous to Brazil, seeds were taken to England where they were germinated, and the plants were sent to the Far East. Rubber plantations were in existence by the late 1800s, and in the 1920s plantations were begun in West Africa. Because of a leaf blight disease, essentially all natural rubber now comes from plantations in Africa and the Far East rather than Central and South America [16].

The production process starts with the trees. Over the years considerable biological research has been done to produce trees that grow faster, produce more latex, and are resistant to wind and disease damage. Once such an improved tree has been identified, buds are grafted from the tree onto root stock. All such trees are referred to as clones and will have the same characteristics. It typically takes 6–7 years of growth before a tree is ready for rubber recovery. Peak rubber production is reached at 12–15 years of age. Another major development in improving tree performance has been the use of tree stimulants, which has resulted in an overall yield increase of 30% without adverse effects on the trees.

The production process of natural rubber in the tree is not yet fully understood. However, it involves a long series of complex biochemical reactions that do not involve isoprene as a monomer, even though the resulting polymer is 100% *cis* 1,4 polyisoprene. Because the tree makes the product, the rubber production process is really one of recovery.

The recovery process starts with tapping of the tree, which involves manually removing thin sections of bark at an angle so that as the latex is exuded from the damaged living latex cells, it will flow down to be collected in a cup. The depth of the cut is critical, as a tool shallow cut will not allow optimum latex yield, and too deep a cut will damage the tree. The damaged latex vessels will seal off after several hours. A preservative is added to the collection cup to prevent coagulation of the latex. The tapper then collects the latex and takes it to a collection station, from which it is shipped to the rubber factory. Additional preservative is added at the collection station to assure stable latex. The latex contains around 30–45% rubber. Approximately 3% of the solid material is nonrubber materials, consisting primarily of proteins, resins, and sugars.

The latex is processed in one of several ways, depending on the desired final product. If it is going to be used in the final product application in the latex form (such as for dipped goods and adhesives), it will be concentrated to 60% or higher solids. The most common process is centrifugation. The latex separates into the high-solids product and a low-solids skim material that contains a much higher percentage of the nonrubbery components; rubber produced from skim is generally of a lower quality because of these impurities. A variety of chemicals can be added to the latex to provide the necessary preservation and mechanical stability. As the final use of the latex will involve destabilizing the rubber particles, care must be taken not to overstabilize the latex concentrate. Being a natural product, the latex tends to change upon aging, a factor that also must be compensated for in the process.

Dry rubber is produced from the latex first by dilution, then by coagulation with organic acids, and finally by formation into sheets or crepe. Rubber smoked sheets are made by working the coagulated sheets between rolls to remove as much of the nonrubbery components as possible, followed by drying for up to a week in a smokehouse. The smoke serves as a fungicide that prevents biological attack of the rubber. To provide a more well-defined uniform product, Technically Specified Rubber (TSR) processes have been developed. These processes involve converting the coagulated rubber into rubber crumb, which is further washed, dried, and baled. Constant-viscosity grades of natural rubber have been developed by chemically reacting the aldehyde groups, which otherwise would lead to cross-linking upon storage. In the pale crepe process the latex is carefully selected for colored bodies (from carotene) and treated with sodium bisulfite to stop enzyme activity. The rubber is extensively washed through rollers to remove serum, as this can lead to yellowing. The wet crepe is dried under carefully controlled temperatures and in the absence of light to assure optimum properties. The previously described processes involve considerable investment and are practical only for plantation operations. Small shareholders may allow their latex to coagulate naturally and sell it to processors. Such a product will vary greatly, depending on the specific history of each rubber slab [17].

Uses

The largest use of natural rubber is in the manufacture of tires. Over 70% of its consumption is in this area. The next largest use is as latex in dipped goods, adhesives, rubber thread, and foam. These uses account for approximately another 10%. The remainder is used in a variety of applications such as conveyor belts, hoses, gaskets, footwear, and antivibration devices such as engine mounts.

Because of the high stereoregularity of natural rubber, the units in the polymer chain can form very orderly arrangements, which result in crystallization upon storage at low temperatures or upon stretching. Stored crystallized rubber may be converted to its original amorphous state by heating. Several crystalline forms are reported, with melting points varying from 14 to 36°C [18]. Although crystallization upon storage can be a problem to users, the ability to crystallize reversibly upon stretching accounts for many of the unique properties of natural rubber. Specifically, the ability of natural rubber to be used as a gum polymer (unfilled vulcanizate) depends on this property. The crystallites that form act both as filler and as temporary cross-links, providing high tensile properties [19]. In tire fabrication this property is reflected in natural rubber stocks possessing high green (uncured) strength and excellent building tack. In the final product, strain-induced crystallization provides tear and cracking (cut growth) resistance.

Although many other polymers can crystallize, only natural rubber has been found to have the necessary combination of rate of crystallization, degree of crystallization, and melting point to provide all the properties discussed above. Natural rubber has good flexibility and resilience. In truck and bus tires, it is used extensively with blends of polybutadiene to give the low heat buildup needed as well as wear and cut growth resistance. In passenger tires, natural rubber is used in the sidewalls and carcass areas. These areas require the building tack, ply adhesion, and hot strength properties that it imparts. Synthetic rubber is used almost exclusively in passenger tire treads.

The largest use of natural rubber latex is in the dipped goods area. Products include balloons, surgical and examination gloves, and prophylactics. The rapid spread of AIDS has led to a dramatic increase in the surgical and examination glove market. Latex also is used to make adhesives, rubber thread, and foams. However, natural rubber is being replaced in these two latex areas by urethanes and synthetic rubber latex. Some of the replacement has been driven by skin allergy problems which are caused by the natural proteins in the latex.

Although natural rubber latex is a mature product, research continues on improving its uniformity, stability, and performance [20].

Polyisoprene

Faraday discovered in 1826 that natural rubber was composed of a hydrocarbon with a ratio of five carbons to eight hydrogen atoms; and in 1860, G. Williams isolated isoprene by collecting the distillate from the heating of natural rubber. By 1887, scientists in France, England, and Germany had converted isoprene back into a rubbery material. Because this offered a potential for manufactured "natural" rubber, research was undertaken to find ways to obtain isoprene from sources other than rubber itself [21].

Monomer Production

The primary source of isoprene today is as a by-product in the production of ethylene via naphtha cracking. A solvent extraction process is employed. Much less isoprene is produced in the crackers than butadiene, so the availability of isoprene is much more limited. Isoprene also may be produced by the catalytic dehydrogenation of amylenes, which are available in C-5 refinery streams. It also can be produced from propylene by a dimerization process, followed by isomerization and steam cracking. A third route involves the use of acetone and acetylene, produced from coal via calcium carbide. The resulting 3-methyl-butyne-3-ol is hydrogenated to methyl butanol and subsequently dehydrogenated to give isoprene. The plants that were built on these last two processes have been shut down, evidently because of the relatively low cost of the extraction route.

Polymer Production Process

The free-radical catalysts were found to produce a product that did not have the tack, green strength, or gum tensile of natural rubber. Whereas natural rubber is an essentially pure *cis*-1,4 structure, the emulsion product was of mixed microstructure. This precluded the ability of the latter to undergo strain-induced crystallization, which is required to obtain many of the desired natural rubber properties.

In 1955 investigators from the Firestone Tire and Rubber Company and the B. F. Goodrich Company announced the synthesis of polyisoprene with over 90% *cis*-1,4 structure. The work at Firestone was based on lithium metal catalysts, whereas the work at Goodrich was the result of using Ziegler–Natta type coordination catalysts [22,23].

Use

Although considerable interest was generated by these discoveries, their commercial success has been rather limited. The lithium-based polymers were found to produce up to 94% *cis*, which still was not high enough to provide the properties of natural rubber. Polymers made with the coordination catalysts have *cis* contents of up to 98%, providing products that can more closely serve as replacements for natural rubber than the lithium-based polymers. In comparison with natural rubber, they offer the advantage of a more highly pure rubber (no nonrubber material) and excellent uniformity. For economic reasons, polyisoprene has seen only limited success. Several of the plants built to produce polyisoprene have been either shut down or converted for use to produce other polymers. In terms of synthetic rubber production in 2003, only 258,000 metric tons were produced (excluding centrally planned economy countries, CPEC), which represented only 3% of total synthetic rubber production. In comparison, in 2003 natural rubber usage was 7,554,000 metric tons [24]. Evidently because of its strategic importance, the Former Soviet Union (FSU) continues to rely heavily on polyisoprene.

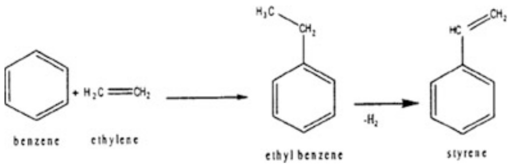
Styrene–Butadiene Rubber

The largest-volume synthetic rubber consumed is SBR. In 2009, SBR solid rubber accounted for 34% of all synthetic rubber. If SBR latex and carboxylated SBR latex are included, its share increases to 50%. The major application of solid SBR is in the automotive and tire industry, accounting for approximately 70% of the use. Therefore, SBR has been tightly tied to the tire business [25].

Initially, SBR was developed as a general purpose alternate material to natural rubber. In the United States the thrust came early in World War II when the US supply of natural rubber was cut off. The basic technology was developed in Germany in the late 1920s, and by 1939 Germany had 175,000 metric tons of capacity in place. The first US production was 230 metric tons in 1941, but by 1945 there were more than 850,000 metric tons of capacity. Basically, in a period of 5 years the emulsion SBR business as we know it today was put in place. By 1973 US capacity had increased to almost 1,400,000 metric tons, but in 1989 it had contracted to 881,000 tons close to the 1945 capacity [26].

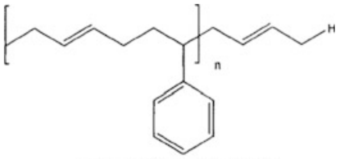
Monomer Production

The production of butadiene monomer is discussed below in section "Polybutadiene." The largest volume of styrene is produced by the alkylation of benzene with ethylene to give ethyl benzene, which is then dehydrogenated to give styrene [27].



Polymer Production Process

SBR is produced by two different processes: emulsion and solution. The emulsion process involves a free-radical mechanism, whereas solution SBR is based on alkyllithium catalysis.



copolymer of butadiene and styrene

Emulsion Process

The formula developed to provide SBR during World War II was standardized, with all rubber plants owned by the US government. The standard recipe is listed below[28].

Component	Parts by weight
Butadiene	75
Styrene	25
n-Dodecyl mercaptan	0.5
Potassium peroxydisulfate	0.3
Soap flakes	5.0
Water	180

Initiation occurs through reaction of the persulfate with the mercaptan, as shown below:

$$K_2S_2O_8 + 2RSH \rightarrow 2RS\bullet + KHSO_4$$
$$RS\bullet + M \rightarrow RSM\bullet$$

Chain propagation occurs by the growing chain free radical attacking either the butadiene or styrene monomer. The active radical chain can react with mercaptan to form a new mercaptyl radical and a terminated chain. The mercaptyl radical then can initiate an additional chain. The molecular weight of the chain P can be controlled by the concentration of mercaptan via this chain transfer mechanism.

$$P\bullet + RSH \rightarrow PH + RS\bullet$$
$$RS\bullet + M \rightarrow RSM\bullet$$

Termination also can occur by the reaction of two free radicals, through either combination or disproportionation reactions.

$$P\bullet + P\bullet \rightarrow P - P \text{ combination}$$
$$\text{(bimolecular coupling)}$$

$$P\bullet + P\bullet \rightarrow P - CH = CH_2$$
$$+ PH \text{ disproportionation}$$
$$\text{(hydrogen-free radical transfer)}$$

The mercaptyl radical also can react with growing chains, to lead to termination [29].

Polymerization is initially carried out at 50°C until conversion of 70–75% is reached, at which time the polymerization is terminated by the addition of a free-radical scavenger such as hydroquinone. Polymerization beyond this point results in excessive free-radical attack on the polymer chains. Products made under such conditions have poor properties due to excessive branching and gelation. Unreacted butadiene and styrene are removed by flashing and steam stripping. Antioxidant is added to the latex, followed by coagulation with the addition of polyelectrolytes and salt-acid. The coagulated crumb then is washed, dried, baled, wrapped, and packaged for shipment. Because of the soap and other chemicals in the formulation, most emulsion polymers will contain about 7% of nonrubber residues. The emulsion process flow sheet is shown in Fig. 16.1.

It was soon discovered that polymers made at lower temperatures had significantly better properties, especially in tire treads. This was mainly because they contained fewer low-molecular-weight species. A lower temperature process, using "redox" chemistry, was eventually developed. It used peroxides or hydroperoxides with a reducing agent such as a water-soluble transition metal salt which were active even at 0°C.

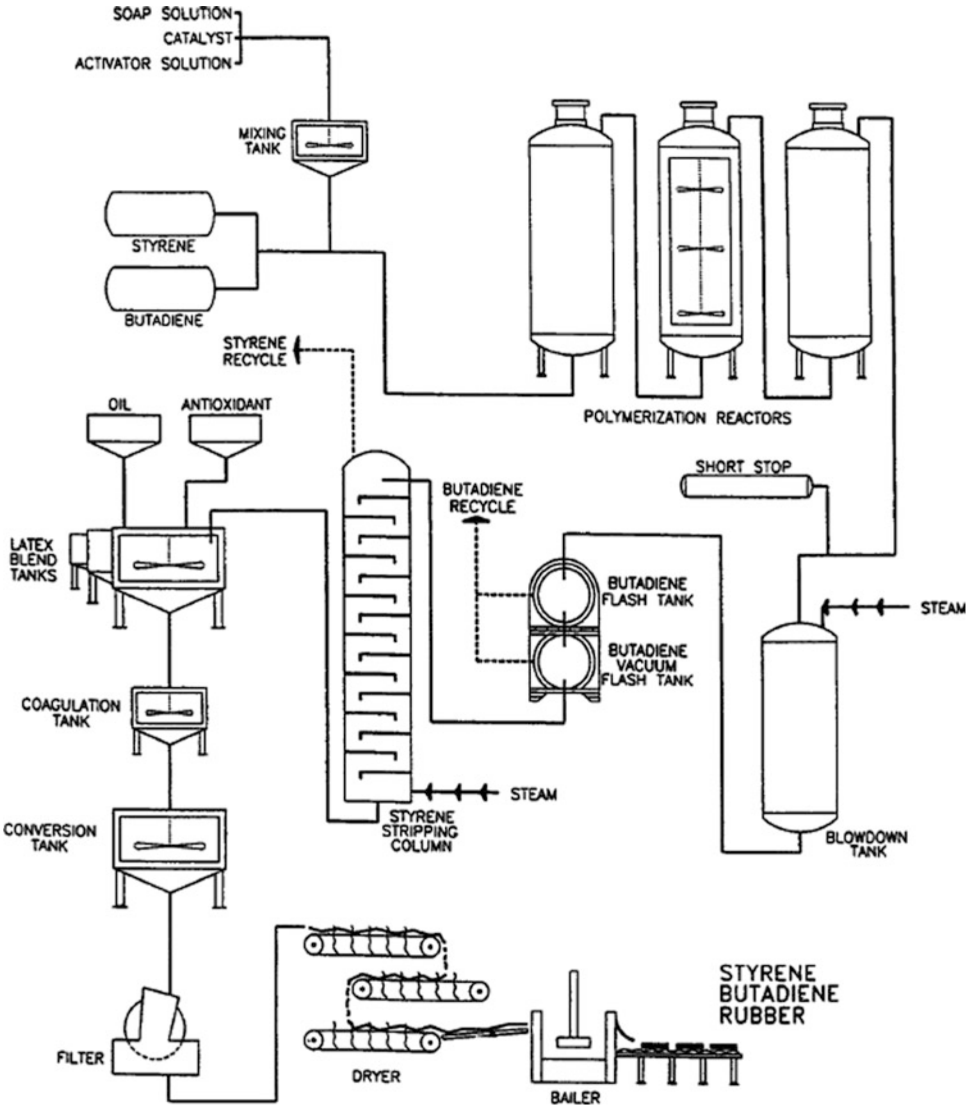


Fig. 16.1 The production of styrene-butadiene rubber (modified by Graves DF from a drawing in the Vanderbilt Rubber Handbook, 1990 Edition, Copyright the R. T. Vanderbilt Company, Inc., by permission)

Mercaptans are also used as chain transfer agents to provide a mechanism for molecular weight control. Commercially these types of polymerization are carried out at 5°C and are referred to as "cold" polymerizations to differentiate them from the previously discussed "hot" systems. A typical formula is listed below [30].

Component	Parts by weight
Butadiene	71
Styrene	29
tert-Dodecyl mercaptan	0.18
p-Methane hydroperoxide	0.08
Ferrous sulfate heptahydrate	0.03
Trisodium phosphate decahydrate	0.50
Tetrasodium ethylenediaminetetraacetate	0.035
Sodium formaldehyde sulfoxylate	0.08
Rosin acid soap	4.5
Water	200

The improved tire wear of cold polymerization SBR led to the very rapid replacement of hot SBR for most applications. This change was relatively easy to make, as all the equipment could be used with the only modification required being the addition of reactor cooling, which is achieved with either the reactor jacket, internal coils, or both.

It later was found that even more improvements could be realized by polymerizing to very high molecular weights and then adding petroleum-based oils to the latex prior to coagulation. The oil is absorbed by the rubber which, upon coagulation, produces oil-extended polymers. For tread applications, oils of higher aromatic content were preferred because of their excellent compatibility with the rubber. Typically 37.5 parts of oil are added, although grades containing up to 50 parts have been produced. Very high molecular weight polymers thus can be processed without requiring excessive energy to mix them. The oil also allows these tough polymers to be processed without excessive degradation. Carbon black master batches also are produced. In this process carbon black is added to the latex prior to coagulation, and the black, along with oil, is incorporated into the latex in the coagulation step. These products offer the user the advantage of not having to handle free black in their mixing operation, and can provide additional compounding volume for manufacturers with limited mixing capacity.

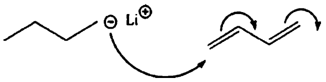
Not all emulsion SBR is converted to dry rubber for use. There is a variety of applications where the latex can be used directly in the final fabrication process. This technology logically grew out of the latex technology developed for natural rubber. For latex applications the particle size distribution can be critical because of its effect on viscosity and performance variables, as when used to provide impact strength in plastics such as ABS [31]. Careful control of the mechanical stability of the latex also is critical, as these systems must destabilize under relatively mild conditions such as those in a coating operation. A number of processes have been developed to control particle size via partial destabilization of the latex. Among the commercial methods are careful control of a freeze–thaw cycle, controlled shear agitators, high-pressure colloid mills, and the addition of chemicals such as hydrocarbons or glycols.

A special variation of SBR latex containing terpolymerized vinyl pyridine is used in the tire industry to provide adhesion of organic fiber tire cords to rubber stock. The vinyl pyridine SBR latex is combined with resins and coated on the fiber by a dipping process. The adhesive is set by a controlled temperature and tension process to control the shrinkage properties of the cord.

Solution Process

The discovery of the ability of lithium-based catalysts to polymerize isoprene to give a high *cis* 1,4 polyisoprene was rapidly followed by the development of alkyl lithium-based poly-butadiene. The first commercial plant was built by the Firestone Tire and Rubber Company in 1960. Within a few years the technology was expanded to butadiene–styrene copolymers, with commercial production under way toward the end of the 1960s.

The copolymerization with alkyl lithium to produce uniformly random copolymers is more complex for the solution process than for emulsion because of the tendency for the styrene to form blocks. Because of the extremely high rate of reaction of the styryl-lithium anion with butadiene, the polymerization very heavily favors the incorporation of butadiene units as long as reasonable concentrations of butadiene are present. This observation initially was somewhat confusing because the homopolymerization rate of styrene is seven times that for butadiene. However, the cross-propagation rate is orders of magnitude faster than either, and it therefore dominates the system. For a 30 mole percent styrene charge the initial polymer will be almost pure butadiene until most of the butadiene is polymerized. Typically two-thirds of the styrene charged will be found as a block of polystyrene at the tail end of the polymer chain:



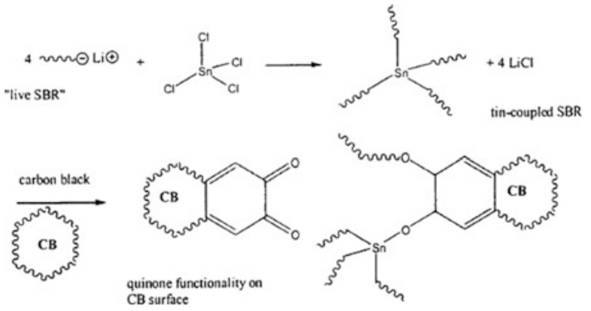
Several methods have been proposed to overcome this problem. In one, the styrene and part of the butadiene are charged initially with butadiene metered at a rate equivalent to its incorporation into the chain. A second approach involves adding both monomers at a relatively slow rate so that the equilibrium monomer concentration reaches a pseudosteady state that will produce polymer at the desired composition [32]. This process can be done in either a batch or a continuous mode [33].

In addition to these reaction engineering approaches to produce uniformly random copolymer, the chemistry may be changed by the addition of polar agents such as amines and ethers. This action results in bringing the reactivity rates much closer together. The change in chemistry also is reflected in the microstructure of the butadiene portion of the polymer. Whereas in the nonpolar system the vinyl content of the butadiene portion is around 10%, in polar systems vinyl contents of 30–40% typically are obtained when a 20% styrene polymer is randomized. Higher styrene contents require higher modifier levels, resulting in even higher vinyl contents. An added complication with polar modifiers is their ability to react with the growing chain, resulting in undesired termination.

In spite of these complications, all recent US expansions or announced plants for SBR have been for solution polymers. The ability to better design the polymer structure and produce special functional polymers (described below) accounts for most of this shift.

Functional Solution SBR

The driving force toward functional solution SBR is its improved hysteresis properties for passenger tire treads. The fact that anionic SBR has a much more narrow molecular weight distribution compared with emulsion gives it lower hysteresis. However, the big advantage is the relatively stable growing chain ends which can be chemically modified to improve interaction with carbon black and silica in tire compounds [34]. This modification can lead to a dramatic reduction in rolling resistance, which is critical for automotive manufacturers who must meet government-mandated fuel economy targets. The most active functional end-groups contain either organotin or certain amines. Termination with tin tetrachloride is the easiest and most popular method which generates a four-armed star polymer. The polymer–tin bonds break down during mixing of the compound and both lower the compound viscosity and create active sites for reaction with carbon black surfaces. The creation of this "carbon-bound rubber" effectively prevents the carbon black from agglomerating on a micro-scale to form hysteretic, three-dimensional networks. The breaking of these networks during the deformation of a tire tread is a major source of rolling resistance. Polymers with amine end-groups also show good activity with carbon black. Termination with silane esters is usually used to obtain interaction with silica fillers producing the same decrease in hysteresis. There have recently been efforts to make functional anionic initiators, some of which could be used to make low hysteresis rubber [35]. However, the main use for such polymers has been in adhesives and other nontire applications. One study comparing an emulsion polymer, its solution counterpart, and a chemically modified version of the solution polymer showed a 23% hysteresis reduction in going from emulsion to solution and an additional 15% reduction for the chemically modified polymer, to provide an overall reduction of 38% [36]. To date, this type of chemical modification is only possible using anionic techniques.



Polybutadiene (BR)

Next to SBR, polybutadiene is the largest volume synthetic rubber produced. Consumption was approximately 2,166,000 metric tons in 2009 [37].

Monomer Production

Butadiene monomer can be produced by a number of different processes. The dominant method of production is as a by-product from the steam cracking of naphtha to produce ethylene. The butadiene is recovered from the C-4 fractions by extractive distillation [38]. "On-purpose" butadiene is generally produced by dehydrogenation or oxidative dehydrogenation of four-carbon hydrocarbons [39,40].

Polymer Production Process

Polybutadiene is usually produced by alkali metal, and transition metal coordination solution processes. Most production is based on the solution processes because of the ability to obtain preferred microstructures by these routes.

Alkali-metal-based polymerization (usually organolithium) produces a product with about 36% *cis*, 54% *trans*, and 10% vinyl. The polymerization process is conducted in an aliphatic hydrocarbon under an inert atmosphere in either a batch or a continuous mode. Because of the characteristics of this polymerization system, polymers of extremely narrow molecular weight distribution and low gel can be produced [41]. The narrowest distribution is produced via batch polymerization. Coupled (star branched) and end-functional polybutadienes are possible using organolithium technology due to the living anion on the chain end which is available for further reactions. Upon the addition of polar agents, such as ethers or amines, the organolithium initiators can produce polybutadienes with vinyl contents up to 100% [42]. The vinyl content can be controlled by the ratio of modifier to catalyst and the polymerization temperature, with lower temperatures favoring increased vinyl formation. Even with high vinyl contents such polymers do not crystallize because of the atactic nature of the vinyl units. High *cis* polybutadiene is produced via solution processes using Ziegler–Natta type transition metal catalysts. The major commercial catalysts of this type are based on titanium, cobalt, nickel, and neodymium [43]. Typically the transition metal is used in the form of a soluble metal salt, which can react with an organoaluminum or organoaluminum halide as a reducing agent to give the active species. Because of the active nature of transition metals, the polymer solutions are treated to deactivate or remove such materials from the final product.

All of these catalysts produce products with 90% or higher *cis* content. The neodymium system is reported to produce the highest *cis* (98–99%) with the most linear chain structure [44]. The highest branched *cis* BR is produced with the Co system with Ni giving intermediate branching.

All the solution processes require high efficiency in recovering the solvent. The most widely used process consists of termination of the polymerization and the addition of antioxidant to the polymer solution. The solution may be treated to remove catalyst residue and then transferred into an agitated steam stripping vessel in which unreacted monomer and solvent are flashed off, leaving the rubber as a crumb slurry in water. The water–crumb slurry then is dewatered and dried. The recovered monomer/solvent is recirculated to a series of distillation columns to recover

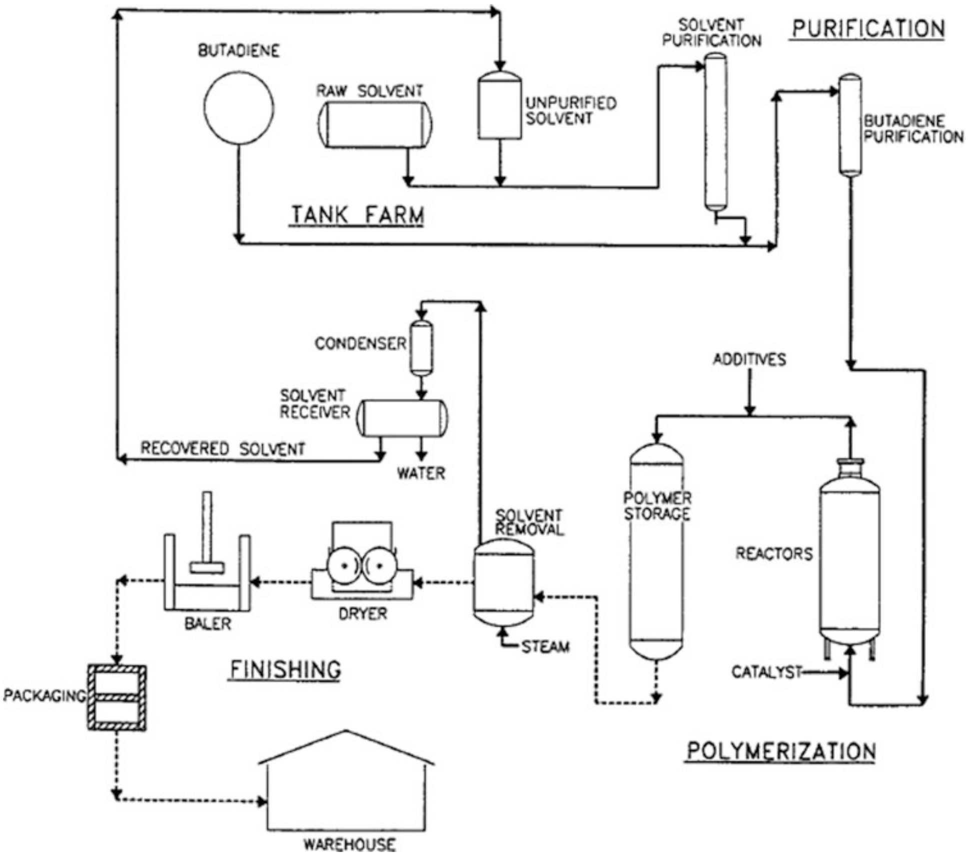


Fig. 16.2 Flow diagram for a typical solution process for the manufacture of polybutadiene (courtesy of Firestone Polymers, Akron, Ohio)

monomer and purify the solvent. As both the anionic and the coordination catalyst systems are highly sensitive to impurities such as water, the purification system is very critical for satisfactory process control (Fig. 16.2).

Uses

The major use for polybutadiene is in tires, with over 70% of the polymer produced used by the tire industry. Cured polybutadiene has excellent low-temperature properties, high resiliency, and good abrasion resistance due to its low glass transition temperature. However, this same fundamental property also leads to very poor wet skid resistance. For this reason, polybutadiene is blended with other polymers such as natural rubber and SBR for use in tread compounds. In general, polybutadiene is a poorer-processing polymer than SBR, but this is generally not a problem as it is blended with other polymers in use. The very high *cis* polymers have the potential for strain-induced crystallization, which can lead to improved green strength and increased cut growth resistance in the cured product. High *cis* polybutadiene is reported to have a melting point of 6°C [45].

The other major use for polybutadiene is as an impact modifier in plastics, in particular high impact polystyrene (HIPS) and acrylonitrile–butadiene–styrene resin (ABS). In the HIPS application the rubber is dissolved in the styrene monomer, which is then polymerized via a free-radical mechanism. A complex series of phase changes occurs, resulting in small rubber particles containing even smaller polystyrene particles being incorporated into a polystyrene matrix. The rubber is added to increase impact strength. Because of the unique morphology that is formed, low levels of rubber (typically around 7%) provide rubbery particles having a volume fraction of 30–40%. This morphology leads to high impact at very low rubber levels, providing good stiffness and hardness [46].

There is also a fairly large market for high *cis* BR in solid core golf balls. In this application, the polymer is compounded with zinc acrylate and the mixture is cured with peroxide [47]. This produces an ionically cross-linked compound that has outstanding resilience. The covers are also ionomers with superior cut resistance. In the last few years the golf ball market has been shifting away from the traditional wound ball to these new solid core balls that use polybutadiene.

Ethylene–Propylene Rubber

There are two general types of polymers based on ethylene and propylene: ethylene–propylene rubber (EPM) and ethylene–propylene terpolymer (EPDM). EPM accounts for approximately 20% of the polyolefin rubber produced. Comprising a totally saturated polymer, these materials require free-radical sources to cross-link. EPDM was developed to overcome this cure limitation. For EPDM a small amount (less than 15%) of a nonconjugated diene is terpolymerized into the polymer. One of the olefinic groups is incorporated into the chain, leaving its other unsaturated site free for vulcanization. This ensures that the polymer backbone remains saturated, with corresponding stability, while still providing the reactive side group necessary for conventional cure systems. The nonconjugated dienes used commercially are ethylidene norbornene, 1,4 hexadiene, and dicyclopentadiene. The selection of the termonomer is made on the basis of the reactivity of the termonomer, both in polymerization and in vulcanization. The estimated 2009 worldwide consumption (excluding CPEQ) was 932,000 metric tons.

Monomer Production

Ethylene and propylene are produced primarily by the cracking of naphtha. They also are available from the fractionation of natural gas. Ethylidene norbornene is produced by reacting butadiene with cyclopentadiene. 1,4 Hexadiene is produced from butadiene and ethylene. Dicyclopentadiene is obtained as a by-product from the cracking of heavy feedstocks to produce ethylene.

Polymer Production

There are two processes used to produce EPM/EPDM: solution and suspension. In either case a Ziegler–Natta type catalyst is used (aluminum alkyl or aluminum alkyl chlorides and a transition metal salt). The most generally used transition metal is vanadium in the form of the tetrachloride or the oxytrichloride [48]. The solution process is similar to that used for other solution polymers. The polymer cement can be finished by stream stripping and drying of the resulting crumb [49].

In the suspension process, the polymer is suspended in the monomer propylene. This process offers the advantages of being able to operate at higher solids owing to the lower viscosity of a suspension compared with a solution at comparable solids. Other advantages are simple heat removal by the evaporative cooling of the propylene, more uniform reactor temperature profile, and ease of production of high molecular weight or semi-crystalline polymers [50].

A specially developed titanium-based catalyst has been used in the suspension process for EPM and EPDM where the termonomer is low-boiling. The advantages claimed, in addition to those characteristic of the suspension process, are better structural control and high catalyst efficiency, resulting in a high-purity product without requiring catalyst removal [51].

The polymer composition for both EPM and EPDM is usually in the 40/60 to 60/40 ethylene/propylene ratio. Outside these ranges, the polymer will start to crystallize because of either polyethylene or polypropylene blocks.

Use

EPM/EPDM polymers exhibit outstanding resistance to heat, ozone, oxidation, weathering, and aging due to the saturated backbone. They have low density, are miscible with aliphatic and naphthenic oils, and maintain acceptable properties at high filler loadings. They are used in single-ply roofing, wire and cable, automotive parts, impact modification of polypropylene, and viscosity index additives for automotive oils. They also can be used in producing thermoplastic olefin elastomers by blending with polypropylene, which may be partially grafted or cross-linked by dynamic vulcanization. These “polymer alloys” will be discussed in the section “Thermoplastic Elastomers.” Although at one time EPDM was expected to become the major polymer for tires, this market has not materialized for a variety of processing and performance reasons [52].

Butyl Rubber

Butyl rubber is one of the older synthetic rubbers, having been developed in 1937. Because of the saturated nature of a polyolefin elastomer, the commercial polymer is actually a copolymer of isobutylene and isoprene. The isoprene is added to provide cure sites. In addition, halogenated (bromo or chloro) derivatives are available. The halogenated products improve the mixing and cure compatibility with the more common unsaturated rubbers such as natural or SBR.

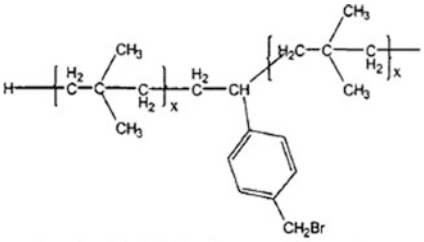
Monomer Production

Isobutylene is obtained as a by-product from petroleum and natural gas plants. The monomer must be highly purified to assure high molecular weight.

Production Process

Butyl rubber is produced at very low temperature (below −90°C) to control the rapid exotherm, and to provide high molecular weight. The process consists of charging isobutylene along with isoprene (2–4%) with an inert diluent such as methyl chloride to a reactor to which a Friedel–Crafts catalyst is added. The polymerization is very rapid, and the polymer forms in a crumb or slurry in the diluent. Heat is removed via the reactor jacket. The slurry is steam-stripped to remove all volatiles. The catalyst is neutralized, and antioxidants are added to the slurry prior to drying [53]. The halogenated derivatives are produced by the direct addition of the halogen to a solution of the isobutylene–isoprene polymer.

During the last 10 years another type of butyl rubber was developed which is derived from a copolymer of isobutylene and *p*-methylstyrene [54]. They are subsequently brominated to varying degrees producing different grades of the elastomer. Bromination occurs selectively on the methyl group of the *p*-methylstyrene providing reactive benzylic bromine functionality, which can be used for grafting and curing reactions.



brominated poly(-isobutylene-*p*-methylstyrene)

Properties and Use

The most important characteristics of butyl rubber are its low permeability to air and its thermal stability. These properties account for its major uses in inner tubes, tire inner liners, and tire curing bladders. Because of the poor compatibility of butyl with other rubbers (with respect to both solubility and cure), the halobutyls are preferred. The brominated *p*-methylstyrene-containing butyl rubbers are used in a number of grafting reactions for tire applications and adhesives. Other uses for butyl rubber are automotive mechanical parts (due to the high damping characteristics of butyl), mastics, and sealants [55].

Nitrile Rubber

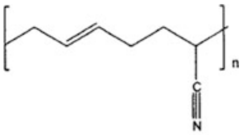
Nitrile rubber was invented at about the same time as SBR in the German program to find substitutes for natural rubber [56]. These rubbers are copolymers of acrylonitrile–-butadiene, containing from 15 to 40% acrylonitrile. The major applications for this material are in areas requiring oil and solvent resistance. The estimated worldwide consumption in 2003 was 303,000 metric tons [57].

Monomer Production

The production of butadiene is discussed in the diene section "Polybutadiene." Although several routes have been developed to produce acrylonitrile, almost all now is produced by the catalytic fluidized-bed ammoxidation of propylene.

Polymer Production

The polymerization process parallels the emulsion process used for SBR. Either a hot or a cold process can be used, with the cold polymerization providing the same improved processing and vulcanizate properties as seen in SBR. Polymerizations are carried to 70–80% conversion and terminated to avoid gel formation. The latex must be stripped to remove unreacted butadiene and acrylonitrile.



butadiene-acrylonitrile copolymer

Properties and Use

As the acrylonitrile content increases in the polymer chain, the properties change predictably. The glass transition temperature increases approximately 1.5°C for each percent increase in acrylonitrile. Properties such as hysteresis loss, resilience, and low-temperature flexibility will correspondingly change. The oil resistance increases with increased acrylonitrile content, as does the compatibility with polar plastics such as PVC. The major market for nitrile rubber is in the automotive area because of its solvent and oil resistance. Major end uses are for hoses, fuel lines, O-rings, gaskets, and seals. In blends with PVC and ABS, nitrile rubber acts as an impact modifier. Some nitrile rubber is sold in latex form for the production of grease-resistant tapes, gasketing material, and abrasive papers. Latex also is used to produce solvent-resistant gloves [58].

Hydrogenated Nitrile Rubber

During the last 25 years several companies have developed hydrogenated grades of nitrile rubber to improve both its thermal stability and solvent resistance. Although the hydrogenation of a polydiene backbone was done as early as the 1920s, real commercial products with acrylonitrile were not introduced until the mid-1980s [59].

Hydrogenated NBR (HNBR) is produced by first making an emulsion-polymerized NBR using standard techniques. It then must be dissolved in a solvent and hydrogenated using a noble metal catalyst at a precise temperature and pressure [60]. Almost all the butadiene units become saturated to produce an ethylene–butadiene–acrylonitrile terpolymer. These "post-polymerization" reactions are very expensive so HNBRs usually command a premium price. HNBR is usually cured with peroxides, similar to ethylene–propylene elastomers, because it has no unsaturation for a conventional sulfur cure system.

Uses

HNBR has many uses in the oil-field, including down hole packers and blow-out preventers, because of its outstanding oil resistance and thermal stability. For the same reasons, it has also found uses in various automotive seals, O-rings, timing belts, and gaskets. Resistance to gasoline and aging makes HNBR ideal for fuel-line hose, fuel-pump and fuel-injection components, diaphragms, as well as emission-control systems.

Chloroprene Rubber

Chloroprene rubber (Neoprene—trade name of DuPont) was one of the earliest synthetic rubbers, first commercialized in 1932. It has a wide range of useful properties but has not become a true general purpose synthetic rubber, probably because of its cost. It does possess properties superior to those of a number of general purpose polymers, such as oil, ozone, and heat resistance; but for these properties other specialized polymers excel. Polychloroprene thus is positioned between the general purpose elastomers and the specialty rubbers.

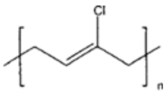
Monomer Production

Chloroprene monomer production starts with the catalytic conversion of acetylene to monovinylacetylene, which is purified and subsequently reacts with aqueous hydrogen chloride solution containing cuprous chloride and ammonium chloride to give chloroprene [61].

Production Process

Polychloroprene is produced by using an emulsion process. Two general types of processes are used: sulfur modified and unmodified. In the sulfur modified process, sulfur is dissolved in monomer and is incorporated into the polymer chain. Upon the addition of thiuram disulfide-type materials and under alkaline conditions, some of the sulfur bonds are evidently cleaved to give the soluble polymer. In the unmodified process chain transfer agents are used. If neither the sulfur modified nor the chain transfer system is used, the resulting polymer is a gelled tough material. Typical polymerization systems consist of rosin acid soap emulsifier and persulfate catalyst. Conversions of 80–90% are obtained. Polymerizations are run at around 40°C. For the modified polymer, the thiuram disulfide is added after polymerization, and the latex is aged to allow the peptization (chain scission) reaction to occur. Acidification stops the peptization reaction. The latex is vacuum-stripped and coagulated using a cold drum dryer process. The coagulated rubber is washed and dried.

The polymerization produces primarily *trans*-1,4-polychloroprene. The *trans* content can be increased somewhat by lowering the polymerization temperature [62,63]:



polychloroprene

Properties and Uses

Polychloroprene is stable to oxidation and ozone. It also is flame resistant, and its oil resistance is better than that of general purpose rubbers. Its major disadvantage, other than cost, is relatively poor low-temperature properties. Because of the high stereoregularity, polychloroprene will strain crystallize, giving good tensile to unfilled stocks. At low temperatures the polymer can crystallize, making processing more difficult. Polymer made at lower temperatures will show higher unfilled tensile properties and more rapid crystallization due to the higher *trans* content (i.e., less disruption of the crystal structure). The major end uses are conveyor belts, V-belts, hoses, and mechanical goods such as wire insulation, O-rings, and gaskets. It also has found use in single-ply roofing and adhesives [64].

Silicone Elastomers

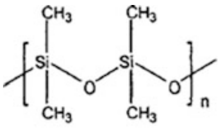
Silicone elastomers represent a rather unique group of polymers in that they consist of alternating silicon–oxygen bonds to form the polymer chain backbone. Side groups off the silicon atoms are selected to provide very specific properties that differentiate one type from another. The most common side group is the dimethyl structure. Replacement of small amounts of the methyl group with vinyl provides sites for cross-linking. Phenyl groups are used to improve low-temperature properties. Fluorosilicones are produced by replacing the methyl with trifluoro-propyl units. The addition of bulky phenyl side groups leads to an increase in the glass transition temperature. However, the disruption caused by such groups leads to the desired reduction or elimination of crystallization, which is critical for low-temperature properties. Such polymers have glass transition temperatures around −110°C, which is the range of the lowest T_g carbon-backbone polymers.

Monomer Production

The actual polymerization process involves a ring-opening reaction of dimethyl-substituted cyclic siloxanes. The preparation of the cyclic materials starts with the production of pure silicon via the reduction of quartz with coke in an electric arc furnace. The silicon metal then reacts with methyl chloride to give a mixture of silicones, from which dimethyldichlorosilane is removed by distillation [65]. Subsequent hydrolysis gives the cyclic dimethylsiloxane.

Polymer Production

The polymerization process involves an equilibrium ring-opening reaction carried out in the bulk state, which can be catalyzed by acids or bases:



polydimethylsiloxane

Uses

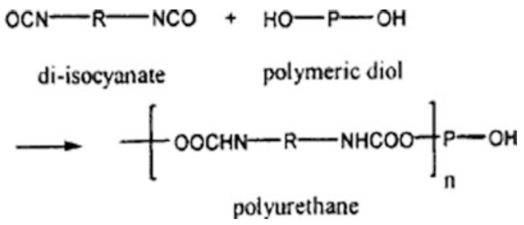
Silicone rubber offers a set of unique properties to the market, which cannot be obtained by other elastomers. The Si–O backbone provides excellent thermal stability and, with no unsaturation in the backbone, outstanding ozone and oxidative stability. The very low glass transition temperature, combined with the absence of low-temperature crystallization, puts silicones among the materials of choice for low-temperature performance. The fluoro-substituted versions provide solvent, fuel, and oil resistance along with the above-mentioned stability advantages inherent with the silicone backbone.

The gum polymer has rather poor tensile properties when cured, but these properties can be greatly improved by the use of silica-reinforcing agents. These systems exhibit some of the greatest improvements in properties by filler addition; and because this improvement is significantly higher for silica than for other reinforcing agents, it is assumed that direct bonding occurs between the silica and the polymer. The silicone materials also may have very low surface energy, which accounts for their nonstick characteristics. Because of the inherent inertness of the materials, they have been widely used for medical purposes within the body. The largest use of silicone is in sealant and adhesive applications.

Several different methods have been developed to cure silicones. Free-radical cures are possible for those polymers containing vinyl groups. The largest-volume process, however, involves room-temperature ulcanizetions, which can employ either a one-component or a two-component system. In the one-component system, a cross-linking agent such as methyltriacetoxysilane is used. With exposure to moisture, hydrolysis of the cross-linking agent leads to the silanol reactive cure site, so such materials must be compounded and stored free of moisture [66].

Polyurethane Rubber

A wide range of materials is included in this class. The common feature is the use of chain extension reactions to provide products with acceptable commercial properties. The chain extension reaction effectively reduces the actual number of chain ends, thereby eliminating the generally poor properties observed when very low-molecular-weight polymers are cross-linked. The chain extension step involves the reaction of a difunctional polymeric polyol with difunctional organic isocyanates to give the polyurethane:



The most used polyols have a polyester or a polyether backbone. A wide variety of isocyanates are used, with toluene di-isocyanate, *m*-phenylene di-isocyanate, and hexamethylene di-isocyanate the most common.

Raw Materials

The largest-volume polyether used is obtained from propylene oxide polymerized under basic conditions. Polyester polyols are produced from a number of different materials involving diacids and diols to give the ester linkage. Aliphatic polyesters generally are used for elastomers to impart chain flexibility.

The production of isocyanates is based on the reaction of phosgene with primary amines. Toluene di-isocyanate is the most frequently used di-isocyanate.

Uses

The urethane elastomers are complex-segmented or block polymers. Soft, noncrystalline blocks are provided by the polyether or aliphatic polyester long chains, whereas stiff, hard blocks are produced by the reaction of aromatic diisocyanates with low-molecular-weight materials such as diols or diamines. The hard blocks can phase-separate to provide a physical rather than a chemical cross-link, similar to those to be discussed for the butadiene–styrene TPE, although the domain size is much smaller for the polyurethanes. Chemical cross-links can be introduced by using tri-functional materials or by adjusting the stoichiometry to allow additional reactions with the urethane or urea structures to give allophanate or biuret linkages.

Polyurethane rubbers can have high tensile strengths, excellent tear strength, and good abrasion and chemical resistance. The greatest disadvantage is the hydrolytic instability of the urethane linkage. A major use is in automotive bumpers and facias. These materials are made in a reaction injection molding process. Castable urethanes are used to produce solid rubber wheels and printing rolls. Millable urethanes can be processed on conventional rubber equipment. Unsaturation can be introduced to allow the use of conventional rubber cure systems. Cures also are possible by reaction of the active hydrogens with materials such as high molecular weight polyfunctional isocyanates that are nonvolatile at cure temperatures.

Thermoplastic polyurethane elastomers are processed by injection molding and other processes used for thermoplastics. Small gears, seals, and even automotive fender extensions can be produced by this means.

By selecting from the large number of possible reactions and stoichiometry, properties can be tailored to meet a very wide range of applications [67–69].

Modified Polyethylene Rubbers

Elastomeric polymers can be produced by the chlorination or chlorosulfonation of polyethylene. Both products start with polyethylene, either in solution or in aqueous suspension, which then is reacted to give the specified degree of substitution to obtain the desired properties. Sufficient substitution is necessary to disrupt the regularity of the polymer chain, changing it from the crystalline polyethylene plastic into amorphous elastomers.

Chlorinated Polyethylene

The chlorinated products contain around 40% chlorine. These materials must be stabilized with metal salts, like other chlorinated elastomers and plastics. Peroxide cross-linking generally is used. Being saturated, the materials have excellent weather and ozone resistance and can be used over a temperature range of −65 to 300°F. The high chlorine content imparts oil resistance and relatively slow rates of burning. Typical applications where this combination of properties is required include hoses for chemical or oil resistance, tubing, and belting. In comparison with plasticized PVC, these materials have better low-temperature properties and do not suffer a loss of plasticizer because none is required.

Chlorosulfonated Polyethylene

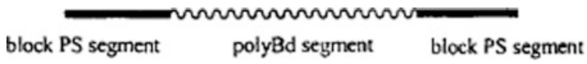
Reaction with sulfur dioxide in addition to chlorine introduces cross-linking sites into the polymer chain. Sulfur contents in the range of 1.0–1.5% are used, with chlorine contents of 25–40%. Curing is accomplished by using metallic oxides, sulfur-bearing organic compounds, and epoxy resins. These materials have outstanding ozone resistance and show little color change upon light exposure. Good resistance to oils, heat, oxidation, weather, and corrosive materials also is exhibited. Applications include pond and pit liners, coated fabrics, light-colored roofing membranes, wire and cable insulation, chemical hose, and belting [70,71].

Thermoplastic Elastomers

This class of elastomeric materials is called “thermoplastic” because they contain thermally reversible cross-links of various types. The types of crosslinks vary from phase-separated polystyrene domains, such as in SBS elastomers to ionic cluster cross-links in the ionomers. The beauty of these noncovalent interactions to form cross-links is that when the material is heated, the cross-links are broken. This allows the polymer to flow and be processed, and also recycled. When cooled, the cross-links reform and the material becomes strong again. All TPEs are two-phase systems where there is a soft, rubbery “continuous phase” and a hard “dispersed phase” which does not flow at room temperature. The first two polymers to be discussed are “block copolymers” and “ionomers.”

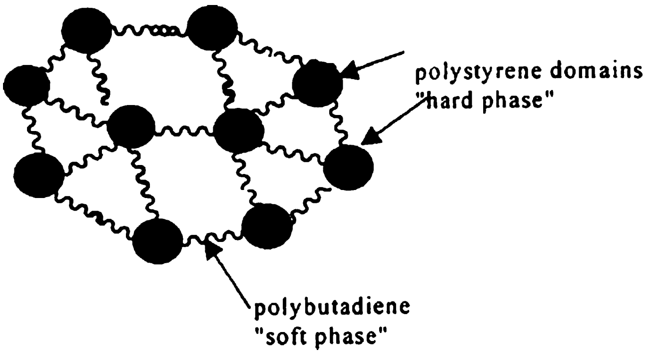
Block Copolymers

A very popular and useful TPE is made from blocks of styrene and butadiene monomers using anionic polymerization techniques, which was described in the solution SBR section above. They are made up of short chains of polystyrene (usually 8,000–15,000 MW), followed by a much longer chain of polybutadiene (about 60,000 MW), and capped off by another short chain of polystyrene, hence the name SBS. Similar polymers are prepared using isoprene instead of butadiene (SIS). The differences between SBS and SIS will be discussed later in the subsection "Uses."



The linear polymers, as shown above, can be built up by the sequential addition of monomer or by coupling the living anionic chains using compounds like dichloro dimethylsilane. Hence, the base polymer would have styrene polymerized first, followed by butadiene, and then addition of the coupling agent. If a multifunctional coupling agent such as silicon tetrachloride is used, a radial block or "star-branched" SBS is formed.

The polystyrene is highly insoluble in the polybutadiene so the PS chains cluster together and phase-separate into domains. Since there is much more polybutadiene (PBD) than polystyrene, the PBD becomes the continuous phase containing dispersed particles of PS which act both as cross-links and reinforcing agents. Every PBD chain is tied to a PS chain on both ends so a very strong cross-linked network is formed. If the PBD is tied to only one PS, then a "diblock" polymer is formed which has very little strength. When this network is heated to above the glass transition of polystyrene (100°C) the PS domains break down and begin to flow, so the polymer can be processed by injection molding or extrusion. Upon cooling to below 100°C, the domains (cross-links) reform and the material becomes strong again. SBS elastomers can have tensile strength as high as conventional thermoset elastomers which may approach 4,000 psi.



The butadiene blocks can be hydrogenated, as mentioned above with hydrogenated nitrile, to form SEBS polymers having better thermal stability and chemical resistance. The EB stands for ethylene–butylene, which are the structures formed after the butadiene segments have been hydrogenated.

Other block copolymers which are useful are based on polyesters, polyurethanes, and ethylene–propylene. The first two have been discussed in other sections and the ethylene—propylene blocks will be discussed below in the subsection "Metallocene Polymers."

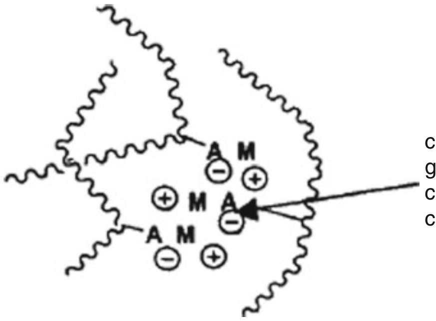
Uses

SBS copolymers are used in a wide variety of applications because of their clarity, toughness, and ease of processing. A major application is hot melt adhesives where they are compounded with hydrocarbon resins and oil [72]. The SIS polymers are very popular in adhesives because the isoprene segments tend to undergo chain scission during aging instead of cross-linking, which is observed in butadiene polymers. This leads to better retention of adhesion after aging; however, the SIS polymers usually have poorer initial strength compared with SBS. Another high-volume use is in toughening of asphalt compounds for paving, crack sealants, and roofing. The SBS improves rutting and low-temperature performance in paving and crack resistance in roofing applications [73]. SEBS would be the elastomer of choice in many adhesive and asphalt applications because of its superior aging properties; however, its high cost is prohibitive. SBS also is used widely to toughen polystyrene and high-impact polystyrene. This polyblending technique is used to toughen a number of plastics. Various other injection molding and extrusion applications include shoe soles and toys. SBS is limited in use because of its poor high-temperature performance.

Ionomers

Ionomers are copolymers in which a small portion of the repeat units have ionic pendant groups on usually a nonpolar backbone. The ionic groups tend to separate themselves into domains similar to the polystyrene segments in the SBS rubber because they are insoluble in the nonpolar polymer chains. Therefore, these ionic clusters serve as cross-links up to temperatures where they tend to disassociate. Most commercial grades of ionic elastomers are based on ethylene and propylene monomers.

The polymer backbone usually contains sulfonyl or carboxylic acid groups and the metal counterion can be zinc, calcium, sodium, or lithium. The properties are highly dependent on the metal cation because they determine the temperature at which the ion clusters disassociate. The ionic cross-links usually impart outstanding tensile and tear strength properties because these cross-links are very good energy absorbers (hysteretic) due to their mobility. A process known as ion hopping provides one source for absorbing energy.



clusters of ionic groups forming a crosslink M+ = metal cation, A− = anion

Uses

Solid golf balls are a good example of ionic cross-linked materials [74]. The covers are usually lithium or sodium cross-linked plastics but the cores are also a type of ionomer. The cores are high *cis* polybutadiene which is compounded with zinc diacrylate (ZDA) and then peroxide cured. During the curing process, the ZDA grafts to the BR backbone creating a material with a very high cross-link density where the cross-links are clusters of zinc carboxylates with very high resilience. Other applications include heat-sealable food packaging, automotive trim, footwear, foamed sheets (mats), and interlayers for bulletproof glass.

Metallocene Elastomers

Metallocene catalysts are the latest innovations to make a big impact in the polymer industry. They have been used mostly to make new poly-olefin plastics, such as very-high-molecular-weight, bullet-proof polyethylene, but they have also been used to make elastomers. The catalysts make very regular "stereospecific" polymers similar to the Ziegler–Natta catalysts. They are based on various metals, such as zirconium, complexed with cyclopentadienide anions. This type of compound is called a "zirconocene" and is used with organoaluminum to make highly regular polymers. The catalyst has the ability to flip back and forth from making *atactic* to *isotactic* polypropylene in the same polymerization. The alternating tacticity of the polymer breaks up the crystallinity of the chains and yields an elastomer. Metallocene catalysts are currently very expensive and cannot yet polymerize dienes such as butadiene, so they have only enjoyed limited commercial success in elastomers. However, this is one of the most intense fields of polymer research and many new product breakthroughs are expected in the near future.

Rubber–Plastic Alloys

A discussion on TPE would not be complete without mentioning the elastomers produced by simply blending rubbers and plastics in an internal mixer using a process known as "dynamic vulcanization." The simplest of these elastomers is based on polypropylene and EPDM rubber. Products are made with various rubber contents (hardness values) by simply mixing the rubber and plastic in a Banbury or an extruder at high temperature while cross-linking the EPDM in situ with a phenolic curing resin or other curing agents such as sulfur or peroxides. The resulting blend processes like polypropylene but is actually an elastomer because of the cross-linked rubber phase which it contains. It can be reprocessed and recycled like other TPEs. The process is described in an excellent review article by Abdou-Sabet [75]. Other alloys are based on nylon and NBR to obtain better thermal stability and solvent resistance. Constant improvements are being made in polymer alloys by using different combinations of plastics and rubbers and also new grafting and cross-linking chemistry to achieve properties more like conventional thermoset rubbers.

Uses

The applications for this type of elastomer have been limited because of the melting point of the plastic phase and poor compression set. For these reasons, alloys probably will never be successful as tire materials but they have many other automotive applications, including instrument panels, cowl vents, body panels, and bumpers. They also are used in appliances, lawn and garden equipment, and as grips on tools.

Plasticized Polyvinyl Chloride

Generally one thinks of polyvinyl chloride as a rigid plastic, which it is, with a glass transition temperature around 85°C. However, the addition of polar chemicals such as dioctyl phthalate can reduce the glass transition temperature below room temperature, producing a rubbery material.

Monomer Production

Vinyl chloride is produced primarily from ethylene, which is converted to ethylene dichloride either by chlorination or oxychlorination. The ethylene dichloride is cracked to give vinyl chloride and hydrogen chloride.

Production Process

Polyvinyl chloride is produced by the free-radical polymerization of vinyl chloride. Bulk, emulsion, solution, and suspension polymerization processes have been used.

The plasticized product can be produced by mixing the polymer and plasticizers at elevated temperatures, also by dry blending in which the plasticizer is absorbed into the resin and then heated. Solution blending is sometimes used, as well as the plastisol process, in which fine polyvinylchloride powder is dispersed in the plasticizer which is relatively stable until it is heated.

Properties and Use

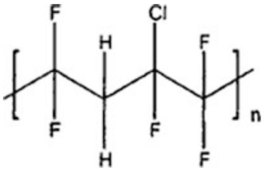
Plasticized polyvinyl chloride can be regarded as the first thermoplastic elastomer, as it is used in an uncross-linked form. Because of the lack of cross-linking, this material exhibits high rates of creep and stress relaxation. As with other TPE, these disadvantages worsen as the temperature is increased. Although the polymer is saturated, it must be stabilized for use to prevent dehydrochlorination at processing temperatures. Because of the high chlorine content, polyvinyl chloride has excellent flame resistance as well as good electrical properties. Even at high plasticizer levels, these materials have marginal elasticity when compared with most other elastomers. Despite their shortcomings, they are used in many areas where they compete with other rubbers.

One of the larger uses is for wire and cable insulation. The flame resistance makes this the material of choice for residential wiring, extension cords, and so on. Inexpensive garden hose represents another large-volume use. Sports balls can be produced by rotational molding of plastisols. The low fabrication cost allows such products to dominate the lower-price-range market [76,77].

Fluorocarbon Elastomers

The fluoroelastomers were developed as specialty materials for high heat applications and solvent resistance. By 2009 there was over 15,800 metric tons of fluoroelastomer capacity worldwide [78]. The elastomers were by-products of Plunkett's 1938 discovery of polytetrafluoroethylene. Copolymers of olefins with vinylidene fluoride were found to be leathery, whereas elastomers were made if tetrafluoroethylene (TFE) or trifluoropropene was used. The first commercial product was produced as a copolymer of vinylidene fluoride and chlorotrifluoroethylene (CTFE) and was called Kel-F. The later, more stable polymers used hexafluoropropene (HFP) with TFE. They were first commercialized by DuPont under the trade name Viton and then a few other companies added their own fluoroelastomers. These elastomers are usually prepared by radical polymerization in emulsion using catalysts such as ammonium persulfate and sometimes with chain transfer agents such as carbon tetrachloride or halogen salts [79]. Caution must be exercised when using these polymerizations because the fluoromonomers can be explosive. They are usually isolated by coagulation and normal rubber drying techniques and sold in the form of pellets, slabs, or rubber crumb.

As a result of being very unreactive, the fluoroelastomers cure very slowly and usually require a high-temperature post-cure. The curatives are designed to remove hydrogen fluoride to generate a cure site which can react with a diamine or bisphenol. They are also cured with organic peroxides. Most fluoroelastomers can be compounded using normal rubber processing equipment such as rubber mills and internal mixers. Processing aids such as dioctyl phthalate or waxes can be used to obtain smoother extrusions and better mold release.



Uses

Fluoroelastomers have outstanding heat resistance when cured. Some vulcanizates have almost indefinite service life at temperatures up to 200°C. The perfluoro polymers, such as DuPont's Kalrez, have short-term usage at up to 316°C and extended service as high as 288°C. This polymer is extremely expensive and is only offered in the form of finished parts, usually O-rings, seals, or gaskets. The fluoroelastomers also have excellent solvent and ozone resistance making them ideal for automotive fuel hoses. Although many of the applications involve small seals, O-rings, gaskets, and hose, the single largest application is flue-duct expansion joints. The polymer's resistance to high temperature and wet acidic flue gases are critical in this application.

References

1. Whitley GS, Davis CC, Dunbrook RF (eds) (1954) Synthetic rubber. Wiley, New York
2. Blackely DC (1983) Synthetic rubber: their chemistry and technology, Chapter 2. Applied Science, London
3. http://www.therubbereconomist.com/The_Rubber_Economist/Home.html
4. Fox TG, Grateh S, Loshaek S (1956) Rheology—theory and applications, vol 1. Academic, New York
5. Billmeyer FW Jr (1984) Textbook of polymer science. Wiley, New York

6. Odian G (1970) Principles of polymerization. McGraw-Hill, New York, pp 19–24
7. Flory PJ (1953) Principles of polymer chemistry, Chapter 7. Cornell University Press, Ithaca, NY
8. Rodriguez F (1982) Principles of polymer systems. McGraw-Hill, New York, pp 138–142
9. Rudin A (1982) The elements of polymer science and engineering. Academic, New York, p 126
10. Rudin A (1982) The elements of polymer science and engineering. Academic, New York, p 402
11. Krause S (1972) J Macromol Sci C7(2):251
12. Aklonis JJ (1972) Introduction to polymer viscoelasticity. Wiley, New York
13. Oberster AE, Bouton TC, Valaitis JK (1973) Die Angewandte Markomoleculare Chemie. 29/30: 291
14. Flory PJ (1949) J Chem Phys 17:223
15. Natta G (1969) in Polymer chemistry of synthetic elastomers, part 1. In J. Kennedy, E. Tornqvist (Eds.) Chapter 7 Wiley, New York
16. Semegen ST, Cheong SF (1978) Vanderbilt rubber handbook. RT Vanderbilt Company, Norwalk, pp 18–41
17. Webster CC, Baulkwill WJ (1989) Rubber, Chapters 1 and 11. Wiley, New York
18. Brandrup J, Immergut E (1975) Polymer handbook, vol III. Wiley, New York, p 54
19. Treloar L (1949) The physics of rubber elasticity. Clarendon, Oxford
20. Poh WN (1989) Elastomers. p. 12
21. International Institute of Synthetic Rubber Producers (1973) Synthetic rubber: the story of an industry. IISRP, New York
22. Stavelv FW (1956) Ind Eng Chem 48:778, presented at Div. Rubber Chem., ACS, Philadelphia, 1955
23. Eng. Patent 827365, to Goodrich–Gulf Chem., Dec 1954, priority data
24. Worldwide Rubber Statistics 2004 (2004) International Institute of Synthetic Rubber Producers, Houston, p 11
25. Worldwide Rubber Statistics 2009 (2009) International Institute of Synthetic Rubber Producers, Houston, p 5
26. Renninger TJ (1990) Presentation to international rubber study group, Ottawa, Sept 1990
27. Faith WL et al (1957) Industrial chemicals. Wiley, New York, p 731
28. Dunbrook RF (1947) India Rubber World 117:203–207
29. Gardon JL (1970) Rubber Chem Technol 43:74–94
30. Storey EB (1961) Rubber Chem Technol 34:1402
31. Kuntz L (1969) J Poly Sci 54:569
32. Bouton T, Futamura S (1974) Rubber Age 3:33
33. Mostert S, Van Amergongen G (1968) British Patent 1,136,189
34. Day G, Moore D (1985) Paper at 26th annual meeting, IISRP, May 1985
35. Quirk RP (1996) Rubber Chem Technol 69(3):444
36. Oshima N, Salcacobore M, Tsutsumi F (1986) Paper at 27th annual meeting, IISRP, May 1986
37. Worldwide Rubber Statistics 2009 (2009) International Institute of Synthetic Rubber Producers, Houston, p. 5
38. Ponder T (1976) Hydrocarb Process 55(10):119–121
39. Womeldroph D (1958) Am Petrol Inst (14 May 1958)
40. Welch M (1978) Hydrocarb Process 57(11):131–136
41. Adams H, Farhat K, Johnson B (1966) Ind Eng Chem Prod Dev 5(2):127
42. Halasa A, Schulz D, Tate D, Mochel V (1980) In: Stone F, West R (eds) Advances in organometallic chemistry, vol 18. Academic, New York
43. Tate D, Bethea T (1985) Encyclopedia of polymer science and engineering, vol 2. Wiley, New York, pp 568–572
44. Lauretti F, Gargani L (1987) 27th annual meeting, IISRP, June 1987
45. DeChirico A, Lamzani P, Eaggi E, Bruzzone M (1974) Makromol Che 175:2029
46. Buckenell C (1977) Toughened plastics. Applied Science, London
47. Maruko T (2000) Solid Golf Ball. US Patent No. 6,071,201
48. Natta G (1961) J Poly Sci 51:411
49. Lukach C, Spurlin H (1964) Copolymerization. In G. Ham (ed.) Interscience, New York
50. Crespi G, DiDrusco G (1969) Hydrocarb Process 48:103–107
51. Vandenberg E, Repka B, in Ref. 12, Chapter 11
52. Scagliotti F, Milani F, Galli P (1985) Paper at 26th annual meeting, IISRP, May 1985
53. Miles D, Briston J (1963) Poly Technol. Chemical Publishing, New York, p 299
54. Powers KW (1993) Para-alkylstyrene/isoolefin copolymers. US Patent No. 5,162,445
55. “Butyl Rubber Reaches 50 Year Mark,” *Elastomerics*, 30–31 (Mar. 1988).
56. Konrad E, Tschunkur E (1934) G. Farbenindustrie. U.S. Patent 1,973,000 to 1
57. Worldwide Rubber Statistics 2000 (2004) International Institute of Synthetic Rubber Producers , Houston, p 9
58. McKenzie G (1986) Paper at 27th annual meeting, IISRP, May 1986
59. Hashimoto K et al (1985) Paper at 26th annual meeting, IISRP, May 1985
60. Buding H (1986) Production of hydrogenated nitrile rubber. US Patent No. 4,581,417
61. Whitley GS, Davis CC, Dunbrook RF (eds) (1954) Synthetic rubber. Wiley, New York, p 770
62. Miles D, Briston J (1963) Poly Technol. Chemical Publishing, New York, p 305
63. Blackely DC (1983) Synthetic rubber: their chemistry and technology, Chapter 2. Applied Science, London, pp 175–194
64. Brodrecht L (1989) Chemical economics handbook. SRI International, Menlo Park
65. Polmateer K (1988) Rubber Chem Techn 16(3):470
66. Semegen ST, Cheong SF (1978) Vanderbilt rubber handbook. RT Vanderbilt Company, Norwalk, pp 216–232
67. Natta G (1969) in Polymer chemistry of synthetic elastomers, Part 1, J. Kennedy and E. Tornqvist (eds) Chapter 8. Wiley, New York
68. Morton M, Fetters L (1977) Polymerization processes, Chapter 9. Wiley, New York
69. Handbook of elastomers. Marcel Dekker, New York, pp 643–659, 1988
70. Broadrecht L, Mulach R, Tauchiya K (1989) Chemical economics handbook—elastomers. SRI International, Menlo Park
71. Semegen ST, Cheong S (1978) Vanderbilt rubber handbook. RT Vanderbilt Company, Norwalk, pp 18–41
72. Kraton Polymers for Adhesives & Sealants, Shell online literature, Shell Chemical Web site.
73. Bull A, Vonk W (1988) Shell Chemicals Technical Manual TPE 8: 15 Report. (Shell Web site)
74. Maruko T (2000) Solid golf ball. US Patent No. 6,071,201
75. Abdou-Sabet S (1996) Rubber Chem Technol 69(3):476
76. Penn W (1971) PVC technology. Applied Science, London
77. Miles DC, Briston JH (1963) Polymer technology. Chemical Publishing, New York, p 159
78. Worldwide Rubber Statistics 2009 (2009) International Institute of Synthetic Rubber Producers, Houston
79. Morton M (1995) Rubber technology. Chapman & Hall, London, p 410

The Agrochemical Industry

17

K.N. Singh and Kavita Merchant

Introduction

Agrochemical Industry comes under life science segment of Chemical Industry and deals with production and distribution of pesticides and fertilizers to increase the crop yields. As Pharmaceuticals are meant for life, Agrochemicals are for livelihood.

Although overall agrochemicals are very small part of chemical industry, their role in sustainable growth is very important as the world's population is estimated to grow from 6.8 billion today to >9.0 billion by 2050. Global demand for food is expected to double by then as per UN projections and estimate.

There is no dispute whatsoever that for meeting the growing food need we will essentially be called upon to produce more and more food with less cultivable land. This necessitates use of agrochemicals in general and pesticides in particular under integrated crop and pest management program to enhance productivity and prevent 30–40% avoidable losses caused by various pests during production, processing and storage. Pesticides were and continue to be an integral part of high-yield farming success.

This chapter describes mainly pesticide active ingredients and formulations manufactured for treatment of seeds, crops and farm products to protect from various insect pests, weeds, plant pathogens, microbes and germs. Pesticides used in non-crop sectors like animal and public health, disinfectant, home care, etc. are also covered. Fertilizers or fundamental nutrients are arbitrarily excluded from discussion in this chapter.

The global pesticide market is divided into three basic categories based on intellectual property rights (IPR) viz.: Proprietary products; Proprietary off-patent products; Generic products.

A proprietary product has a granted patent in force. The market share of proprietary products is approximately 30%, i.e. ~13 billion of the total pesticide market. For proprietary off-patent products, though the patent has expired, there is no generic competition. The majority of these products are owned by the proprietary companies and adding together both proprietary and off-patent proprietary products, the market share is approximately 70%, which is controlled by a few multinational companies viz. Syngenta, Bayer Crop Sciences, BASF, Dow AgroSciences, Monsanto, Du Pont, FMC, etc.

For a generic product, patent protection has expired and local companies have obtained sales authorization from Federal and State Governments that are independent of the proprietary company's registrations. Generic products account roughly for ~30% of total pesticide market, i.e. $13 billion.

Special mention is provided in the chapter to products introduced in the recent past and going off-patent in 2010 and beyond (please refer Annex 17.1). This section will help the readers to identify and develop generic active ingredients to meet the global demand for food, feed, fibre, plastics and energy as well as to maintain the general hygiene of homes.

The global pesticide market with sales revenue of approximately $44 billion during 2009 comprises of the crop protection sector with sales of $38 billion (86%) and non-crop sector with sales of $6 billion (14%) according to data from market research consultancy Phillips McDougall.

The crop protection sector over the period of last 10 years has achieved growth of 34% from $28.1 billion during 1999 to $31.2 billion during 2005 to $38 billion during 2009. There is significant optimism in the industry that there will be a period of sustained growth in the next 5 years. This optimism is mainly due to increased consumption of pesticides in Latin American countries like Brazil, Argentina, Mexico and Asian countries mainly India and China.

The non-crop sector has a wide variety of applications including forestry, timber treatment, public health, industrial (mainly weed control on railways, highways), home and garden markets, etc. There are many drivers affecting growth of this sector including increase in urban population and more disposable income per capita, part of which is used in hygiene and pest control operations.

K.N. Singh (✉) • K. Merchant
Gharda Chemicals Limited, 48 Hill Road, Bandra (W), Mumbai, India
e-mail: knsingh@gharda.com

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_17, © Springer Science+Business Media New York 2012

Pesticides

Pesticide active ingredients (AI) are chemicals designed to be harmful to a target pest. These are purposely converted into plant protection products (PPP) and introduced into environment for preventing, destroying, repelling or mitigating pests.

People often think pesticide means insecticide. However, it also includes herbicides, fungicides and various other substances used to control target pests viz. rodents and other animals, unwanted plants (weeds), fungi or microorganisms like bacteria and viruses.

Farmers use pesticides to control the pests that can destroy or damage food, fibres and other crops. Health agencies use pesticides to control insects and other organisms known to carry diseases. Hospitals use disinfecting pesticides to destroy viruses and other germs on floor and equipments. Pest control operators use pesticides to free houses from cockroaches, termites, ants, bed bugs and other house pests. Pesticides are also used for maintaining lawns, gardens and turfs, for freeing cattle and companion animals from ectoparasites viz. ticks, fleas, mites, etc.

Since most pesticides are designed to be toxic to their target pest, it can be harmful if used improperly. Its use therefore is strictly regulated. A rigorous regulatory evaluation process is applied for checking quality, toxicity, efficacy and residue levels of pesticides before it can be manufactured, sold or used in designated country/state.

Pesticides can be classified in a number of ways such as mode of entry, application, chemical class, mode of action and toxicity. Broad categories are defined in Fig. 17.1.

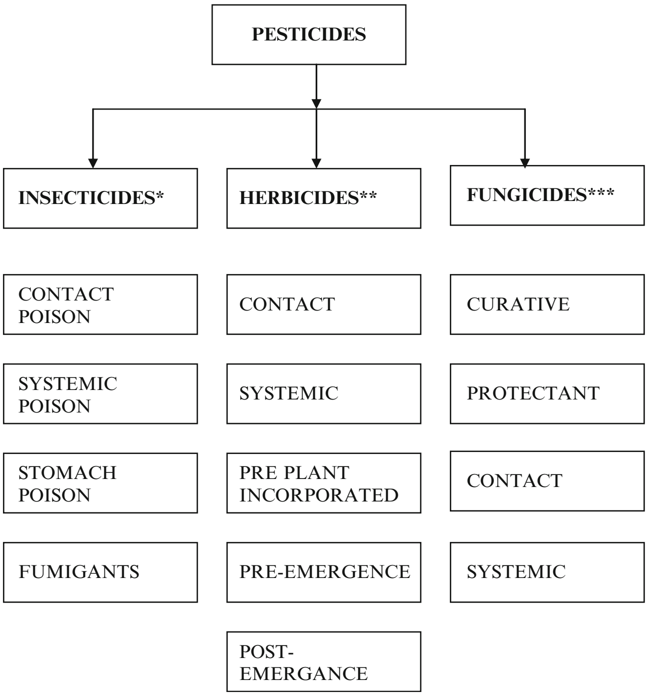


Fig. 17.1 Categories of pesticides.* Including nematicides, acaricides, molluscicides, rodenticides, repellants, attaractants, chemosterilant, antifeedents, growth regulators, etc. ** Including plant growth regulators, defoliants or desiccants, etc. ***Including bactericides, virucides, etc.

Insecticides

An insecticide is a chemical that kills the insects or inhibits their growth. Broad types of insect pests controlled by insecticides are:

- Insect pests of agricultural importance viz. field crops, vegetable crops, ornamental and horticultural crops including forestry, turf and lawns.
- Insects feed on stored products and cause economic loss.
- Household and disease carrying insects viz. cockroaches, furniture beetle, termite, ants, mosquitoes, houseflies, bedbugs, silverfish, fleas, bees, wasps, etc.
- Pest attacking cattle and companion animals such as houseflies, flesh fly, head lice, ticks, fleas and mites, etc.

Classification Based on Mode of Entry

Contact poison:
Insects get killed when they come in contact with the insecticides. It reaches into insect's body via contact with the cuticle and penetration through the sutures, membranes, tracheal system and bases of setae, etc.

This can be achieved by the direct application of the insecticide on the insects or by the insects crawling or passing over the treated surface having residue of the toxic material.

Systemic poison:
Systemic insecticides penetrate into the plant tissues through the roots, stem and leaves of the plant and translocate through the vascular tissues. Insects that feed on such treated plants are killed by stomach poison effect.
Systemic insecticides are applied in a variety of ways viz. seeds treatment, foliar treatment, broadcast application, in-furrow application to soil, etc.

Stomach poison:
When ingested by insects they act on their digestive system to kill them. All the systemic insecticides are stomach poison but all the stomach poisons are not the systemic insecticide.

Fumigants:
Fumigants are gaseous poison, which penetrate into the body through respiratory system and are used to kill insects, nematodes, rodents, etc.

Fumigants are employed to control stored products, household articles, soil inhabiting insects and nematodes, agricultural pests and live plants during export and import.

Classification Based on Chemical Class and Mode of Action

Following classification (Table 17.1) is based on IRAC (Insecticide Resistance Action Committee) "mode of action classification" version Sept 2010 [1]. It provides a guideline to crop protection professionals, growers and researchers for selection of an insecticide for use in sustainable insecticide resistance management program.

Herbicides

Herbicide is a chemical that kills the weeds or inhibits their growth. Weeds are the plants, which grow where they are not wanted. They grow in the fields where they compete with crops for water, soil nutrients, light and space and thus reduce crop yields. Annuals weeds live and produce their seeds in a single growing season. Biennials need two growing seasons; in one season they complete vegetative growth and second season they complete their reproductive stage. The multiplication of both the annuals and the biennials weeds is through seed. However, Perennials weeds live indefinitely and are propagated not only through seeds but also often vegetatively such as rhizomes, stolons, bulbs and tubers.

Weeds are divided into three sub classes: (1) grassy weeds, (2) broad-leaved weeds (3) sedges.

Herbicides have largely replaced mechanical methods of weed control in countries where intensive and highly mechanized agriculture is practiced. They provide a more effective and economical means of weed control than cultivation, hoeing and hand pulling.

Classification Based on Mode of Entry

Contact:
Contact herbicides kill only the plant parts on which the chemical is deposited and are most effective against annual weeds.

Table 17.1 Classification of insecticides based on chemical class and mode of action

Group	Mode of action	Subgroup and Chemical class	Active ingredients
1	Acetycholinesterase (AChE) inhibitors Nerve action	1A Carbamates	Aldicarb Aminocarb Bendiocarb Carbaryl Carbofuran Formetanate Methiocarb Pirimicarb Propoxur Thiodicarb Thiofanox Trizamate Xylylcarb
		IB Organophosphates	Acephate Azamethiphos Azinphos-methyl Chlorfenvinphos Chlorpyrifos Diazinon Dichlorvos (DDVP) Dimethoate Ethion Fenthion Heptenophos Malathion Methamidophos Omethoate Parathion Phenthoate Phorate Quinalphos Temephos Tebufos Tetrachlorvinphos Triazophos
2	*Gamma-amino butyric acid (GABA)*-gated chloride channel antagonists Nerve action	2A Cyclodienes Organochlorines	Chlordane Endosulfan
		2B Phenylpyrazoles (Fiproles)	Ethiprole Fipronil

(continued)

Table 17.1 (continued)

Group	Mode of action	Subgroup and Chemical class	Active ingredients
3	Sodium channel modulators		
	Nerve action	3A	Acrinathrin
		Pyrethroids	Allethrin
		Pyrethrins	D-*trans* allethrin
			Bifenthrin
			Bioresmethrin
			Cycloprothrin
			Cyfluthrin
			Cypermethrin
			Alphacypermethrin
			Betacypermethrin
			Thetacypermethrin
			Zetacypermethrin
			Deltamethrin
			Fenpropathrin
			Fenvalerate
			Flucythrinate
			Permethrin
			Pyrethrins
			Resmethrin
			Silafluofren
			Tefluthrin
			Tetramethrin
			Transfluthrin
		3B	DDT
		Organochlorines	Methoxychlor
4	Nicotinic acetylcholine receptor (nAChR) agonists		
	Nerve action	4A	Acetamiprid
		Neonicotinoids	Clothianidin
			Dinotefuran
			Imidacloprid
			Nitenpyran
			Thiacloprid
			Thiamethoxam
		4B	Nicotine
		Biopesticide	
5	Nicotinic acetylcholine receptor (nAChR) allosteric activators		
	Nerve action	Spinosyns	Spinosad
			Spinetoram
6	Chloride channel activators		
	Nerve and muscle action	Biopesticides	Abamectin
			Emamectin benzoate
			Lepimectin
			Milbemytin
7	Juvenile hormone mimics		
	Growth regulation	7A	Hydroprene
		Juvenile hormone analogues	Methoprene
			Pyripoxyfen
		7B	Fenoxycarb
		Carbamate	
8	Miscellaneous non-specific (multi-site) inhibitors		
		8A	Methyl bromide
		Alkyl halide	or other alkyl halides
		8B	Chloropicrin
		Unclassified	
		8C	Sulfuryl fluoride
		Inorganic compound	
		8D	Borax
		Inorganic compound	
		8E	Taratar emetic
		Taratar emetic	
9	Selective homopteran feeding blockers		
	Nerve action	Pyridine	Pymetrozine
			Flonicamid
10	Mite growth inhibitors		
	Growth regulation	10A	Clofentezine
		Tetrazine	Diflovidiazin
		10B	Hexythiazox
		Carboxamid	
		10C	Etoxaxazole
		Diphenyl oxazoline	

(continued)

Table 17.1 (continued)

Group	Mode of action	Subgroup and Chemical class	Active ingredients
11	Microbial disruptors of insect mid-gut membranes (includes transgenic crops expressing *Bacillus thuringiensis* toxins)	*Bacillus thuringiensis* or *Bacillus sphaericus* and the insecticidal proteins they produce	*Bacillus thuringiensis* Subsp. *israelensis* Subsp. *aizawai* Subsp. *tenebrionis* Subsp. *kurstaki* *Bacillus sphaericus* *Bt crop proteins: Cry1Ab, Cry1Ac, Cry1Fa, Cry2Ab, m Cry3A, Cry3Ab, Cry3Bb, Cry34/35Ab1*
12	Inhibitors of mitochondrial ATP synthatase Energy metabolism	12A Thourea	Diafenthiuron
		12B Organotin miticide	Azocyclotin Cyhexatin Fenbutatin oxide
		12C Sulfite ester	Propargite
		12D Bridged diphenyl	Tetradifon
13	Uncoupled of oxidative phosphorylation via disruption of the proton gradient Energy metabolism	13A Aryl pyrrole	Chlorfenapyr
		13B Dinitrophenol	DNOC
		13C Sulfonamide	Sulfluramid
14	Nicotinic acetylcholine receptor (nAChR) channel blockers Nerve action	Nereistoxin analogues	Bensultap Cartap hydrochloride Thiocyclam Thiosultap-sodium
15	Inhibitors of chitin biosynthesis type 0 Growth regulation	Benzoylurea	Bistrifluron Chlofluazuron Diflubenzuron Novaluron Noviflumuron Teflubenzuron Triflumuron
16	Inhibit chitin biosynthesis type 1 Growth regulation	Thiadiazin	Buprofezin
17	Moulting disruptor, Dipteran Growth regulation	Triazine	Cyromazine
18	Ecdysone receptor agonists Growth regulation	Diacylhydrazins	Hromafenazide Alofenozide Ethoxyfenozide Tebufenozide
19	Octopamine receptor agonist Nerve action	Triazapentadiene	Amitraz
20	Mitochondrial complex III electron transport inhibitors Energy metabolism	20A Trifluoromethyl aminohydrazone	Hydramethylnon
		20B Tetronic and Tetramic acid derivatives	Acequinocyl
		20C Strobilurin	Fluacrypyrim
21	Mitochondrial complex I electron transport inhibitors Energy metabolism	21A METI acaricides and insecticides	Fenazaquin Fenpyroximate Pyrimidifen Pyridaben Tebufenpyrad
		21B Botanical	Rotenone

(continued)

Table 17.1 (continued)

Group	Mode of action	Subgroup and Chemical class	Active ingredients
22	Voltage-dependent sodium channel blockers Nerve action	22A Oxadiazine	Indoxacarb
		22B Semicarbazone	Metaflumizone
23	Inhibitors of acetyl CoA carboxylase	Tetronic and tetramic acid derivatives	Spirodiclofen Spiromesifen Spirotetramat
24	Mitochondrial complex IV electron transport inhibitors Energy metabolism	24A Phoshphine	Aluminium phosphide Calcium phosphide Phoshphine Inc Phosphide
		24B Cyanide	Cyanide
25	Mitochondrial complex II electron transport inhibitors Energy metabolism	Pyrazole	Cynopyrafen
28	Ryanodine receptor modulators Nerve and muscle action	Diamides	Chlorantraniliprole Flubendiamide
29	Compounds of unknown or uncertain mode of action	Bridged diphenyl	Benzoximate Bromopropylate
		Hydrazine carboxylate	Bifenazate
		Aryloxyphenoxypropionic acid	Diclofol
		Unclassified	Pyridalyl

Systemic:
Systemic herbicides are absorbed either by roots or foliar parts of a plant and are then trans-located within the plant system to tissues that may be remote from the point of application. Systemic herbicides are effective against both annual and perennial weeds and they are particularly advantageous against established perennial weeds.

Classification Based on Application and Crop Selectivity

Pre-plant incorporated:
Treatment made to soil in advance before planting of seed. This is popularly known as "chemical follow treatment".

Pre-emergence:
Treatment made at planting or afterwards but before the weeds emerge.

Early post-emergence:
Treatment made when the weeds first emerge.

Post-emergence:
Treatment made after weeds emerge from the soil and during their active growth and development stage. It may be a contact or a systemic herbicide and may be soil applied or directed to the weeds.

Classification Based on Chemical Class and Mode of Action

Following classification (Table 17.2) is based on HRAC (Herbicide Resistance Action Committee) "mode of action classification" version January 2005 [2]. It is developed by HRAC in cooperation with WSSA (Weed Science Society of America) and provides a guideline to crop protection professionals, growers and researchers for selection of best suited herbicides to combat specific weeds in herbicide resistance management program.

Fungicides

A fungicide is a chemical, which inhibits or kills the pathogens. Most fungicides need to be applied before disease occurs (protectants) or at the appearance of symptoms (curative) to be effective.

Protectant fungicide is effective when applied prior to fungal infection. Prophylactic application is mandatory to prevent damage to growing crops whereas curative fungicide is capable of eradicating the fungus after it has caused infection.

Classification Based on Mode of Entry

Contact: A fungicide that remains on the surface where it is applied but does not go deeper into the plant system.

Table. 17.2 Classification of herbicides based on chemical class and mode of action

Group				
HRAC	WSSA	Mode of action	Chemical class	Active ingredients
A	1	Inhibitors of acetyl CoA carboxylase ACCase These chemicals block an enzyme called ACCase. This enzyme is helpful in formation of lipids in roots of grass plants. Without lipids susceptible weeds die	Aryloxyphenoxypropionate “FOPS”	Clodinafop-propargyl
				Cyhalofop-butyl
				Diclofop-methyl
				Fenoxaprop-p-ethyl
				Fluazifop-p-butyl
				Haloxyfop-R-methyl
				Propaquizafop
				Quizalofop-p-ethyl
			Cyclohexanediones “DIMS”	Alloxydim
				Butroxydim clethodim
				Cycloxydim
				Profoxydim
				Tepraloxydin tralkoxydim
			Phenylpyrazolin “DEN”	Pinoxeden
B	2	ALS/AHAS inhibitors These chemicals block the normal function of an enzyme called acetolactase (ALS) acetohydroxy acid (AHAS). This enzyme is essential in amino acid synthesis, without protein plants starve to death	Sulfonylureas	Amidosulfuron
				Azimsulfuron
				Bensulfuron-methyl
				Chlorimuron-ethyl
				Chlorsulfuron
				Cyclosulfamuron
				Ethoxysulfuron
				Foramsulfuron
				Imazosulfuron
				Metsulfuron-methyl
				Nicosulfuron
				Oxasulfuron
				Rimsulfuron
				Sulfosulfuron
				Tritosulfuron, etc.
			Imidazolinone	Imazapic
				Imazamethabenz-methyl
				Imazamox
				Imazapyr
				Imazaquin
				Imazethapyr
			Triazolopyrimidine	Cloransulam-methyl
				Diclosulam
				Florasulam
				Flumetsulam
				Metosulam
				Penoxsulam
			Pyrimidinyl(thio)benzoate	Bispyribac-Na
				Pyribenzoxim
				Pyriftalid
				pyrithiobac-Na
				Pyriminobac-methyl
			Sulfonylaminocarbonyl-triazolinone	Flucarbazone-Na
				Propoxycarbazone-Na
C1	5	Inhibition of photosynthesis at photosystem II. Site A These chemicals interfere with photosynthesis and disrupt plant growth, ultimately leading to death These chemicals interfere with photosynthesis and disrupt plant growth, ultimately leading to death	Triazine	Ametryne
				Atrazine
				Cyanazine
				Desmetryne
				Dimethametryne
				Prometon
				Prometryne
				Propazine
				Simazine
				Simetryne
				Terbumeton
				Terbuthylazine
				Terbutryne
				Trietazine
			Triazinone	Hexazinone
				Metamitron
				Metribuzin
			Triazolinone	Amicarbazone
			Uracil	Bromacil
				Lenacil
				Terbacil

(continued)

Table. 17.2 (continued)

Group				
HRAC	WSSA	Mode of action	Chemical class	Active ingredients
C1	5	Inhibition of photosynthesis at photosystem II. Site A	Pyridazinone	Pyrazon
				Chloridazon
			Phenyl-carbamate	Desmedipham
				Phenmedipham
C2	7	Inhibition of photosynthesis at photosystem II. Site B	Urea	Chlorobromuron
				Chlorotoluron
				Chloroxuron
				Dimefuron
				Diuron
				Ethidimuron
				Fenuron
				Fluometuron (see F3)
				Isoproturon
				Isouron
				Linuron
				Methabenzthiazuron
				Metobromuron
				Metoxuron
				Monolinuron
				Neburon
				Siduron
				Tebuthiuron
			Amide	Propanil
				Pentanochlor
C3	6	Inhibition of photosynthesis at photosystem II	Nitrile	Bromofenoxim
				Bromoxynil
				Ioxynil
			Benzothiadiazinone	Bentazon
			Phenyl-pyridazine	Pyridate
				Pyridafol
D	22	Photosystem-I-electron diversion	Bipyridylium	Diquat
				Paraquat
E	14	Inhibition of protoporphyrinogen oxidase (PPO)	Diphenylether	Acifluorfen-Na
				Bifenox
				Chlomethoxyfen
				Fluoroglycofen-ethyl
				Fomesafen
				Halosafen
				Lactofen
				Oxyfluorfen
			Phenylpyrazole	Fluazolate
				Pyraflufen-ethyl
			N-phenylphthalimide	Cinidon-ethyl
				Flumioxazin
				Flumiclorac-pentyl
			Thiadiazole	Fluthiacet-methyl
				Thidiazimin
			Oxadiazole	Oxadiazon
				Oxadiargyl
			Triazolinone	Azafenidin
				Carfentrazone-ethyl
				Sulfentrazone
			Oxazolidinedione	Pentoxazone
			Pyrimidindione	Benzfendizone
				Butafenacil
			Nitrile	Pyraclonil
			Pyridazinone	Flufenpyr-ethyl
F1	12	Bleaching: Inhibition of carotenoid biosynthesis at the phytoene desaturase step (PDS)	Pyridazinone	Norflurazon
				Flurtamone
			Pyridinecarboxamide	Diflufenican
				Picolinafen
			Amide	Beflubutamid
			Unclassified	Fluridone
				Flurochloridone

(continued)

Table. 17.2 (continued)

Group				
HRAC	WSSA	Mode of action	Chemical class	Active ingredients
F2	27	Bleaching: Inhibition of 4-hydroxyphenyl-pyruvate-dioxygenase (4-HPPD)	Triketone	Mesotrione
				Sulcotrione
			Isoxazole	Isoxachlortole
				Isoxaflutole
			Pyrazole	Benzofenap
				pyrazolynate
				Pyrazoxyfen
			Unclassified	Benzobicyclon
F3	11	Bleaching: Inhibition of carotenoid biosynthesis (unknown target)	Triazole	Amitrole (in vivo inhibition of lycopene cyclase)
	13	Inhibition of all diterpenes	Isoxazolidinone	Clomazone
F3	13	Inhibition of all diterpenes	Urea	Fluometuron
			Diphenylether	Aclonifen
G	9	Inhibition of EPSP synthesis	Glycine	Glyphosate
				Sulfosate
H	10	Inhibition of glutamine synthetase	Phosphinic acid	Glufosinate-ammonium
				Bialaphos = bilanaphos
I	18	Inhibition of DHP (dihydropteroate) synthase	Carbamate	Asulam
K1	3	Microtubule assembly inhibition	Dinitroaniline	Benefin = benfluralin
				Butralin
				Dinitramine
				Ethalfluralin
				Oryzalin
				Pendimethalin
		These chemicals inhibit the cell division in roots		Trifluralin
			Phosphoroamidate	Amiprophos-methyl
				Butamiphos
			Pyridine	Dithiopyr
				Thiazopyr
			Benzamide	Propyzamide = pronamide
				Tebutam
			Benzoic acid	DCPA = chlorthal-dimethyl
K2	23	Inhibition of mitosis/microtubule organization	Carbamate	Chlorpropham
				Propham
				Carbetamide
K3	15	Inhibition of VLCFAs (inhibition of cell growth and division)	Chloroacetamide	Acetochlor
				Alachlor
				Butachlor
				Dimethachlor
				Dimethanamid
				Metazachlor
				Metolachlor
				Pethoxamid
				Pretilachlor
				Propachlor
				Propisochlor
				Thenylchlor
			Acetamide	Diphenamid
				Napropamide
				Naproanilide
			Oxyacetamide	Flufenacet
				Mefenacet
			Tetrazolinone	Fentrazamide
			Organophosphate	Anilofos
				Piperophos
			Triazole	Cafenstrole
L	20	Inhibition of cell wall (cellulose) synthesis	Nitrile	Dichlobenil
				Chlorthiamid
			Benzamide	Isoxaben
			Triazolocarboxamide	Flupoxam
			Quinoline carboxylic acid	Quinclorac (for monocots) (also group O)
M	24	Uncoupling (membrane disruption)	Dinitrophenol	*DNOC*
				Dinoseb
				Dinoterb

(continued)

Table. 17.2 (continued)

Group				
HRAC	WSSA	Mode of action	Chemical class	Active ingredients
N	8	Inhibition of lipid synthesis—not ACCase inhibition These chemicals inhibit the cell division and elongation in the seedling shoots before they emerge above ground	Thiocarbamate	Butylate Cycloate Dimepiperate EPTC Esprocarb Molinate Orbencarb Pebulate Prosulfocarb Thiobencarb = benthiocarb Tiocarbazil Triallate Vernolate
			Phosphorodithioate	Bensulide
			Benzofuran	Benfuresate Ethofumesate
			Chloro-carbonic-acid	TCA Dalapon Flupropanate
O	4	Synthetic auxins These chemicals disrupt plant cell growth in the newly forming stems and leaves; they affect protein synthesis and normal cell division leading to malformed growth and tumours	Phenoxy-carboxylic-acid	Clomeprop 2,4-D 2,4-DB Dichlorprop = 2,4-DP MCPA MCPB Mecoprop = MCPP = CMPP
			Benzoic acid	Chloramben Dicamba TBA
			Pyridine carboxylic acid	Clopyralid Fluroxypyr Picloram Triclopyr
			Quinoline carboxylic acid	Quinclorac (also group L) Quinmerac
			Benzothiazolone	Benazolin-ethyl
P	19	Inhibition of auxin transport	Phthalamate	Naptalam
			Semicarbazone	Diflufenzopyr-Na
Z	17	Unknown	Organoarsenical	DSMA MSMA
			Amide	Bromobutide
			Morphactin	Chloroflurenol
			Monoterpene	Cinmethylin
			Urea	Cumyluron
			Dithiocarbamate	Dazomet Metam
			Organophosphate	Fosamin
			Unclassified	Oxaziclomefone
			Biopesticide	Pelargonic acid
	25		Arylaminopropionic acid	Flamprop-M-methyl/-isopropyl
	26		Pyrazolium	Difenzoquat

Note: While the mode of action of herbicides in Group Z is unknown it is likely that they differ in mode of action between themselves and from other groups

Fungicides have no after-infection activity. Repeated applications are needed to protect new growth of the plant and to replace material that has been washed off by rain or irrigation, or degraded by environmental factors.

Systemic: A fungicide that is absorbed into plant tissue and may offer some after-infection activity. Very few fungicides are truly systemic (i.e. move freely throughout the plant); however, some are upwardly systemic (i.e. move only upward in the plant through xylem tissue) and some are locally systemic (i.e. move into treated leaves and redistribute to some degree within the treated portion of the plant).

Classification Based on Chemical Class and Mode of Action

Following classification (Table 17.3) is based on FRAC (Fungicide Resistance Action Committee) "mode of action

Table. 17.3 Classification of fungicides based on chemical class and mode of action

FRAC Code	Mode of action	Chemical class	Active ingredients
1	Inhibition of B-tubuline assembly in mitosis and cell division	Benzimidazoles	Benomyl
			Carbendazim
			Fuberidazole
			Thiabendazole
		Thiophanates	Thiophanate
			Thiophanate-methyl
2	Inhibition of mitogen-activated-protein (MAP)/Histidine-kinase in osmotic signal transduction	Dicarboxamides	Chlozolinate
			Iprodione
			Procymidone
3	Inhibition of sterol biosynthesis (C14-demethylase) in membrane	Piperazines	Triforine
		Pyridines	Pyrifenox
		Pyrimidines	Fenarimol
			Nuarimol
		Imidazoles	Imazalil
			Oxpoconazole
			Pefurazoate
			Triflumizole
		Triazoles	Azaconazole
			Bitetranol
			Bromuconazole
			Cyproconazole
			Difenconazole
			Diniconazole
			Epoxiconazole
			Etaconazole
			Fenbuconazole
			Flusilazole
			Flutriafol
			Hexaconazole
			Ipoconazole
			Metconazole
			Penconazole
			Propiconazole
			Simeconazole
			Tridimenol
4	Inhibition of RNA polymerase1 in nucleic acid synthesis	Acylalanines (Phenyl amides)	Benalaxyl
			Furalaxyl
			Metalaxyl
		Oxazolidinones (Phenyl amides)	Oxadixyl
		Butyrolactones (Phenyl amides)	Ofurace
5	Inhibition of isomerase in sterol biosynthesis	Morpholines	Aldimorph
			Dodemorph
			Tridemorph
		Piperidines	Fenpropidin
			Piperalin
		Spiroketal-amines	Spiroxamine
6	Inhibition of phospholipid biosynthesis methyltransferase in lipid membrane synthesis	Phosphorothioates	Edifenphos
			Iprobenfos
			Pyrazophos
		Dithiolanes	Isoprothiolane
7	Inhibition of complex II, succinate dehydrogenase (SDHI) in respiratory chain of fungal mitochondria	Phenyl benzamides	Benodanil
			Flutolanil
			Mepronil
		Pyridinyl-ethyl-benzamides	Fluopyram
		Furan-carboxamide	Fenfuram
		Oxathiin-carboxamide	Carboxin
			Oxycarboxin
		Thiazole-arboxamide	Thifluzamide
		Pyrazole-carboxamide	Bixafen
			Furametpyr
			Isopyrazam
			Penflufen
		Pyridine-carboxamide	Boscalid
8	Inhibition of adenosine-deaminase in nucleic acid synthesis	Hydroxy-(2-amino-)pyrimidines	Bupirimate
			Dimethirimol
			Ethirimol
9	Inhibition of amino acid (methionine) biosynthesis	Anilino pyrimidines	Cyprodinyl
			Mepanipyrim
			Pyrimethanil

(continued)

Table. 17.3 (continued)

FRAC Code	Mode of action	Chemical class	Active ingredients
10	Inhibition of B-tubuline assembly in mitosis and cell division	N-phenyl carbamates	Diethofencarb
11	Inhibition of complex III Cytochrome bc1 (ubiquinol oxidase) at Qo site in respiratory chain of fungal mitochondria	Methoxy-acrylates Methoxy-carbamates Oximino-acetates Oximino-acetamides Oxazolidine-diones Dihydro-dioxazines Imidazolinones Benzyl-carbamates	Enestrobin Picoxystrobin Pyraclostrobin Pyrametastrobin Dimoxystrobin Metoministrobin Famoxadone Fluoxastrobin Fenamidone Pyribencarb
12	Inhibition of mitogen-activated-protein (MAP)/Histidine-kinase in osmotic signal transduction	Phenylpyrroles	Fenpiclonil Fludioxonil
13	Inhibition of signal transduction by unknown mechanism	Quinolines Quinazolinones	Quinoxyfen Proquinazid
14	Inhibition of lipid peroxidation synthesis in lipid membrane synthesis	Aromatic hydrocarbons 1,2,4-Thiadiazoles	Biphenyl Chloroneb Dicloran Quintozin Etridiazole
16.1	Inhibition of reductase in melanin biosynthesis in cell wall	Isobenzo-furanone Pyrrolo-quinolinone Triazolobenzo-thiazole	Pthalide Pyroquilon Tricyclazole
16.2	Inhibition of dehydratase in melanin biosynthesis in cell wall	Cyclopropane-carboxamide Carboxamide Propionamide	Carpropamid Diclocymet Fenoxanil
17	Inhibition of sterol biosynthesis (3-keto reductase, C4-de-methylation) in membrane	Hydroxyanilides	Fenhexamid
18	Squalence-epoxidase inhibition in sterol biosynthesis in membrane	Thiocarbamates Allylamines	Pyributicarb Naftifine Terbinafine
19	Inhibition of Chitin synthase in glucan synthesis	Peptidyl pyrimidine nucleoside	Polyoxin
20	Mitosis and Cell division inhibition (proposed)	Phenyl ureas	Pencycuron
21	Inhibition of complex III Cytochrome bc1(ubiquinone reductase) at Qi site in respiratory chain of fungal mitochondria	Cynoimidazole Sulfamoyltriazole	Cyazofamid Amisulbrom
22	Inhibition of B-tubuline assembly in mitosis and cell division	Toluamides (Benzamides)	Zoxamide
23	Amino acid and protein synthesis inhibition	Encopyranuronic acid	Blasticidin-S
24	Amino acid and protein synthesis inhibition	Hexopyranosyl antibiotic	Kasugamycin
25	Amino acid and protein synthesis inhibition	Glucopyranosyl antibiotic	Streptomycin
26	Inhibition of trehalase and inositol-biosynthesis in glucan synthesis	Glucopyranosyl antibiotic	Validamycin
27	Unknown	Cyanoacetamideoxime	Cymoxanil
28	Inhibition of cell membrane permeability in lipid and membrane synthesis	Carbamates	Iodocarb Propamocarb Prothiocarb
29	Respiration inhibition (uncouplers of oxidative phosphorylation)	Dinitrophenyl crotonates 2,6-Dinitroanilines Pyrimidinone hydrazones	Binapacryl Dinocap Fluazinam Ferimzone
30	Inhibitors of oxidative phosphorylation, ATP synthase in respiration	Tri phenyl tin compounds	Fentin acetate Fentin chloride Fentin hydroxide
31	Nucleic acid synthesis inhibition in DNA topoisomerase type II (gyrase)	Carboxylic acids	Oxolinic acid
32	Nucleic acid synthesis inhibition in DNA/RNA synthesis (proposed)	Isoxazoles Isothiazolones	Hymexazole Octhilinone
33	Unknown	Phosphonates	Fosetyl-Al Phosphorus acid and salts

(continued)

Table. 17.3 (continued)

FRAC Code	Mode of action	Chemical class	Active ingredients
34	Unknown	Phthalamic acid	Teclofthalam
35	Unknown	Benzotriazines	Triazoxide
36	Unknown	Benzene sulfonamides	Flusulfamide
37	Unknown	Pyridazinones	Diclomezine
38	Inhibition of ATP production in respiration (proposed)	Thiophene carboxamides	Silthiofems
39	Inhibition of Complex1 NADH oxido-reductase in respiration	Pyrimidinamines	Diflumetorim
40	Inhibition of phospholipid biosynthesis and cell wall deposition in lipids and membrane synthesis	Cinnamic acid amides	Dimethomorph Flumorph
		Valinamide carbamates	Benthiavalicarb Iprovalicarb Valifenlate
		Mandelic acid amides	Mandipropamid
41	Protein and amino acid synthesis inhibition	Tetracycline antibiotic	Oxytetracycline
42	Unknown	Thiocarbamate	Methasulfocarb
43	Delocalization of spectrin-like proteins in mitosis and cell division	Pyridinylmethyl-benzamides (Benzamides)	Fluopicolide
44	Microbial disrupters of pathogen cell membranes in lipids and membrane synthesis	*Bacillus subtilis* and the fungicidal lipopeptides they produce	*Bacillus subtilis* strain QST 713
45	Inhibition of Complex III Cytochrome bc1 (ubiquinone reductase) at Qu (unknown) site in respiratory chain of fungal mitochondria	Triazolo-pyrimidylamine	Ametoctradin
P	Activators of host plant defense system	Benzothiadiazole	Acibenzolar-s-methyl
		Benzisothiazole	Probenazole
		Thiadiazole-carboxamide	Tiadinil Isotianil
		Natural compounds	Laminarin
M	Inactivates fungal pathogens by multi site activity	Inorganic	Copper (different salts)
		Inorganic	Sulphur
		Dithiocarbamates and relatives	Ferbam Mancozeb Maneb Metiram Thiaram Zineb Ziram
		Phthalimides	Captan Captafol Folpet
		Chloronitriles	Chlorothalonil
		Sulfamides	Dichlofluanid Tolyfluanid
		Guanidines	Guazatine Iminoctadin
		Triazines	Anilazin
		Quinones	Dithianon
NC	Unknown	Diverse	Mineral oils Organic oils Potassium bicarbonate Materials of biological origin
U5	Microtubule disruption in mitosis and cell division (proposed)	Thiazole carboxamide	Ethaboxam
U6	Unknown	Phenyl acetamide	Cyflufenmid
U8	Actin disruption (proposed)	Benzophenone	Metraphenone
U12	Cell membrane disruption (proposed)	Guanidines	Dodine

Table. 17.4 WHO toxicity classification of pesticides

WHO Class	Description	LD_{50} for the rat (mg/kg body weight) Oral	 Dermal
Ia	Extremely hazardous	<5	<50
Ib	Highly hazardous	5–50	50–200
II	Moderately hazardous	50–2,000	200–2,000
III	Slightly hazardous	$>2,000$	$>2,000$
IV	Unlikely to present acute hazard	$\geq 5,000$	$\geq 5,000$

Table 17.5 Classification and labelling system

Globally harmonized system category	Classification criteria Oral LD_{50} (mg/kg bw)	 Hazard statement	 Dermal LD_{50} (mg/kg bw)	 Hazard statement
Category 1	<5	Fatal if swallowed	<50	Fatal in contact with skin
Category 2	5–50	Fatal if swallowed	50–200	Fatal in contact with skin
Category 3	50–300	Toxic if swallowed	200–1,000	Toxic in contact with skin
Category 4	300–2,000	Harmful if swallowed	1,000–2,000	Harmful in contact with skin
Category 5	2,000–5,000	May be harmful if swallowed	2,000–5,000	May be harmful in contact with skin

Table 17.6 Classification of pesticides based on toxicity

Toxicity and signal word	Category I	Category II	Category III	Category IV
Acute oral	Up to and including 50 mg/kg	>50–500 mg/kg	>500–5,000 mg/kg	$>5,000$ mg/kg
Acute dermal	Up to and including 200 mg/kg	>200–2,000 mg/kg	$>2,000$–5,000 mg/kg	$>5,000$ mg/kg
Acute inhalation	Up to and including 0.05 mg/L	>0.05–0.5 mg/L	>0.5–2 mg/L	>2 mg/L
Primary eye irritation	Corrosive (irreversible destruction of ocular tissue) or corneal involvement or irritation persisting for more than 21 days	Corneal involvement or other eye irritation clearing in 8–21 days	Corneal involvement or other eye irritation clearing in 7 days or less	Minimal effects clearing in less than 24 h
Primary skin irritation	Corrosive (tissue destruction into the dermis and/or scarring)	Severe irritation at 72 h (severe erythema or edema)	Moderate irritation at 72 h (moderate erythema)	Mild or slight irritation at 72 h (no irritation or slight erythema)
Signal word	Danger	Warning	Caution	Not required

classification” version December 2009 [3] with an objective to help development of a strategy for the management of fungicide resistance risk.

Pesticide Classification Based on Toxicity

Assignment to a Toxicity Class is based typically on results of acute toxicity studies such as the determination of LD_{50} values in animal experiments, notably rodents. The Toxicity Class generally does not address issues of other potential harm of the agent, such as bioaccumulation issues of carcinogenicity, mutagenicity, etc.

World Health Organization Toxicity Classification of Pesticides [4]

See Table 17.4.

Globally Harmonized System for Classification and Labelling [4]

See Table 17.5.

US EPA Toxicity-based Classification of Pesticides [5]

See Table 17.6.

Chemistry of Agrochemicals

Rationale of Design

Launching a successful pesticide in the market requires a great deal of money, effort and time. First step is to develop a chemical compound with defined biological activity. It is tested in glass house/screen house conditions for favourable bio effectiveness. Its biochemical mode of action is established or predicted. A large number of molecules have to undergo initial screening against variety of insect pests, diseases and weeds to discover selected few candidates for further evaluation. This lead discovery can be done by "random screening" on the basis of historic compound collections, or a rational approach can be followed to develop structure-based design. It is a process in which structural information about binding of a molecule (usually an inhibitor) to an enzyme (very often at or near the active site) is used to design other molecules with improved binding properties in an iterative process, i.e. modifying structure gradually and repeating the evaluation process till desired result is obtained. Besides this, 3D-QSAR (quantitative structure activity relationship) studies or biochemically inspired design approach can also be employed for rationale designing of new molecules [6].

Importance of Physico-chemical Properties

Most biologically influential properties are Solubility and Reactivity. Hence, it is essential to take into consideration a suitable physicochemical profile to ensure uptake and translocation of pesticide.

Solubility is important for developing formulation in appropriate form and for bio-disposition of active ingredient in the living system after application. This includes absorption, distribution/translocation and metabolism.

The solubility is expressed in terms of its affinity/philicity or repulsion/phobicity for either an aqueous (hydro) or lipid (lipo) solvent.

Majority of active ingredients have some degree of solubility in both aqueous and lipid media as there is a need for them to move through both aqueous (plasma, extracellular fluid, cytoplasm, etc.) and lipid media (biological membranes) in the biological systems.

It is possible to estimate the solubility properties, i.e. (hydrophilic vs. lipophilic) by examining the structure and noting whether its structural features promote affinity for aqueous or lipid media.

The relative solubility of a pesticidal agent can be determined in the laboratory as partition coefficient (P; the ratio of the solubility of the compound in an organic solvent, usually *n*-octanol, to the solubility of the same compound in aqueous environment and is often expressed as a log value).

The relative solubility of a pesticidal agent is the sum of the contributions of each group and substituent to overall solubility. The presence of oxygen and nitrogen containing functional groups usually enhances water solubility while nonionizable hydrocarbon chains and ring systems enhance lipid solubility.

If the activity of pesticide is dependent on physical properties like solubility, partition coefficient and vapour pressure and not on the presence or absence of some chemical groups it is termed as structurally non-specific activity on the other hand if the activity is dependent on factors such as presence or absence of certain functional groups, intramolecular distance and shape of molecules and not easily co-related with any physical property it is termed as structurally specific activity.

Structural Features

Structurally specific activity is dependent on the interaction of the molecule with a cellular receptor and a small change in structure often leads to change in activity.

Receptor could be cellular constituent or it could be protein or enzyme. Receptors have a specific chemical affinity and structural requirements for a pesticidal agent for necessary interaction. The effect could be agonist in nature, i.e. generated after combining with the receptor or antagonistic, i.e. produced after blocking the receptor.

Stereochemistry or space arrangement of the atoms/3D structure of the molecule is a very important consideration while designing a pesticidal agent.

Any change in stereo specificity results in change of activity to certain degree. Since isomeric pairs have different physical properties, they differ in biological activity.

The following pesticide activity cases are possible for stereo isomeric pesticides:

1. Both of the stereoisomers exhibit comparable pesticide activity.
2. Only one of the stereoisomers exhibits pesticide activity. The other stereoisomer is inactive.
3. Both stereoisomers exhibit pesticide activity, but one is more active than the other.

Following classes stereoisomers are most important for bioactive compounds.

Geometric and Optical Isomers

Geometric isomers are referred as *cis* and *trans* isomers. When a molecule has a carbon-to-carbon double bond (−C=C–), a substituent group (or atom) can be positioned at the same side (*cis* isomers) of a double bond or at opposite side (*trans* isomers). These are not mirror images and have different physico-chemical properties and sometimes altogether different activity profiles.

They generally do not fit the same receptor equally well and if these functional groups are responsible for molecules biological activity a very different activity profile is resulted.

Optical isomerism originates from the presence of an atom containing four different substituents (i.e. the presence of an asymmetric atom, mostly referred to as a "chiral atom)". In most molecules, chirality is associated with a tetrahedral carbon atom. However, it is not restricted to tetrahedral carbon. Examples of chirality at the phosphate-P atom are found among some organophosphate pesticides.

Optical isomers or enantiomers have exactly the same physical and thermodynamic properties and differ only in their ability to rotate plane of polarized light.

Chirality plays a very important role because all organisms constitute a chiral environment and most enzymatic pathways are stereoselective.

Chirality is a very common phenomenon. As many as 25% of all pesticide active ingredients are chiral. Among insecticides all pyrethroids (having more than one chiral centres) and 30% of organophosphorus insecticides are chiral. Among herbicides chirality is observed in phenoxies, oxyphenoxypropionates, imidazolinones and cyclohexanediones. Fungicidal classes which exhibit chirality are triazoles and acylalanines.

Enatiomers differ in their toxicity, both to target and non-target species.

Example: two S-enantiomers of metolachlor are ten times more toxic to target weeds than the two R-isomers [7]. All fungicidal activity of metalaxyl and herbicide activity of Clodinafop reside with the R-enantiomer.

In spite of the fact that main biological activities or toxicities are associated with only one or a few enantiomers, most of these pesticides have been produced and used as racemates, which cause excessive contamination and pollution in the environment.

When agrochemicals have chiral structure efforts should be made to define the mode of action, elucidate metabolic pathways and define the human and environmental toxicity of each enantiopure isomer. If there are large differences in the biological activities of individual enantiomers in racemates, it is desirable to develop and use only the enantiopure isomer with the highest sought-after biological activities.

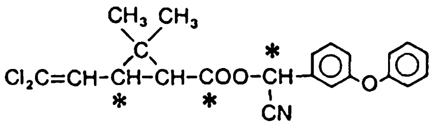
The separation of an enantiomeric pair or chiral synthesis is often a technically difficult and/or laborious and expensive process. However, techniques for asymmetric syntheses and separations of enantiomers with chiral HPLC and GC columns have been developed for several agrochemicals.

For industrial scale preparation of optically pure compounds naturally occurring, chiral compounds such as amino acids, carbohydrates and terpinoids constitute the "chiral pool" that can be used for transferring chirality to reagents. The chiral reagents are then used to synthesize the desired enantiomer. This is in contrast with "conventional" syntheses that result in racemic mixtures and not in a specific enantiomer. An example of a pesticide that uses a terpinoid chiral pool is *S*-Metolachlor, which results in an 80% enrichment of the most herbicide active *S*-enantiomer. Use of chiral solvents is sometimes used as the reaction media for non-chiral synthesis [8].

Production of single or enriched enantiomer pesticide is a green chemistry development, because it reduces the environmental loading of the inactive enantiomer.

The pyrethroid insecticides are typically esters of chrysanthemic acid and provide an excellent example of optical and geometric isomerism. Pyrethroid toxicity is highly dependent on the 3D of the molecule. Acute toxicity of a mixture of 2 isomers depends on the ratio of the amounts of the two isomers in the formulation. For example, the female rat's acute oral LD50 of permethrin increases from 224 to 6,000 mg/kg as the proportion of *trans* isomer increases from 20 to 80%.

Cypermethrin molecule has three chiral centres



This gives rise to eight stereoisomers (Fig. 17.2), four *cis* and four *trans* isomers, the *cis* isomers being the more biologically active.

Of the eight possible isomers, 1R *Cis* S isomer is the most biologically active isomer.

Racemic mixture of 1R *Cis* S and 1S *Cis* R is Alphacypermethrin. Alphacypermethrin is two to three times more effective than Cypermethrin.

Bioisosterism

Bioisosterism is the procedure of the synthesis of structural analogues of a lead compound by substitution of an atom or a group of atoms in the parent compound for another similar electronic and steric characteristics.

Any compound derived from another by such replacement is called bioisostere.

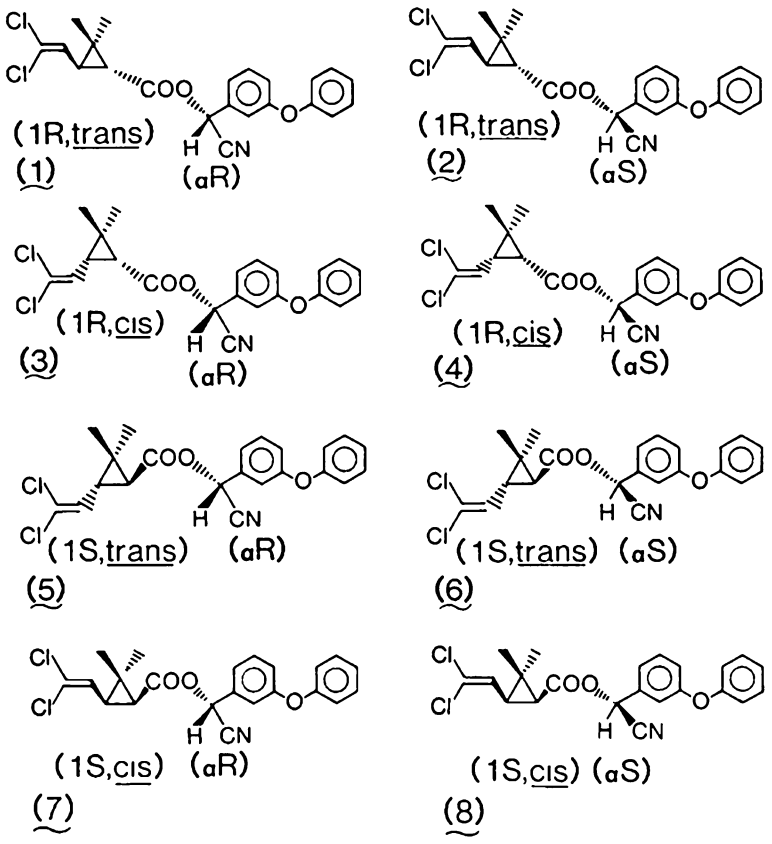
Bioisosterism is a useful strategy for molecular modification to design new agrochemical. Since molecules maintain similar biological properties this approach is used to modify molecules for acceptable efficacy, safety, stability, etc.

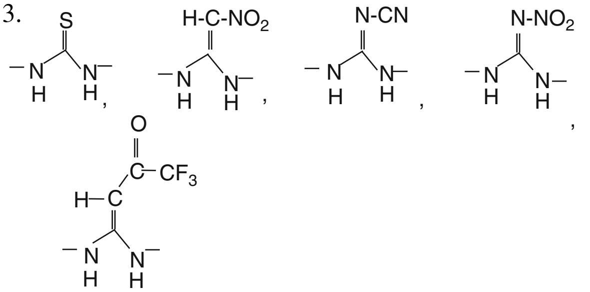
There are many examples of bioisosteres, which may be equivalent in some instances but not in others depending on what factors are important in binding (electronegativity, size, polarity, etc.).

Examples of bioisosterism [9]:

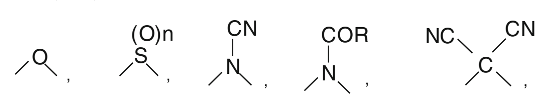
1. $-CO_2H$,$-SO_2NHR$,$-SO_3H$,$-PO(OH)_2$,$-PO(OH)NH_2$,– CO NHCN
2. –F,–Cl,–Br,$-CF_3$,–CN,–SCN,$-N(CN)_2$,$-C(CN)_3$

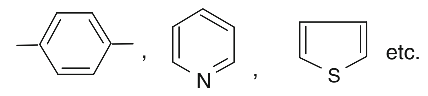
Fig. 17.2 Stereoisomers of cypermethrin

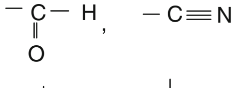


3.


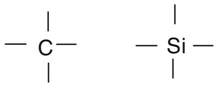
4. –OH,–NHCOR,–NHSO$_2$R,–CH$_2$OH,–NHCONH$_2$,–NHCN, –CH(CN)$_2$

5.


6. — CH=CH)n,


7.


$-\underset{\underset{O}{\|}}{C}-H, \quad -C\equiv N$

8.


$-\overset{|}{\underset{|}{C}}- \quad -\overset{|}{\underset{|}{Si}}-$

Synthesis of Agrochemicals

Synthesis of agrochemicals is typically a multi-step process and different routes can be followed to reach at the target active ingredient. Several factors are considered before deciding and standardizing feasible route for synthesis. Main factors that influence the choice are availability and cost of raw materials and reagents, engineering capabilities for performing desired reactions, amount and profile of impurities obtained by reaction sequence, yield and purity of intermediates and final product, cost benefit ratio, effluent generated during process, status of process patent, etc.

There is a lot of emphasis to shift towards reagents and chemicals which are less polluting for environment for instance use of Diphosgene/Triphosgene/Diarylcarbamates in lieu of phosphene in the manufacture of carbamates and sulfonyl ureas.

Backward synthesis or retro synthesis approach is often followed to work out synthetic route both in the synthesis of specific building blocks, key intermediates, penultimate intermediates or target active ingredient whole range of chemical reactions are deployed, e.g. aromatic electrophilic and nucleophilic substitution, chlorination, diazotization, sandmeyer reaction, diels alder cycloadditions, telemerizations, “N”, “O” alkylations, etc.

Pesticide Formulation (Commonly Known as Products or Plant Protection Products)

The active ingredient in a pesticide formulation controls the pest. Active ingredient is seldom used as such. It needs to be mixed with other inert materials viz. solvents, adjuvants, safeners, synergists, bulk carriers, etc. during the formulation process. Inerts may simply be added as bulk carriers to aid in handling the product or added to increase the stability, performance and safety of the product. The combination of an active ingredient with inert material is referred to as a formulation.

Sometimes two or more pesticides are combined as a premix referred as “Combination formulation” for:

- Providing better performance than either product would alone.
- Increasing the length and spectrum of control of pests.
- Saving money by reducing labour and wear on equipment (due to fewer applications).

When two or more pesticides provide a greater response than the added effects of each material applied separately, the response is considered synergistic. People often mistake additive effects for synergism. However, unlike with additive effects, the chemicals in a synergistic combination are not neutral toward each other. Rather, they interact in some way that increases efficacy beyond simple addition. Thus, not only does synergism save time, labour and wear on equipment, it enhances control compared to applying each of the materials separately.

Different formulations may be used differently. Some are to be used direct from the package, while others need to be diluted typically with water or oil. Some are recommended with the addition of certain adjuvants in spray tank to enhance the biological function.

Depending on intended use a single active ingredient can be formulated in various forms and concentrations in solid, liquid or gaseous forms and can be grouped into following categories:

Dry or solid formulations

(a) Ready to use : Dust (DP), baits (BB), granules (GR)
(b) Concentrates which have to be mixed with water before application: Tablets (TB) wettable powder (WP), water dispersable granule (WG) or Dry flowables (DF), soluble powder (SP)

Liquid formulations

(a) Emulsifiable cocentrates (EC)
(b) Ready-to use low concentrate solutions (RTU)
(c) Ultra low volume (ULV)
(d) Flowables/Suspension concentrates (SC)
(e) Soluble concentrates (SL)
(f) Aerosols(AE)

Other formulations: which can not be classified clearly as solid or liquid include.

(a) Microencapsulated materials
(b) Water soluble packets
(c) Insect attractants/repellents and pheromones
(d) Pesticide/fertilizer combinations
(e) Fumigants

Nanoformulation

The current trend is towards development of Nanoformulation or Nanotechnology. Nanoencapsulated formulations enable smaller quantities of the pesticides to be used effectively over a given period of time. Their design enables them to resist the severe environmental processes that act to eliminate conventionally applied pesticides, i.e. leaching, evaporation and photolytic, hydrolytic and microbial degradation. When the nanoformulations are applied as foliar spray or in the soil, the carrier and pesticide interact with the soil, insect, plant and atmosphere. Currently there is little knowledge how these nanoencapsulated pesticides are degraded in the plant soil and environment.

While, the benefits and prospects of the nanocapsulated pesticides are huge, but addressing some of the following issues are also equally important while designing the nanocarriers. They are

- Cost of the material and processing of the nanoencapsulated formulations should be low compared to the existing formulation.
- Fate of the nanocarriers in the environment, plant and soil.
- Fate of nanocarrier additives like fillers, stabilizers, antioxidants, etc.
- Environmental impact of the degradation of the carriers and additives in response to heat, hydrolysis, oxidation, solar radiation and biological agents.

Role of Adjuvants in Formulations

“A substance other than water, without pesticidal properties, which enhances stability or is intended to enhance the effectiveness of a pesticide when added either in the pesticide formulation or in the spray tank mixture”.

Adjuvants are helping to revolutionise the pesticide business as an important tool for users to improve application and achieve more cost-effective, better targeted and more environmentally acceptable pest control.

Though there is no “magic bullet” or single adjuvant product that can perform all these functions, technology

Table 17.7 Classification of adjuvants based on chemistry

Classes	Sub-classes
Emulsifiable oils	Mineral oil Seed oil (triglycerides) Alkoxylated triglycerides
Fatty acids from vegetable origin	Esterified Alkoxylated and esterified
Surfactants	Based on: alcohols, fatty acids, alkylphenols, alkylamines, sorbitol, sorbitan, trisiloxane, and alkylpolyglycosides
Salts	Ammonium sulphate, ureum, ammonium nitrate, calcium nitrate
Monoterpenic alcohols	
Waxes	Acid wax, ester wax, ethoxylated acid wax
Phospholipids	soya phospholipids
Polymers	Polyacrylamides, polyethylene oxide, polyvinyl

has advanced to the point where many of these functions can be incorporated into fewer high value products.

The physical–chemical properties of an adjuvant determine its usefulness in a certain application. Different chemical structures may have similar physical-chemical properties required for a certain adjuvant function. That's why so many different structures are sometimes recommended for the same function.

On the basis of chemistry adjuvants can be classified as in Table 17.7.

Characteristics of Main Formulation Types

See Table 17.8.

Formulation Types and Their Codes [10]

The following types of formulations are listed as per the "Catalogue of Pesticide Formulation types and International Coding Systems", GCPF (GIFAP) Technical Monograph No 2, 6th Edition, 2008. (See Table 17.9).

Adjuvant Types with Terms and Definitions

The following terminology for adjuvants has been approved by ASTM (American Society for Testing and material), and approved by the Adjuvants and Inerts committee (AIC) of the Chemical Producers and distributors Association (CDPA) [11]. (See Table 17.10).

Challenges and Way Ahead for Pesticide Industry

Resistance to Pesticides

Pesticide resistance is the adaptation of pest population targeted by a pesticide resulting in decreased susceptibility to that chemical. By the nature of co-evolution process, resistance most often develops to pesticides that are initially very effective and frequently used for control of pests. Pests develop a resistance to a chemical through genetic mutation and selection.

The United Nations Environmental Program listed pest resistance to pesticides as one of the most serious threat to global agriculture. World wide there are >550 insects and mite species, >150 plant pathogens species, and >270 weed biotypes that have been reported to have developed resistance to pesticides and more than 10% crop is lost annually due to pest resistance.

Rotation of different classes of pesticides and in some cases pre-mixes/tank mixes of two or more pesticides are useful strategies since an individual pest is less likely to develop resistance to two or more differing classes of pesticides. Such strategies are incorporated in Integrated Pest Management (IPM) program, which is being promoted globally by pesticide companies as an effective tool to combat pesticide resistance.

IPM programs use current and comprehensive information on the life cycles of pests and their interaction with the environment. This information, in combination with available pest control options, is used to minimize pest damage by the most economical means, and with the least possible hazard to people, property, and the environment.

Common strategy could be as follows to delay the resistance to pesticides:

1. Set action thresholds

   The emphasis is on control by judicious use of pesticides, not eradication of pests. By not killing 100% pests there is some susceptible population left that dilute resistant genes in the given population. Thus, before taking any pest control action, first set an action threshold, a point at which pest control action must be taken to stop the population to reach economic injury level where control of pests become difficult and un-economical.
2. Monitor and identify pests

   IPM programs work to monitor and identify the pests accurately, so that appropriate control decisions can be made in conjunction with action thresholds. The monitoring and identification removes the possibility of overuse, misuse and unnecessary uses of pesticides.

Table 17.8 Characteristics of main formulation types

Type	Main ingredients	Advantages	Disadvantages
Granule (GR)	Active ingredient Binder Stabiliser Carrier	Easy to handle and pack Low drift Low phytotoxicity Long residual activity	May be consumed by non-target organisms especially birds Need specialist application equipment like horn seeder
Tablets (TB)	Active ingredient Binder Lubricant Disintegrant	Easy to use Minimal packaging Less risk of incorrect dosage	Only suitable for highly active AIS High value low volume application
Wettable powder (WP)	Active ingredient Absorbent/carrier Wetting agent Antifoaming agent Dispersant	Easy to produce, pack and handle Tolerant to low temperature Low phytotoxicity No solvents	Produce dust Difficult to measure and mix Poor wettability and suspensibility May block lines and nozzles Poor stability Sticks inside packaging material
Water dispersible granule (WG)	Active ingredient Filler/carrier Disintegrant Wetting agent Dispersant	Low dusting Easy to pack, dispose handle and measure No solvent Tolerant to freezing	Long development effort Need expensive production equipment Dispersion may be affected by low temperature
Emulsifiable concentrates (EC)	Active ingredient Stabiliser Solvent Emulsifier	Easy to produce, handle and mix Useful for water-insoluble low melting AIS	Expensive to pack and transport Crystallization at low temperature Chances of phytotoxicity May be corrosive to metal and plastic Increased dermal toxicity May be sensitive to water hardness
Soluble concentrates (SL)	Active ingredient Antifreeze Wetter Water miscible solvent	Easy to produce Low phytotoxicity No flammability hazard	Expensive to pack and transport Frost sensitive May corrode metal
Suspension concentrates (SC) or flowables	Active ingredient Thickener Wetting agent Preservative Dispersant Anti-freezing agent Anti-foaming agent Water	No solvent Easy to measure and mix Can contain high concentration of AI Improves adhesion Compatible with aqueous concentrates Improved bio-efficacy as finer particle size	Difficult to produce successfully Can bleed, settle and cake out in storage Sensitive to freezing Can cause phytotoxicity Sensitive to AI purity and polymorphic form
Oil-in water emulsions (EW)	Active ingredient Emulsifier Solvent Thickener Anti-freezing agent Anti-foaming agents Preservative Water	Minimal skin and eye irritation Less or no solvent Minimal phytotoxicity Low or no flammability Ease of incorporating adjuvants Environmentally safer	Container disposal problem May be temperature sensitive Stabilization more difficult than for SC
Micro emulsions (ME)	Active ingredient No or minimum solvent Emulsifier Co-surfactant Water	Thermodynamically stable Good bio-efficacy Low flammability Easy application	High emulsifier level Low loading of active
Capsule suspension (CS)	Active ingredient Polymer Surfactant Water	Water based Low toxicity Reduced phytotoxicity Reduced loss on evaporation Reduced leaching Extended activity Lower doses Controlled release of AI	Expensive process High cost of inert ingredient Problems in achieving desired release of AI
Aerosols (AE)	Active ingredient Solvent Emulsifier Inhibitors Water Fragrances Propellant	Ready to use Safe and convenient for user Long shelf life Short application time	Inhalation risk Many propellants have environment issues and are highly flammable

Table 17.9 Formulation types and codes

Code	Description	Definition
AE	Aerosol dispenser	A container-he1d preparation which is dispersed generally by a propellant as fine droplets/particles upon actuation of a valve
AL	Any other liquid	A liquid not yet designated by a specific code, to be applied undiluted
AP	Any other powder	A powder not yet designated by a specific code, to be applied undiluted
BR	Briquette	Solid block designed for controlled release of active ingredient into water
CB	Bait concentrate	A solid or liquid intended for dilution before use as a bait
CP	Contact powder	Rodenticidal or insecticidal formulation in powder form for direct application formerly known as tracking powder (TP)
CS	Capsule suspension	A stable suspension of capsules in a fluid normally intended for dilution with water before use
DC	Dispersible concentrate	A liquid homogeneous preparation to be applied as a solid dispersion after dilution in water
DP	Dustable powder	A free-flowing powder suitable for dusting
DS	Powder for dry seed treatment	A powder for application in dry state directly to the seed
DT	Tablet for direct application	Formulation in the form of tablets to be applied individually and directly in the field, and/or bodies of water, without preparation of a spraying solution or dispersion
EC	Emulsifiable concentrate	A liquid, homogenous preparation to be applied as an emulsion after dilution in water
EG	Emulsifiable granule	A granular formulation, which may contain water insoluble formulants, to be applied as an oil-in water emulsion of the active ingredient(s) after disintegration in water
EO	Emulsion, water in oil	A fluid, heterogeneous preparation consisting of a dispersion of fine globules of pesticide in water in a continuous organic liquid phase
EP	Emulsifiable powder	A powder formulation, which may contain water insoluble formulants, to be applied as an oil-in water emulsion of the active ingredient(s) after dispersion in water
ES	Emulsion for seed	A stable emulsion for application to the seed either direct treatment or after dilution
EW	Emulsion, oil-in water	A fluid, heterogeneous preparation consisting of a dispersion of fine globules of pesticide in an organic liquid in a continuous water phase
FS	Flowable concentrate for seed treatment	A stable suspension for application to seed, either directly or after dilution
FU	Smoke generator	A combustible formulation, generally solid, which upon ignition releases the active ingredient(s) in the form of smoke
GA	Gas	A gas packed in pressure bottle or pressure tank
GE	Gas generating product	A preparation which generates a gas by chemical reaction
GL	Emulsifiable gel	A gelatinized formulation to be applied as an emulsion in water
GR	Granule	A free-flowing solid preparation of a defined granule size range ready for use
GS	Grease	Very viscous preparation based on oil or fat
GW	Water soluble gel	A gelatinized formulation to be applied as an aqueous solution
HN	Hot fogging concentrate	A formulation suitable for application by hot fogging equipment, either directly or alter dilution
KK	Combi-pack solid/liquid	A solid and a liquid formulation, separately contained within one outer pack, intended for simultaneous application in a tank mix
KL	Combi-pack liquid/liquid	Two liquid formulations, separately contained within one outer pack, intended for simultaneous application in a tank mix
KN	Cold fogging	A preparation suitable for application by cold fogging concentrate equipment, either directly or after dilution
Code	Description	Definition
LN	Long-lasting insecticidal net	A slow or controlled release formulation in the form of netting, providing physical and chemical barriers to insects
LS	Solution for seed treatment	A solution for application to the seed either directly or after dilution
MC	Mosquito coil	A coil which burns (smoulders) without producing a flame and releases the active ingredient into the local atmosphere as a vapour or smoke
ME	Micro-emulsion	A clear to opalescent oil and water containing liquid, to be applied directly or after dilution in water, when it may form a diluted micro-emulsion or a conventional emulsion
OD	Oil dispersible	A stable suspension of active ingredient(s) in a water-immiscible fluid, which may contain other dissolved active ingredient(s) intended for dilution with water before use
OF	Oil miscible flowable	A stable suspension of concentrate fluid intended for dilution in an organic liquid before use

(continued)

Table 17.9 (continued)

Code	Description	Definition
OL	Oil miscible liquid	A liquid, homogenous preparation to be applied as a homogenous liquid after dilution in an organic liquid
OP	Oil dispersible powder	A powder preparation to be applied as a suspension after dispersion in an organic liquid
PA	Paste	A water-based film forming composition
PR	Plant rodlet	A small rodlet, usually a few centimetres in length and a few millimetres in diameter containing active substance
PS	Seed coated with a pesticide	Self-defining
RB	Bait (ready for use)	A preparation designed to attract and be eaten by the target species
SC	Suspension concentrate/ flowable concentrate	A stable suspension of active substance(s) in a fluid intended for dilution with water before use
SD	Suspension concentrate for direct application	A stable suspension of active ingredient(s) in a fluid, which may contain other dissolved active ingredient (s) intended for direct application, to rice paddies, for example
SE	Suspo-emulsion	A f1uid, heterogeneous preparation consisting of a stable of active substance(s) in the form of solid particles and of fine globules in a continuous water phase
SG	Water soluble granules	A preparation consisting of granules to be applied as a true solution of active substance after dissolution in water but may contain insoluble inert ingredients
SL	Soluble concentrate	A liquid homogenous preparation to be applied as a true solution of the active substance after dilution with water
SO	Spreading oil	A preparation designed to form a surface layer on application to water
SP	Water soluble powder	A powder preparation to be applied as a true solution of the active substance after solution in water but which may contain insoluble inert ingredients
ST	Water soluble tablet	Formulation in form of tablets to be used individually, to form a solution of the active ingredient after disintegration in water. The formulation may contain water-insoluble formulants
SU	Ultra low volume (ULV) suspension	A suspension ready for use though ULV equipment
TB	Tablet	Pre-formed solids of uniform shape and dimensions, usually circular, with either flat or convex faces
TC	Technical material	A material resulting from manufacturing process comprising the active ingredient, together with associated impurities. This may contain small amounts of necessary additives
TK	Technical concentrate	A material resulting from manufacturing process comprising the active ingredient, together with associated impurities. This may contain small amounts of necessary additives and appropriate diluents
UL	Ultra low volume (ULV) liquid	A homogenous liquid ready for use through ULV equipment
VP	Vapour-releasing product	A preparation containing one or more volatile ingredients, the vapours of which are released into the air. Evaporation rate normally is controlled by using suitable preparations and/or dispensers
WG	Water dispersible	A preparation granule consisting of granules to be applied after dis-integration and dispersion in water
WP	Wettable powder	A powder preparation to be applied as a suspension after dispersion in water
WS	Water dispersible powder for seed treatment	A powder to be dispersed at high concentration in water before application as a slurry to the seed
WT	Water dispersible tablet	Formulation in the form of tablets to be used individually, to form a dispersion of the active ingredient after disintegration in water
XX	Others	Temporary category of all the formulations not listed above
ZC	A mixed formulation of CS and SC	A stable suspension of capsules and active ingredient(s) in fluid, normally intended for dilution with water before use
ZE	A mixed formulation of CS and SE	A fluid, heterogeneous formulation consisting of a stable dispersion of active ingredient(s) in the form of capsules, solid particles and fine globules in a continuous water phase, normally intended for dilution with water before use
ZW	A mixed formulation of CS and EW	A fluid, heterogeneous formulation consisting of a stable dispersion of active ingredient(s) in the form of capsules, and fine globules in a continuous water phase, normally intended for dilution with water before use

Note: Suffix “SB” should be added to the formulation code, if the material is packed in a sealed water soluble bag (e.g. WP-SB)

3. Prevention and control
   Selecting pest resistant varieties best suited for local growing conditions, maintaining healthy crops and rotation between different crops together with judicious use of pesticides in conjunction with action threshold.

Spurious Pesticides

One of the biggest challenges for the industry is rampant use of spurious pesticides. D’Arcy Quinn, anti-counterfeit director at Crop Life International, estimated up to 40% of

Table 17.10 Approved adjuvants terminology

Term	Definition
Acidifier	Material that can be added to spray mixtures to lower pH
Activator	Material that increases biological efficacy of agrochemicals
Amphoteric surfactant	A surfactant capable of forming, in aqueous solution, either surface active anions or surface active cations, depending on the pH
Anionic surfactant	A surface-active agent in which the active portion of the molecule containing the lipophilic segment forms exclusively a negative ion (anion) when placed in aqueous solution
Antifoaming agent	Material used to inhibit or prevent the formation of foam
Attractant	Material that attracts specific pests
Buffer/Buffering agent	A compound or mixture that, when contained in solution, causes the solution to resist change in pH. Each buffer has a characteristic limited range of pH over which it is effective
Canopy penetrating agent	An adjuvant that increases the penetration of the spray material into the crop canopy
Cationic surfactant	A surface-active agent in which the active portion of the molecule containing the lipophilic segment forms exclusively a positive ion (cation) when placed in aqueous solution
Colourant	A material used to alter the colour of tank mix
Compatibility agent	A surfactant that allows simultaneous application of liquid fertilizer and agrochemical formulations, as a uniform tank mix, or improves the homogeneity of the mixture and the uniformity of application
Crop oil concentrate	An emulsifiable petroleum oil-based product containing 15–20 % w/w surfactant and a minimum of 80 % phytoblend oil
Crop oil (emulsifiable)	An emulsifiable petroleum oil-based product containing upto 5 % w/w surfactant and remainder of phytoblend oil
Crop oil (nonemulsifiable)	Same as phytoblend oil
Defoaming agent	Material that eliminates or suppresses foam in the spray tank mix
Disposition aid	Material that improves the ability of agrochemical sprays to deposit on targeted surfaces
Dormant oil	A horticulture mineral oil applied during the dormant phase of the targeted plant
Drift control agent	A material used in liquid spray mixture to reduce spray drift
Emulsifier	A surfactant that promotes the suspension of one immiscible liquid in another
Evaporation reducing agent	A material that reduces the evaporation rate of a spray mix during or after application or both
Extender	A material that increases the effective life of an agrochemical after application
Foam suppressant	Same as defoamer
Foaming agent	A material that increases volume or stability of foam formed in spray mixture
Humectant	A material that increases the equilibrium water content and increases the drying time of an aqueous spray deposit
Modified vegetable oil	An oil extracted from seeds which has been chemically modified (e.g. methylated)
Modified vegetable oil concentrate	An emulsifiable, chemically modified vegetable oil product containing 5–20% w/w surfactant and remainder chemically modified vegetable oil
Naptha-based oil	A petroleum oil containing majority of the naphtha fraction
Nonionic surfactant	A material having no ionizable polar end groups but comprised of hydrophilic and lipophilic segments
Oil	Paraffinic oil or petroleum oil or vegetable oil
Paraffinic oil	A petroleum oil (derived from peraffinic crude) whose paraffinic carbon type content is typically greater than 60%
Penetrant	A material that enhances the ability of an agrochemical to enter a subtract or penetrate a surface
Petroleum oil	Oil derived from petroleum, contains a mixture of hydrocarbons that are broadly classified as paraffins, napthenes, aromatics, or other unsaturates or combinations thereof
Phytoblend oil	A highly refined paraffinic material with a minimum unsulfonated residue of 92% w/w
Spreader	A material which increases the area that a droplet of a given volume of spray mixture will cover on a target
Spreader/sticker	A material that has the properties of both a spreader and a sticker
Sticker	A material that assists the spray deposit to adhere or stick to the target and may be measured in terms of resistance to time, wind, water, mechanical action, or chemical action
Surfactant/surface active agent	A material that when added to a liquid medium modifies the properties of the medium at a surface or interface
Vegetable oil	Oil extracted from seeds typically those of corn, cotton, peanut, rapeseed, sunflower, canola or soya beans
Vegetable oil concentrate	An emulsifiable vegetable oil product containing 5–20% w/w surfactant and a minimum of 80% w/w vegetable oil
Wetting agent	Synonymous with spreading agent in function

pesticides in some countries would be faked. Unlike registered products, which undergo rigorous independent testing before being placed in the market, illegal pesticides are neither tested nor suitable for use. Nothing guarantees that a counterfeit product contains what is described on the label; a reality that can have damaging consequences for crops, human health and environment. They pose threats to farmers through crop and reputation destruction, as well as dangerous exposure during application. Residues may be carried on food to consumers.

All the relevant players from industry, regulatory, customs and enforcement agencies need to be brought together to determine the best ways solve the problem.

Safe Use of Pesticides

If pesticides are used improperly or excessively, they can have harmful effects on pests, crops and environment. Lack of proper knowledge among farmers/users leads to improper handling, use and disposal of pesticides. This is a common scenario in developing and underdeveloped countries.

If labels are read and instructions regarding protective clothing and correct dose are followed carefully, risk in pesticide handling and chances of harming crops by overdose can be reduced drastically. The number of applications and amount of pesticide used can be decreased by better timing, e.g. spray at economic threshold level when the pest is most vulnerable; young caterpillars are easier to kill than bigger ones.

One such example is Global Safe Use Pilot Projects launched by the pesticide industry in June 1991 in 3 countries, Guatemala, Kenya and Thailand with aid from Global Crop Protection Federation (GCPF). The Safe Use program in Guatemala that was aimed to “provide education and training and ensure that pesticides are used rationally and safely in the agricultural production process” claims to have reduced cases of pesticide poisoning considerably [12].

Increasing Development and Registration Cost

Bringing new molecules in the market is getting more and more expensive. It takes, on an average, 10–12 years from discovery of new Active substance to launching it in the market. Discovery and development cost of a new crop protection product is on the rise. This cost was approximately $150 million in 1995 and reached to more than $250 million in 2010. The cost is going to rise further as data requirement is becoming more complex and cost of data development is on a rise too as meeting regulatory framework for pesticides is mandatory before it is manufactured and used in a given country and states. This approval process is governed by national and regional/federal laws depending on locally applicable regulations.

Further, registered pesticides are subject to periodic review. Depending on the prevailing national regulations, authorities may review approvals at set intervals or at any time if new information comes to light.

Besides nationals authorities intergovernmental organizations, such as OECD, FAO, WHO and the Codex Commission also play a vital role in the regulation of pesticides.

The FAO has developed a voluntary Code of Conduct for the distribution and use of pesticides, which is very broad in scope and sets minimum standards at all points of the manufacture, marketing and use chain. The Codex Alimentarius Commission along with National Authorities sets limits on the permitted levels of pesticide residues in produce, so that the quality of food on supermarket shelves is assured. Other major international bodies like OECD and WHO provide forum for the discussion of pesticide regulatory issues at international level.

Companies seeking approval have to submit data package to the national regulatory authority. This data is used to support the proposed uses of the product. The data can be generated or commissioned by the company itself, derived from published material or purchased from third parties by paying data compensation. The cost of preparing the application for a new active substance is considerable, in the range of US$15–20 million.

As part of determination of chemical equivalence, a generic manufacturer is required to supply to the authority conducting the risk assessment, a data package covering its products. This package must include acute toxicology information, and, in addition, must provide information relating to manufacturing process and impurity profile of the technical material as produced in its plants as well as Five batch analysis of typical manufacturing lots, supporting the impurity limits. The authority then reviews the data package, and on a confidential basis, makes a comparison between the impurity profile and manufacturing data with the reference material, i.e. the primary sponsor of the specification, and those from the secondary manufacturer, in order to judge the equivalence of the second product with the reference material. Assuming that the impurity profiles are within the variance parameters specified in the FAO/WHO manual, the products may be judged to be equivalent for purposes of quality comparison for risk assessment. If, however, there are differences among the provided data as defined in the manual, the authority may require the generic manufacturer to supply additional data, even including sub-acute and chronic toxicology and ecotoxicology data, to support the equivalence assessment.

A complete assessment of the impurity content in a technical material under equivalence evaluation is very important since differences in impurity content may significantly affect the toxicological properties of pesticide products. This widely accepted principle is the basis for the rigorous requirement for complete “fingerprinting” (an impurity-by-impurity comparison over the entire range of impurities) during the equivalence assessment process.

Each country has different requirements for product registration, and this can be a daunting prospect for manufacturers seeking to offer their product for sale in many countries. Moves are underway to harmonize product registration requirement across regions or trading partners. OECD member governments and pesticide regulators have agreed to take a number of steps towards full harmonization of data requirements by 2014. Comprehensive information on registration system and requirements for each country can be obtained from their respective website referred in Annex 17.2 for certain countries.

Other examples of harmonization are the NAFTA agreement led to cooperative efforts to harmonize pesticide regulatory requirements between the US, Canada and Mexico and to Joint Review projects in the region. Similar agreements are in place in Latin America (Mercosur) and Asia Pacific (ASEAN).

Entry of Biotechnology/Transgenic Crops

There is a general belief that genetically modified (GM) crops overall reduce pesticide use but there are contradictions in practicality. For example recent surveys in the USA suggest that Bt-corn that targets corn borer has not lowered pesticide use, since most pesticides are directed against other corn pests. In India, consumption of insecticides increased many fold in Bt cotton for the control of sucking pests viz. whitefly, mealy bugs, thrips, aphids and bugs.

The introduction of glyphosate resistant soya beans in 1996 gave farmers an opportunity to reduce the cost of weed control; however, the increasing areas and continuous use of glyphosate has resulted in an increase in the types of weeds that are now glyphosate resistant.

Hence, new-generation “Roundup Ready” (RR) 2 soya beans seed, to be introduced on a widespread basis soon. Dicamba tolerance trait is expected to be stacked with the high yielding Genuity Roundup Ready 2 yield soya beans trait. These next generation technologies are aimed at offering farmers multiple modes of action and superior application timing and flexibility but the cost will be 40–45% more than the original RR seeds they are displacing.

Further, these new GM crops may give rise to super weeds and cause outbreaks of new pests and diseases that would need newer and more potent chemicals to control them. It remains to be seen how much impact GM crops will have on the pesticide industry for the development, registration and use of new pesticides to combat these super weeds, pests and diseases.

Delivery Systems and Application Technology

Application technology is now seen as an enabling technology, which can add significant value to pesticide products, and at the same time improve operator safety and reduce the dose rate and wastage of pesticides applied to crops, thereby reducing environmental impact and increasing food safety.

As per Dr. Zabkiewicz of the Forest research Institute, New Zeeland, the application of pesticides to leaf surfaces was a complex and often inefficient process. He pointed out that the efficiencies of the various stages of spray application are as follows:

Deposition 80%, retention 50%, Uptake 50% and Translocation 10%. This means for a systemic pesticide it is likely that no more than 2% of the active material actually reaches the target site [13].

New more precise delivery techniques are being investigated as a tool to increase efficacy and reduce input costs for growers. The emphasis on safety of users and environment is driving the move towards minimal contact with the product, more accurate placement of the active ingredient, and the use of safer formulations such as water-based suspensions, emulsions, microcapsules and water dispersible granules. The production of nanoparticle-based formulations has shown to increase the activity of an otherwise poorly dispersing product.

Seeds are an effective way of ensuring correct placement. Seed coating formulations are a value added technology that has found application in the establishment of a safe environment that protects the seed when it is most vulnerable, i.e. the time of storage and during initial seed germination. Seed treatments are estimated to enhance productivity by 8–10%. Seed treatment market is continuously growing. It was approximately US$900 million in 2001 and reached US $2.5 billion by 2009. Most GM seeds in US/Canada have minimum 4 coats. Use of microcapsule treatments or coated seeds will apart from improvements in safety give the potential for controlled release.

Cleaner Technology and Products

Manufacturing of certain pesticides involves use of hazardous chemicals and wastes. Clean production is defined as any practice which eliminates the use or formation of hazardous substances at source through the use of non-hazardous chemicals in production processes, or through product or process redesign, and thereby prevents releases of hazardous substances into the environment by all routes, directly or indirectly.

Recently, China Ministry has formed a new management scheme about clean production in pesticide industry.

Molecules with high eco toxicity, human carcinogens/mutagens, endocrine disruptors, persistent organic pollutants and persistent bio accumulative are being picked up by regulatory authorities for thorough investigation and those with unfavourable profiles are being phased out.

Preference is for pesticides that are significantly safer to humans or environment.

R&D efforts are also directed towards evolution of new agrochemicals, in terms of mode of action, structure and specificity for better yield.

Summary

In agriculture, there is no single solution that will be capable of addressing the problems that farmers and the growing world population face today and in the future. Neither biotechnology nor chemistry has all the answers. Instead, combinations of solutions and collaborations across the public and private sectors will be required to save the lives of billions of people from hunger and diseases of malnutrition.

Among toughest challenges the world is facing today are reducing carbon credits and saving rapidly decreasing non-renewable fossil fuels reserves. The alternative energy race to replace fossil fuels is gearing up. The concepts of renewable supply, reduced greenhouse gases, and energy security are driving biofuels opportunities globally. Currently, biofuels provide approximately 2% of the world's transportation fuel and is expected to reach 7% in 2020. This has created a big dilemma of Food vs. Fuel as the risk of diverting farmland or crops for biofuels production will affect the availability of the food, feed and fibre supply on a global scale. Hence, productivity improvements will remain the most important factor in sustainable agriculture. Low yields are not sustainable in the decades ahead.

For a sustainable agriculture there is a need for comprehensive and effective international pest and pesticide management schemes, education and training in proper pesticide management, solid national and legislative framework, global harmonization of pesticide registration requirement and fixing loopholes in the legislative framework to counter illegal trafficking in counterfeit pesticide products.

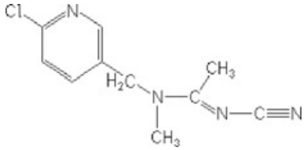
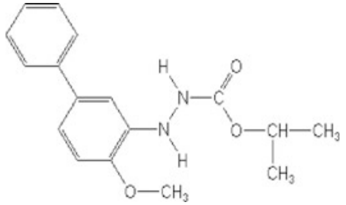
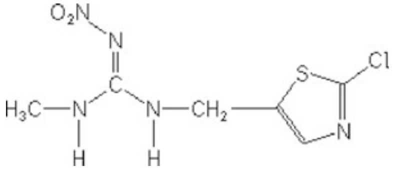
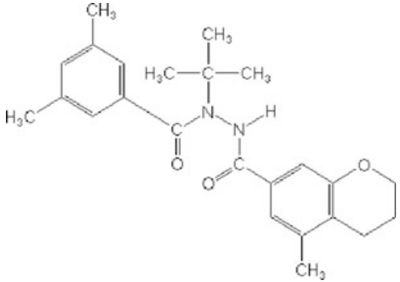
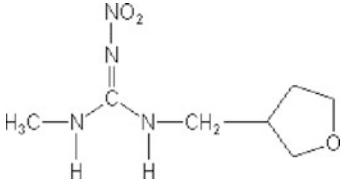
Recent trends in agrochemical industry, viz. consolidation of multinational agrochemicals companies, rapid growth of transgenic crop markets, vertical integration of agrochemical firms with seed companies and food processors, use of better quality seeds, and thrust towards environment friendly products and technologies, are driving the Agrochemical industry forward and paving the way towards ever green revolution.

Acknowledgments The authors express their sincere thanks and gratitude to Dr. K.H. Gharda, Chairman and Managing Director, Gharda Chemicals Limited, Mumbai, for his help and support in preparing this chapter. We also wish to thank Mr. Nitin Girase, Manager-Product Development and Ms Revti Chandrashekhar, Sr. Officer, Regulatory Affairs, Gharda Chemicals Limited, for their help in literature search, and preparation of manuscript.

References

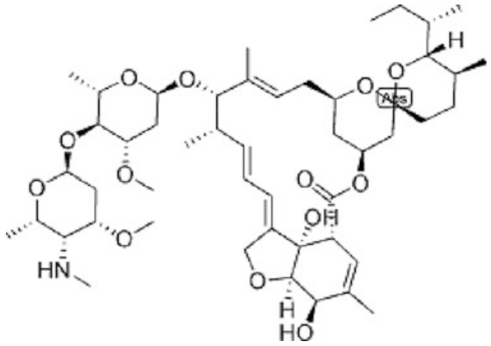
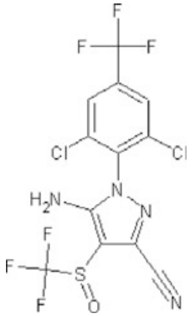
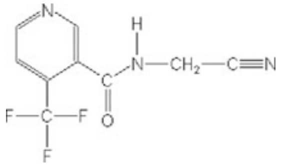
1. IRAC 'mode of action classification' version Sept 2010
2. HRAC 'mode of action classification' version Jan 2005
3. FRAC 'mode of action classification' version Dec 2009
4. The WHO recommended classification of pesticides by hazard and guidelines to classification 2009
5. US EPA—Label Review Manual-Chapter 7
6. Magnus W (2002) Walter, structure-based design of agrochemicals. Nat Prod Rep 19:278–291
7. Spindler F, Früh T (1998) Chiral acylanilides and chiral triazole-related fungicides. In: Kurihara N, Miyamoto J (eds) Chirality in agrochemicals. Wiley, Chichester, p. 143
8. US Environmental Protection Agency, EFED Interim policy for stereoisomeric pesticide, Appendix II-sterioisomerism: Background
9. Malte AM, Lilani A.T (2002) The agrochemical industry (Chapter 17). In: Riegel ER, Kent JA (eds) (2002) Kent and Riegel's handbook of industrial chemistry and biotechnology, vol. 1. Springer New York, p. 774
10. Global Crop Protection Federation (GCPF formerly GIFAP) (2008) Catalogue of pesticide formulation types and international coding systems, vol 2, 6th edn, Technical Monograph. CropLife International, Brussels
11. American Society for Testing and Materials E 1519 (1999) Annual book of ASTM Standards, Vol. 11.05. Designation E 1519–95, Standard terminology relating to agricultural tank mix adjuvants. pp. 905–906. ASTM International, West Conshohocken
12. Murray DL, Taylor PL (2000) Claim no easy victories: evaluating the pesticide industry's global safe use campaign. World Development 28(10):1735–1749
13. Crop protection monthly, 29 August 2002, issue no 153

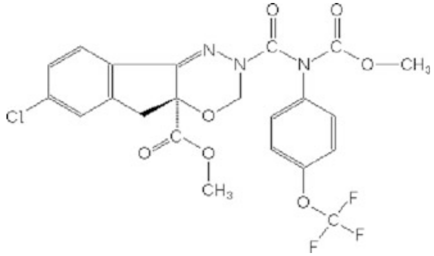
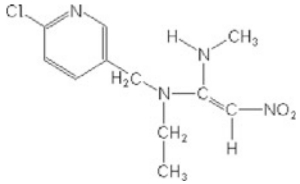
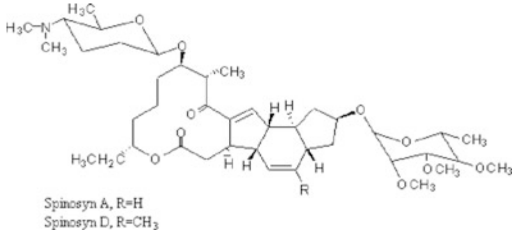
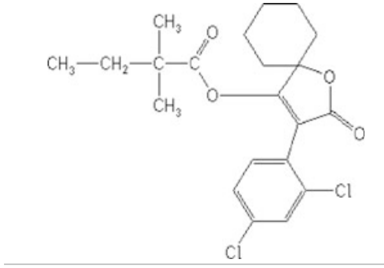
Annex 17.1 Products going off patent from 2010 to 2015 and beyond

Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
Insecticides becoming generic between 2010 and 2015					
1	Acetamiprid 	Neonicotinoid	Systemic with translaminar activity having both contact and stomach action. Acetylcholine receptor (nAChR) agonist Uses: for the control of sucking pests especially aphids and thrips	1995	US 2011 EU 2010 SPC 2015
2	Bifenazate 	Hydrazine carboxylate	Neuronal inhibitor, non-systemic having contact and residual action Uses: effective against phytophagous mites	1999	US 2011 EU 2012
3	Clothianidin 	Neonicotinoid	Translaminar and root systemic activity. Acetylcholine receptor (nAChR) agonist Uses: for the control of sucking and chewing pests on a range of crops	2002	US 2009 EU 2009 SPC 2014
4	Chromafenozide 	Diacylhydrazine	Ecdysteroid agonist. which competes at the binding site of the receptor protein of the insect moulting hormone Uses: for the control of various Lepidopteran species	1999	–
5	Dinotefuran 	Neonicotinoid	Systemic, with contact and stomach action. Nicotinic Acetylcholine receptor agonist/antagonist Uses: to control a wide range of sucking pests, including whiteflies, mealybugs and plant hoppers, in vegetables, fruit, turf and rice crops	2002	–

(continued)

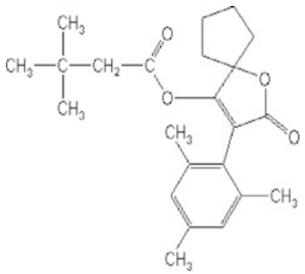
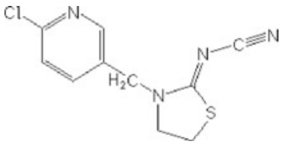
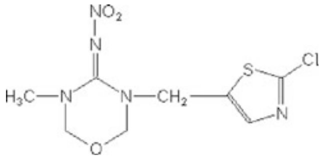
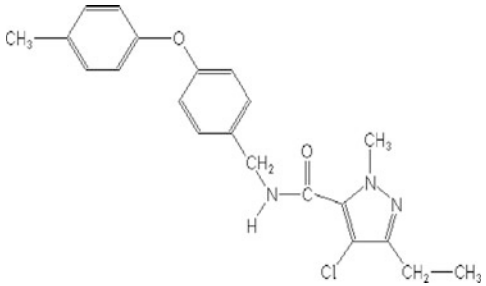
Annex 17.1 (continued)

Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
6	Emamectin benzoate 	Biopesticide	Non-systemic, acts by causing insect paralysis. Chloride channel activator Uses: effective against caterpillars in particular but also helps control mites, thrips and leaf miners	1997	US 2011 EU 2011 SPC 2015
7	Fipronil 	Phenylpyrazole	Broad spectrum with contact and stomach action. GABA-gated chloride channel antagonist Uses: for the control of a broad range of pests including thrips, rootworms, wireworms, weevils and termites	1993	US 2010 EU 2008 SPC 2008
8	Flonicamid 	Pyridine compound	Systemic, selective with long-term activity. Distrub insect feeding pattern Uses: To control aphids, thrips and whitefly in a range of situations	2000	–

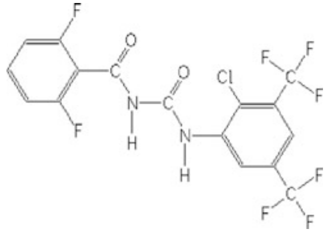
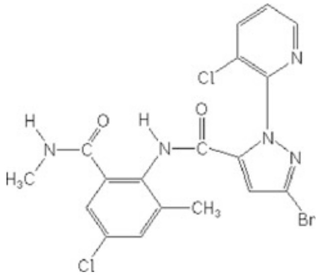
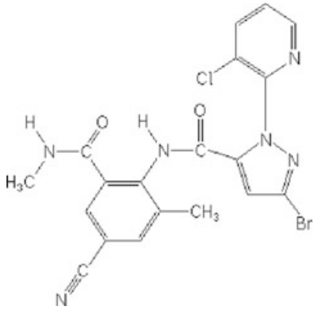
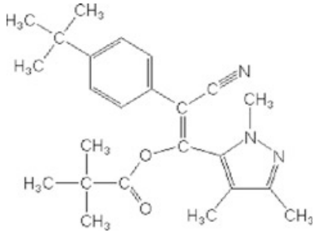
9	Indoxacarb  	Oxadiazine	Contact and stomach action. Voltage-dependent sodium channel blocker Uses: on a wide range of crops to control caterpillars in general and Helicoverpa in particular	2000	US 2012 EU 2011
10	Nitenpyram  	Neonicotinoid	Acetylcholine receptor (nAChR) agonist. Systemic with translaminar activity, contact and stomach action Uses: To control sucking insects on rice and in greenhouse crops. Mainly used to kill external parasites in livestock and domestic pets	1995	–
11	Spinosad  	Biopesticide	Contact and stomach action. Acts by exciting the insect nervous system, leading to involuntary muscle contractions and paralysis Uses: for control of Lepidoptera caterpillars	1997	US 2013 EU 2009 SPC 2014
12	Spirodiclofen  	Tetronic acid	Selective, non-systemic, distrupts mite development Uses: foliar applied acaricide active against a range of mites on fruit and nut	2003	–

(continued)

Annex 17.1 (continued)

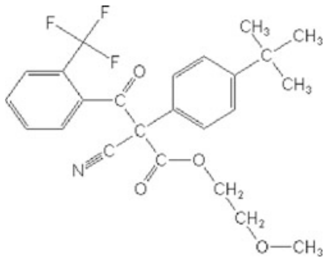
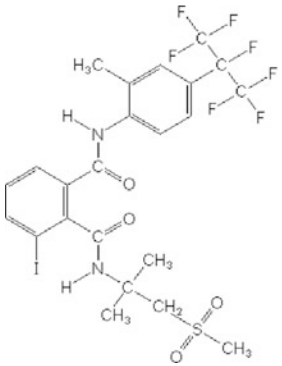
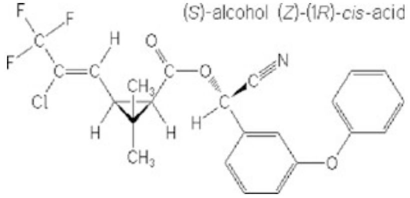
Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
13	Spiromesifen 	Tetronic acid	Non-systemic. Inhibitors of lipid synthesis Uses: against whiteflies and two-spotted spidermite	2002	–
14	Thiacloprid 	Neonicotinoid	Contact and stomach action with some systemic properties. Acetylcholine receptor (nAChR) agonist Uses: for control of sucking and chewing insects	1999	US 2007 EU 2007 SPC 2012
15	Thiamethoxam 	Neonicotinoid	Acetylcholine receptor (nAChR) agonist. Broad spectrum, systemic with contact and stomach action Uses: To control a wide range of pests including aphids, whiteflies, thrips, mealybugs, wireworms and ground beetles	1997	–
16	Tolfenpyrad 	Pyrazole	Broad spectrum, contact activity, exhibits antifeedant activity especially against lepidoptera Uses: for use in vegetables, particularly cruciferous leafy vegetables, fruits and other high-value markets. Also shows some fungicidal activity against powdery mildew	2002	–

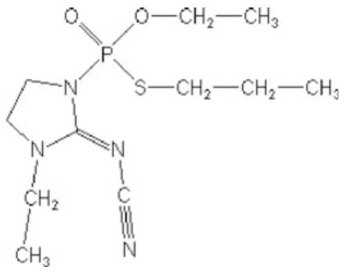
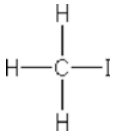
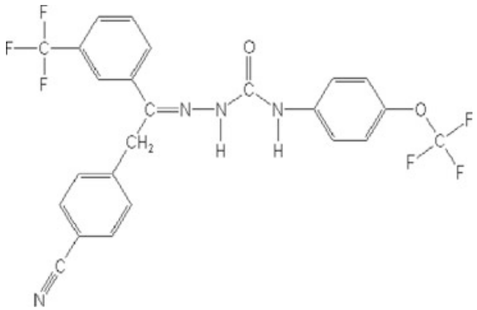
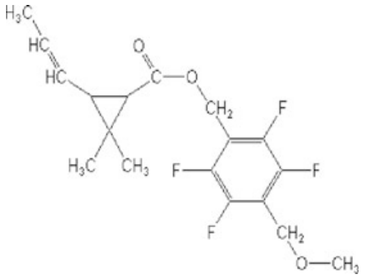
Insecticides (likely to become generic post 2015)

1	Bistrifluron 	Benzoylurea	Chitin synthesis inhibitor Uses: for the control of aphids, whiteflies and other lepidopterous insects	Under development	–
2	Chlorantraniliprole 	Anthranilic diamide	Exhibits larvicidal activity post ingestion by targeting and disrupting the Ca^{2+} balance causing muscle contraction Uses: Broad spectrum control of caterpillars in field, vegetable crops and vine	2007	–
3	Cyantraniliprole 	Diamide	Exhibits larvicidal activity as an orally ingested toxicant by targeting and disrupting the Ca^{2+} balance, second generation ryanodine receptor Uses: Foliar and systemic activity on a wide range of crops that is effective against many key chewing and sucking pests including whitefly, thrips, aphids and fruitflies	2008	–
4	Cyenopyrafen 	Pyrazole	Mitochondrial complex II electron transport inhibitor Uses: An acraicide for fruit, vegetables and tea	2006	–

(continued)

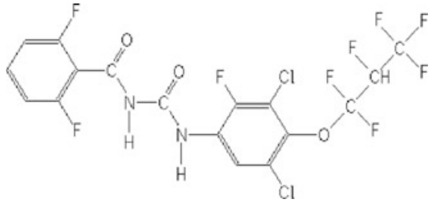
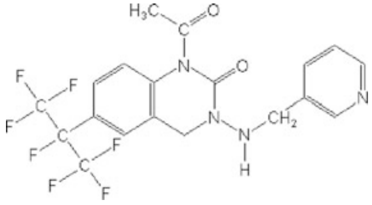
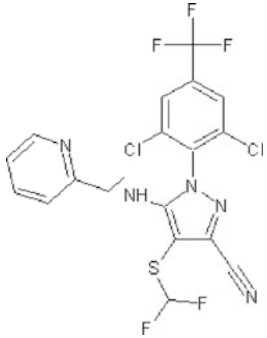
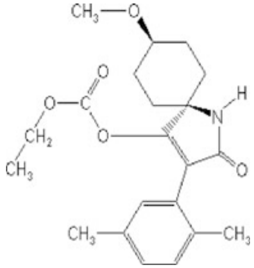
Annex 17.1 (continued)

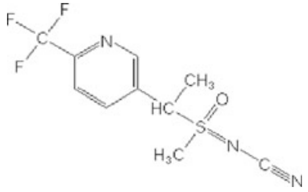
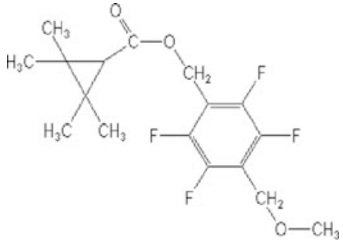
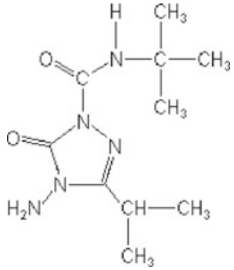
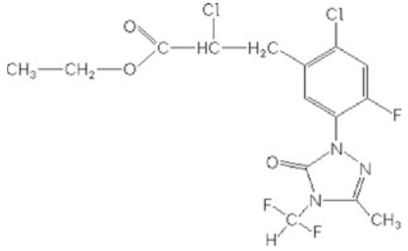
Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
5	Cyflumetofen 	Bridged diphenyl	Highly selective and causes loss of motor coordination leading to paralysis and death Uses: novel acaricide active against phytophagous mites at all development stages	2004	–
6	Flubendiamide 	Benzene-dicarboxamide	Exhibits larvicidal activity as an orally ingested toxicant by targeting and disrupting the Ca^{2+} balance Uses: lepidopteran insecticide for use in wide range of annual and perennial crops	2006	–
7	Gamma-cyhalothrin  	Pyrethroid	Non-systemic acting by contact and/or ingestion. Sodium channel modulator Uses: broad-spectrum control of insect pests in a wide range of crops	2003	–

8	Imicyafos 	Phosphonothioate	High contact activity with fast diffusion in soil Uses: A novel nematicide that is effective against knot root, root lesions, cyst nematodes	2006	–
9	Iodomethane 	Alkyl iodide	Fumigant action thought to be via the nucleophilic displacement (SN2) reaction in various amino acids and peptides within target organisms Uses: to control insects, plant parasitic nematodes, soil borne pathogens and weed seeds	2004	–
10	Metaflumizone 	Semicarbazone	Broad-spectrum, attacks insect nervous system causing paralysis Uses: against key insect pests including Lepidoptera, Coleoptera, Hymenoptera, Isoptera and Diptera	2008	–
11	Metofluthrin 	Pyrethroid ester	Vapour active, knock-down action. Sodium channel modulator Uses: as a mosquito repellent impregnated on to paper strips and positioned in outdoor areas such as gardens and campsites	2004	–

(continued)

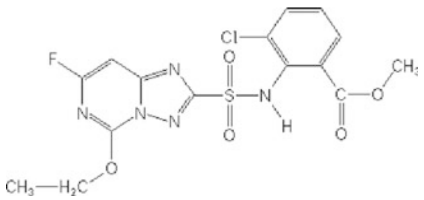
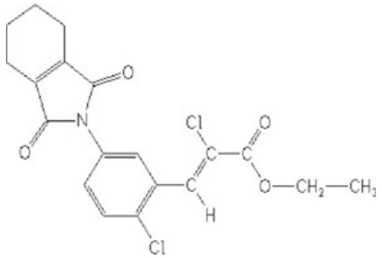
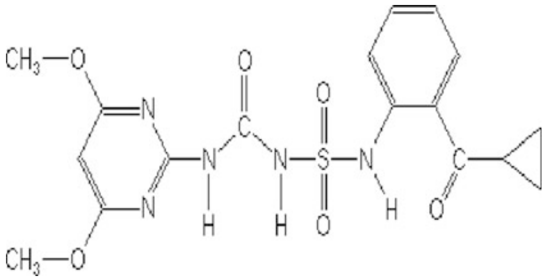
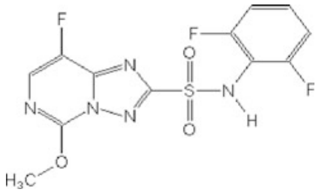
Annex 17.1 (continued)

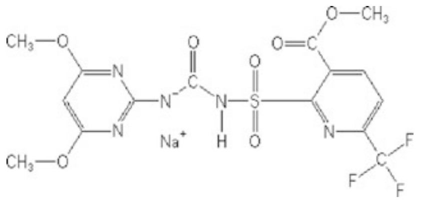
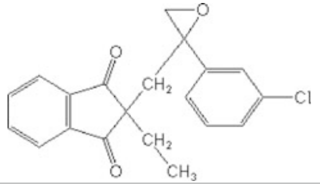
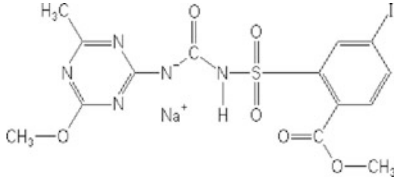
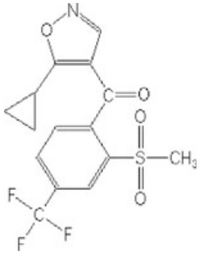
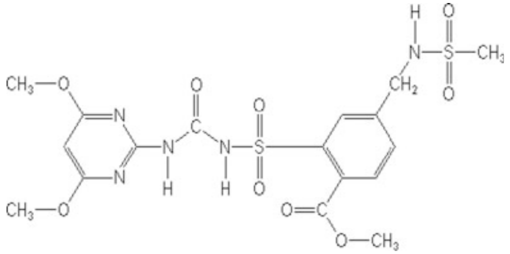
Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
12	Noviflumuron 	Benzoylurea	A chitin synthesis inhibitor Uses: insect growth regulator to control termites, ants, cockroaches, fleas and flies in a wide range of pest management solutions	2003	–
13	Pyrifluquinazon 	Quinazinalone	Modifies insect behaviour, rapidly stopping feeding such that insects starve to death Uses: to control aphids, thrips, mealybugs and scale on vegetables and ornamentals	2006	–
14	Pyriprole 	Phenylpyrazole	Causes uncontrolled activity of parasites nervous system and hence its death, broad-spectrum, non-systemic Uses: Insecticide, Ectoparasiticide, Acaricide for the treatment for fleas and ticks on animals	2004	–
15	Spirotetramet 	Tetramic acid	Stomach acting, broad spectrum, systemic insecticide that is rapidly translocated, inhibition of lipogenesis in treated insects Uses: for the control of a wide range of sucking insects including aphids, whiteflies and psyllids on crops such as fruit, vegetables and potatoes	2007	–

16	Sulfoxaflor 	Sulfilimine	Mode of action not fully understood Uses: for the control of aphids in cotton and peaches	2007	–
17	Tetramethylfluthrin 	Pyrethroid ester	Mode of action not fully understood Uses: A novel insecticide	2008	–
Herbicides becoming generic between 2010 and 2015					
1	Amicarbazone 	Triazolone	Contact action, absorbed by foliage, photosynthesis II inhibition Offers pre- and post-emergence control of annual dicotyledonous weeds and grasses in sugar cane and other crops	1999	US 2009 EU 2010
2	Carfentrazone-ethyl 	Triaolinone	Contact, absorbed by foliage with limited translocation. Cell membrane disruption—PPO inhibitor Post-emergence control of a wide range of broad-leaved weeds	1997	US 2009 EU 2009 SPC 2012

(continued)

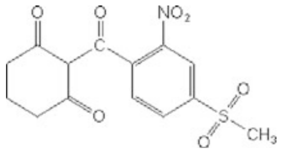
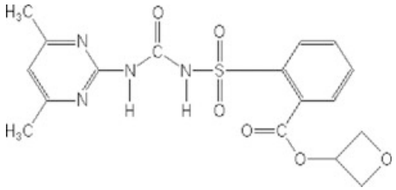
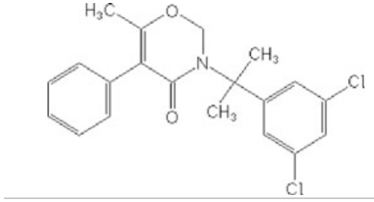
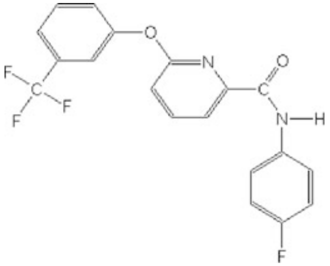
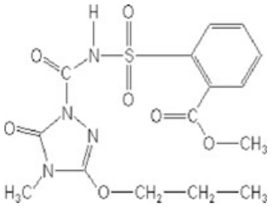
Annex 17.1 (continued)

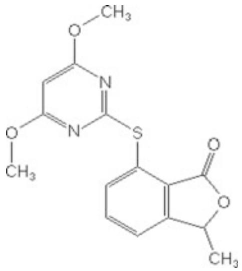
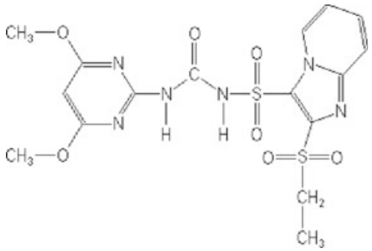
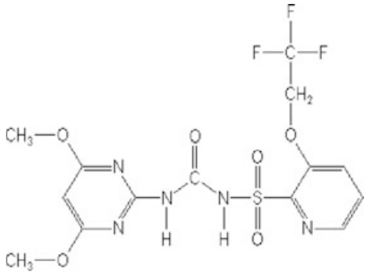
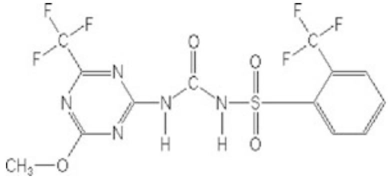
Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
3	Cloransulam-methyl 	Sulfonanilide	Inhibits acetolactate synthesis. The primary site of activity is in plant meristem It is used to control broad-leaved weeds in soya beans and other crops	1997	–
4	Cinidon-ethyl 	Dicarboximide	Contact action Post-emergence control of annual broad-leaved weeds	1998	US 2008 EU 2007 SPC 2012
5	Cyclosulfamuron 	Sulfonyl urea	Inhibition of acetolactate synthase (ALS) stopping cell division and plant growth Post-emergence control a variety of dicotyledonous and sedge weeds in cereals and rice	1997	US 2010 EU 2011
6	Florasulam 	Triazolopyrimidine	Selective, absorbed by roots and shoots. Inhibits plant enzyme acetolactate synthase (ALS) Post-emergence control of broad-leaved weeds	1999	US 2008 EU 2009 SPC 2013

7	Flupyrsulfuron-methyl-sodium 	Sulfonylurea	Inhibition of acetolactate synthase Post-emergence control of broad-leaved weeds in cereals	1998	US 2012 EU 2012 SPC 2013
8	Indanofan 	Unclassified	Selective, systemic, acts by the inhibition of very long chain fatty acid biosynthesis Post-emergence control of sedges such as *Echinochloa oryzicola*, and some dicotyledonous in rice crops	1999	US 2010 EU 2010
9	Iodosulfuron-methyl-sodium 	Sulfonylurea	Acetolactate synthase (ALS) inhibitor Post-emergence control grass and broad-leaved weeds in cereals	2000	–
10	Isoxaflutole 	Isoxazole	Broad-spectrum, systemic absorbed by roots and leaves. Acts by carotenoid biosynthesis inhibition Post-emergence control of broad-leaved weeds and grasses in maize and sugar cane	1998	EU 2010
11	Mesosulfuron-methyl 	Sulfonylurea	Amino acid synthesis inhibitor Post-emergence control of grasses and some broad-leaved weeds in cereals	2001	–

(continued)

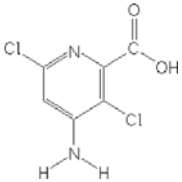
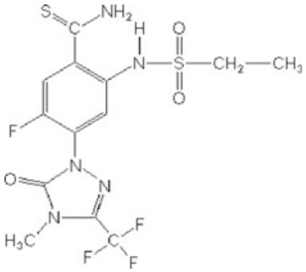
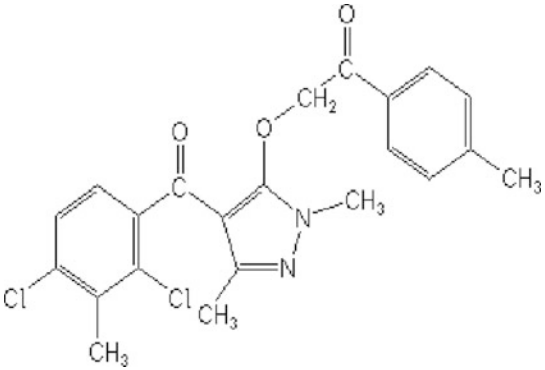
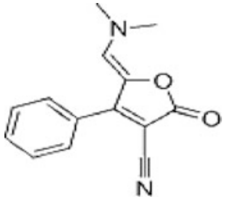
Annex 17.1 (continued)

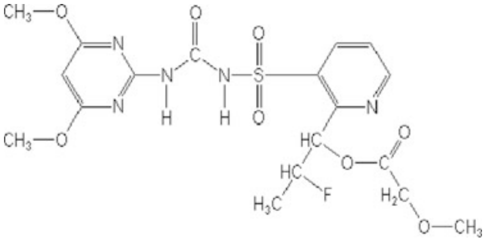
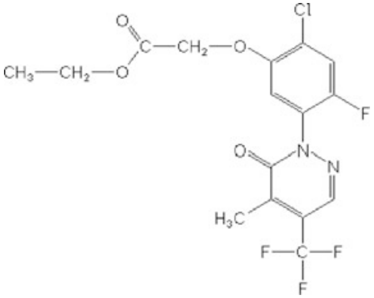
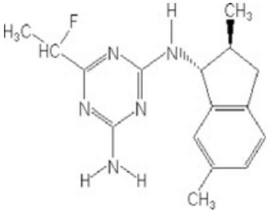
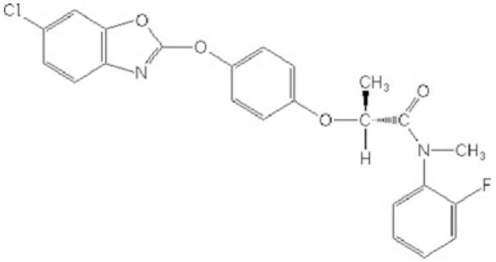
Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
12	Mesotrione 	Triketone	Selective, absorbed by roots and translocated, inhibition of carotenoid biosynthesis Post-emergence control of some grasses and broad-leaved weeds mainly in maize crops	2001	US 2008 EU 2005 SPC 2010
13	Oxasulfuron 	Sulfonylurea	ALS inhibitor Absorbed by shoots and roots and translocated. Post-emergence herbicide for grass and broad-leaved weed control	1996	–
14	Oxaziclomefon 	Unclassified	Blocks plant growth by inhibiting cell expansion in grass roots Effective in controlling grasses and annual sedges in paddy fields	2000	–
15	Picolinafen 	Pyridine compound	Selective, with rapid foliar absorption Post-emergence to control broad-leaved weeds in cereals	2001	US 2011 EU 2011 SPC 2017
16	Propoxycarbazone 	Triazolone	Absorbed by leaves and roots and translocated. Inhibits amino acid synthesis Residual post-emergence herbicide for grass and some broad-leaved weeds mainly in winter wheat	2000	US 2014 EU 2012 SPC 2016

17	Pyriftalid 	Pyrimidinyloxybenzoic	Inhibits amino acid synthesis, taken up mainly by roots Used to control grass weeds in rice and other crops	2000	–
18	Sulfosulfuron 	Sulfonylurea	Systemic, absorbed by roots and leaves and translocated. Acetolactate synthase (ALS) inhibitor Used for annual grass and broad-leaved weed control in cereals	1999	US 2013 EU 2011
19	Trifloxysulfuron 	Sulfonylurea	Inhibitor of acetoacetate synthase, absorbed by shoots and roots and translocated Post-emergence control of grasses, sedges and broad-leaved weeds in cotton, sugar cane and other crops	2000	–
20	Tritosulfuron 	Sulfonylurea	Inhibits amino acid synthesis Uses: to control a wide range of broad-leaved weeds in cereals and maize	–	US 2012 EU 2011

(continued)

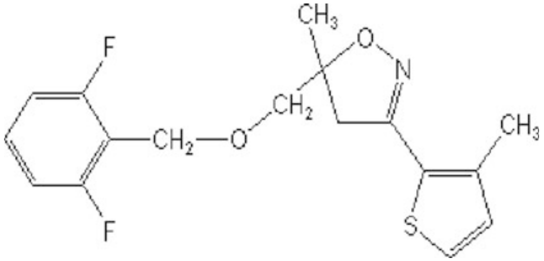
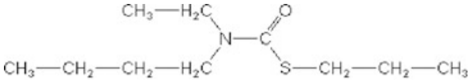
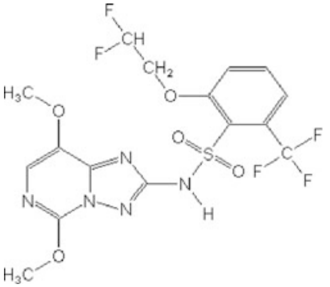
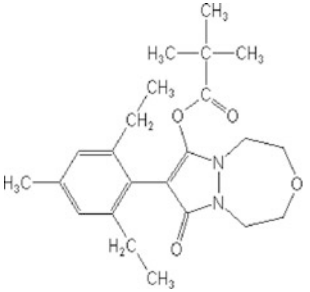
Annex 17.1 (continued)

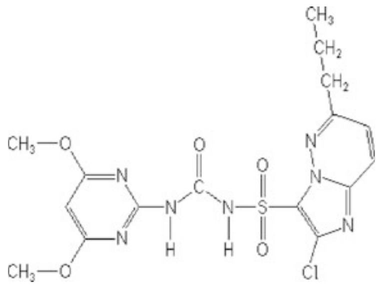
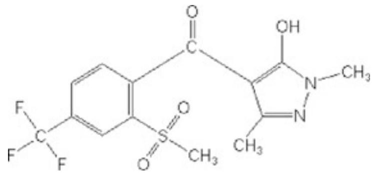
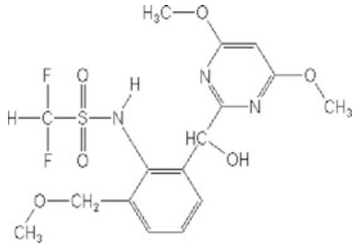
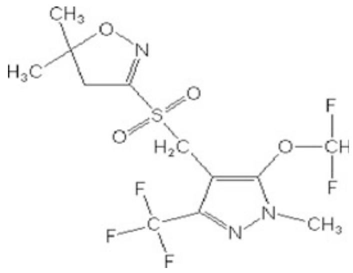
Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
Herbicides (likely to become generic post 2015)					
1	Aminopyralid  	Pyridine carboxylic acid	Systemic, post-emergence action, absorbed by leaves and roots Uses: for the long-term control of noxious and invasive broad-leaved weeds in grassland	2005	–
2	Bencarbazone  	Triazalone	Protox inhibitor, absorbed by roots and foliage and translocated Uses: Post-emergent control of broad-leaved weeds in cereals, sugar cane and corn	2005	–
3	Benzofenap  	Benzoylpyrazole	Systemic, absorbed through roots. Bleaching inhibition of 4-hydroxyphenyl-pyruvate-dioxygenase Uses: mainly to control broad-leaved weeds and sedges in rice crops	2002	–
4	Difunon  	Phenylfuranone	Selective contact and residual action. Bleaching: inhibition of carotenoid biosynthesis Uses: Post-emergence herbicide active against grassy weeds	2005	–

5	Flucetosulfuron 	Sulfonylurea	Broad spectrum, absorbed by roots, stem and leaves, acetolactate synthase (ALS) inhibitor Uses: post-emergence herbicide for the control of broad-leaf weeds, some grass weeds and sedges particularly *Echinochloa* sp. in rice and some cereal crops	2003	–
6	Flufenpyr-ethyl 	Pyridazinone	Rotoporphyrinogen oxidase inhibitor Uses: on potatoes, maize, soya and wheat	2003	–
7	Indaziflam 	Fluoroalkyltriazine	Non-selective, broad spectrum activity Uses: broad spectrum of weeds, including species which are difficult to eliminate such as annual bluegrass, goosegrass, ryegrass and goosefoot	2008	–
8	Metamifop 	Aryloxyphenoxy propionate	Acetyl CoA carboxylase (ACCase) inhibitor, causes chlorosis leading to growth retardation Uses: post-emergent herbicide for the control of a wide range of annual grass weeds in cereal crops and rice	2002	–

(continued)

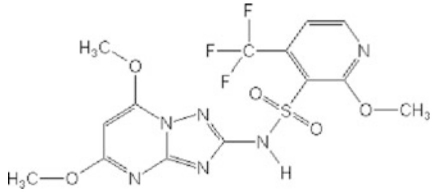
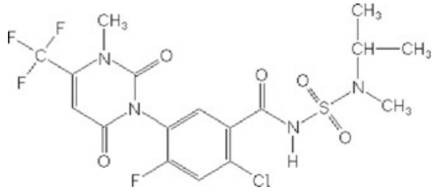
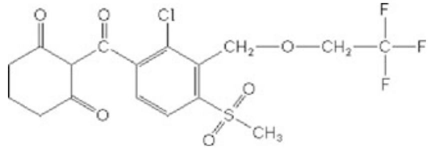
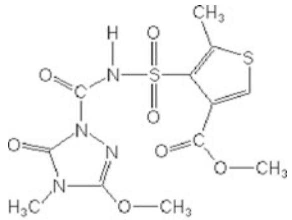
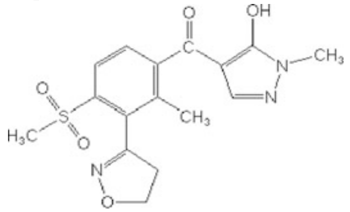
Annex 17.1 (continued)

Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
9	Methiozolin 	Oxazole	Mode of action not fully understood Uses: A novel herbicide for turf	2008	–
10	Pebulate 	Thiocarbamate	Selective, systemic, absorbed through roots and translocated Uses: for control of grass and broad-leaved weeds	2001	–
11	Penoxsulam 	Triazolopyrimidine sulfonamide	Broad spectrum, absorbed mainly by leaves and, to lesser extent, roots and translocated, exhibits knockdown and residual control Uses: on rice crops for the control of sedges, broad-leaved weeds, aquatic plants and certain grasses	2004	–
12	Pinoxaden 	Phenyl pyrazoline	Systemic. acetyl CoA carboxylase (ACCase) inhibitor, inhibiting fatty acid synthesis	2006	–

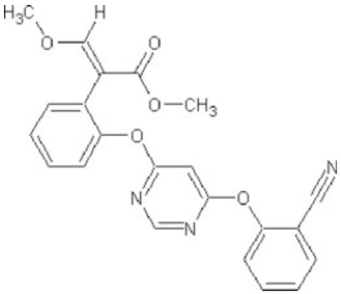
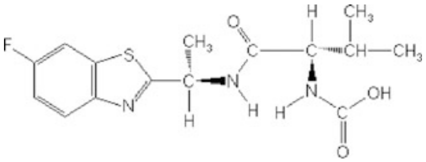
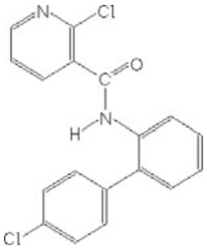
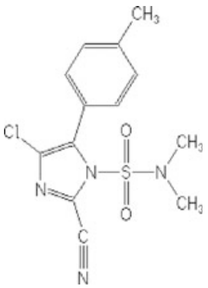
13	Propyrisufuron 	Pyrimidinylsulfonylurea	Mode of action not fully understood Uses: new herbicide for use on paddy rice	2008	–
14	Pyrasulfotole 	Pyrazolone	Selective, acts by the inhibition of carotenoid biosynthesis specifically inhibits the enzyme 4-hydroxyphenylpyruvate dioxygenase (HPPD) Uses: for post-emergent control of various broad-leaf weeds in cereals, used with the safener mefenpyr-diethyl	2004	–
15	Pyrimisulfan 	Sulfonanilide	Acetolactate synthase (ALS) inhibition Uses: for the control of perennial weeds predominately in rice crops	2004	–
16	Pyroxasulfone 	Pyrazole	Sulfonylioxazoline acetolactate synthase (ALS) inhibitor, mode of action not fully understood. Good residual action Uses: pre-emergent grass and broad-leaved weed herbicide which is selective in wheat and corn	2005	–

(continued)

Annex 17.1 (continued)

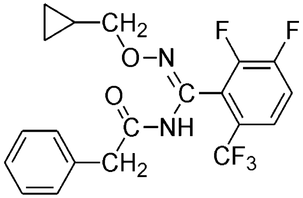
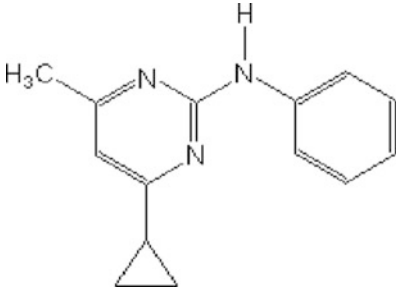
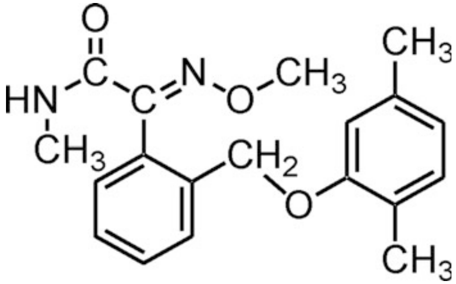
Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
17	Pyroxsulam 	Triazolopyrimidine sulfonamide	Broad spectrum having both foliar and soil activity, translocates and inhibits acetolactate synthase (ALS) Uses: control of annual grass and broad-leaf weeds in winter wheat	2007	–
18	Saflufenacil 	Uracil/amide	Select, contact and residual action, PPO inhibitor Uses: solo or mixed with glyphosate for rapid control of many dicotyledonous weeds by pre-plant and pre-emergence applications to a wide range of food crops	2009	–
19	Tembotrione 	Triketone	Uses: post-emergence broad spectrum herbicide for the control of grassy and broad-leaf weeds in corn and other crops	2007	–
20	Thiencarbazone-methyl 	Triazolone	Acetolactate synthase (ALS) inhibitor, systemic, absorbed by roots and leaves, contact and residual action Uses: in mixtures with isoxaflutole for pre-emergence control of grassy and broad-leaved weeds in corn, soya and other crops	2008	–
21	Topramezone 	Benzoylpyrazole	Systemic, highly selective, rapidly absorbed and translocated Uses: post-emergence herbicide for broad-leaved weeds and grasses mainly in corn	2006	–

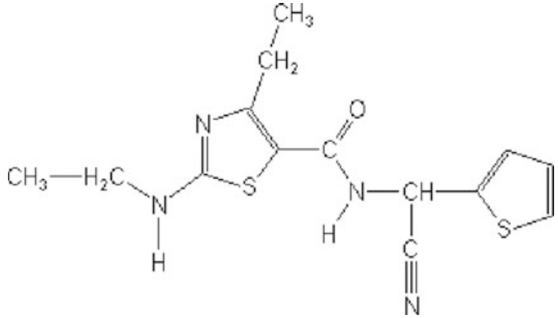
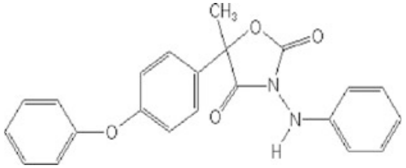
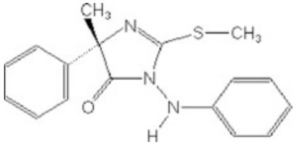
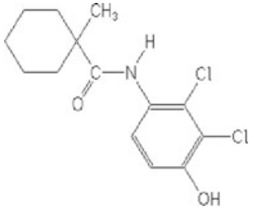
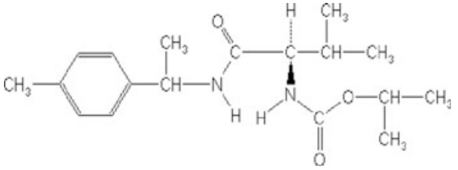
Fungicides becoming generic between 2010 and 2015

1	Azoxystrobin 	Strobilurin	Systemic translaminar and protectant action having additional curative and eradicant properties. Respiration inhibitor. Broad spectrum fungicide mainly for cereals	1992	EU 2010 SPC 2011
2	Benthiavalicarb 	Carbamate	Protective and curative action with residual effects, inhibits phospholipid biosynthesis Uses: against Oomycetes fungal plant pathogens	2003	–
3	Boscalid 	Carboxamide	Protectant, foliar absorption, translocates, inhibits spore germination and germ tube elongation Uses: against a broad range of fungal pathogens for use on a wide range of crops including fruit, vegetables and ornamentals	2002	–
4	Cyazofamid 	Cyanoimidazole	Foliar and soil preventative action with some residual activity. Respiration inhibitor Recommended for controlling Oomycete and Plasmodiophora diseases on potatoes and tomatoes	2001	US 2008 EU 2008 SPC 2013

(continued)

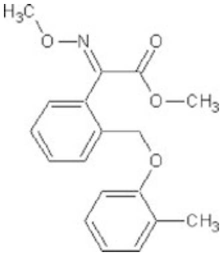
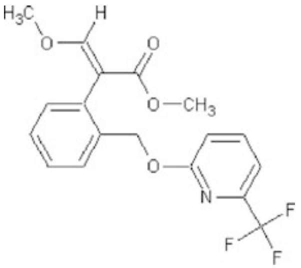
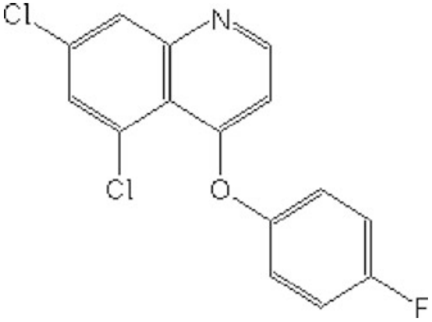
Annex 17.1 (continued)

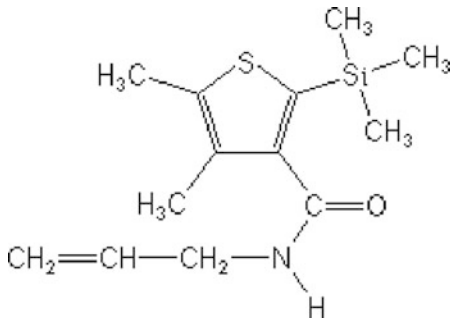
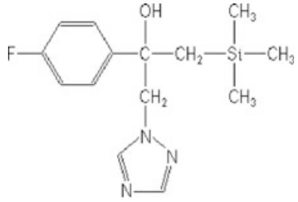
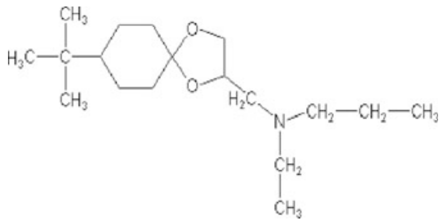
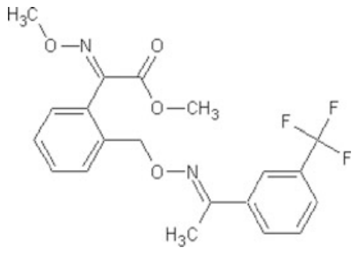
Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
5	Cyflufenamid 	Amide	Preventative and curative activity, acts by inhibiting the infection process Uses: for the control of powdery mildew in cereals and other crops	2002	–
6	Cyprodinil 	Anilinopyrimidine	Systemic, absorbed through foliage. Inhibits protein synthesis Uses: to control a range of pathogens including *Tapesia yallundae*, *Botrytis* spp., *Alternaria* spp. and *Rhynchospium secalis*	1998	US 2008 EU 2008 SPC 2013
7	Dimoxystrobin 	Strobilurin	Protectant, curative and translaminar action. Respiration inhibitor Uses: for disease control in cereals, vegetables and turf	2003	–

8	Ethaboxam 	Thiazole carboximide	Systemic with protective and curative action Used to control Oomycetes in vines and late potato blight	1995	–
9	Famoxadone 	Oxazole	Protectant with residual activity, interferes with plant respiratory system Used to control a broad spectrum of plant pathogenic fungi including downy mildew and blights	1998	US 2010 EU 2010 SPC 2015
10	Fenamidone 	Imidazole	Protective and curative action. Respiration inhibitor Used against Oomycete diseases such as downy mildew and certain leaf spot diseases	2001	US 2014 EU 2012
11	Fenhexamid 	Hydroxyanilide	Foliar applied with protective action. Disrupts membrane function. Inhibits spore germination Used to control *Botrytis cinerea* and related pathogens in fruit, vegetables and ornamentals	1999	US 2009 EU 2009 SPC 2013
12	Iprovalicarb 	Carbamate	Systemic with protective, curative and eradicative action. Lipid synthesis inhibitor Used to control Oomycete pathogens	1999	US 2012 EU 2011

(continued)

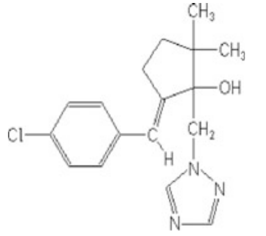
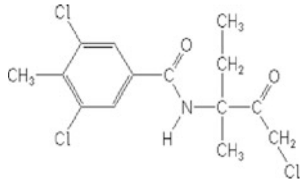
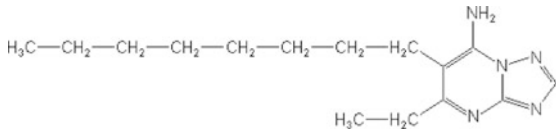
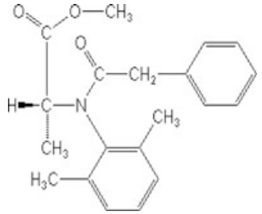
Annex 17.1 (continued)

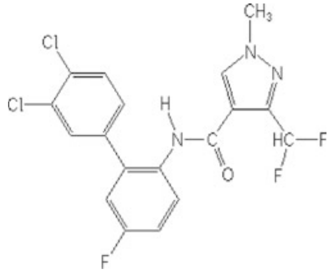
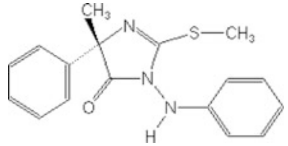
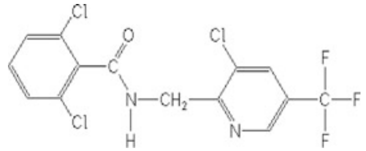
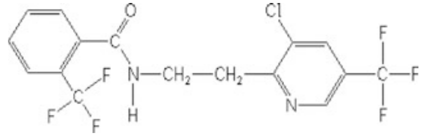
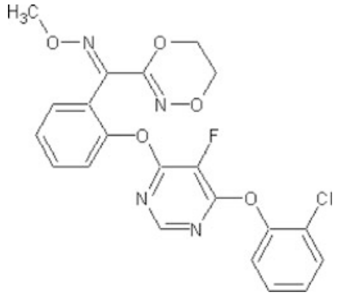
Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
13	Kresoxim-methyl 	Strobilurin	Protective, curative, eradicative action and long residual effects, acts by blocking electron transfer and respiration of the fungi Recommended for the control of scab on apples and pears and other fungal diseases on a wide range of crops	1998	US 2007 EU 2007 SPC 2011
14	Picoxystrobin 	Strobilurin type-methoxyacrylate	Strobilurin type-methoxyacrylate Uses: to control a wide range of diseases including brown rust, tan spot and powdery mildew	2001	US 2008 EU 2008 SPC 2013
15	Quinoxyfen 	Quinoline	Systemic with protective properties, translocates and inhibits appressoria development stopping infections Used to control *Erysiphe graminis*—powdery mildew in cereals	1996	US 2009 EU 2009 SPC 2012

16	Silthiofam 	Thiophene	Protective action with long-lasting activity Used for cereals mainly to control "take-all"	1999	EU 2012
17	Simeconazole 	Conazole	Sterol demethylation inhibitor. Systemic, absorbed by roots and translocates A broad spectrum biofungicide used to control a range of diseases on rice, fruit, soya and vegetables	2001	–
18	Spiroxamine 	Morpholine	Systemic with protective, curative and eradicative action. Disrupts membrane function Used to control powdery mildew, leaf spots, rusts and other diseases	1998	US 2008 EU 2008 SPC 2012
19	Trifloxystrobin 	Strobilurin	Broad spectrum with preventative and curative action. Respiration inhibitor Used for cereals which is particularly active against Ascomycetes, Deuteromycetes and Oomycetes	1999	US 2011 EU 2011 SPC 2013

(continued)

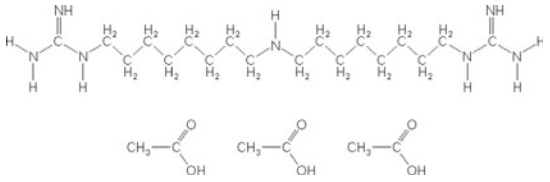
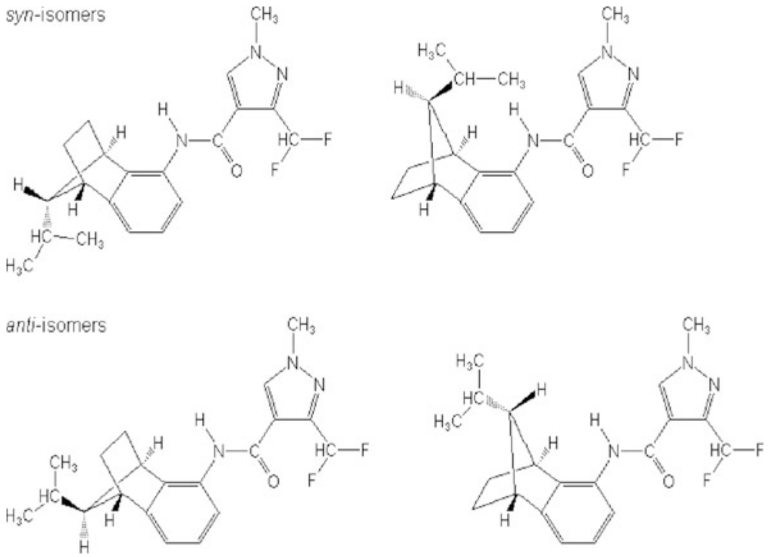
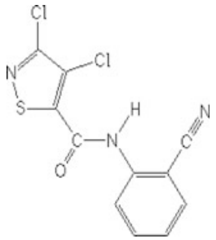
Annex 17.1 (continued)

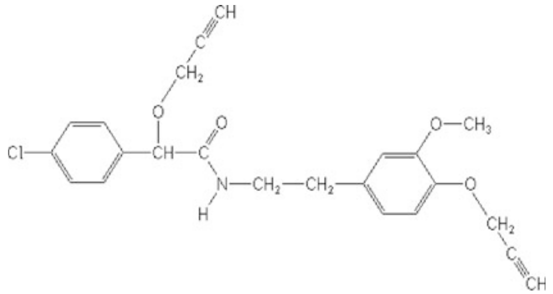
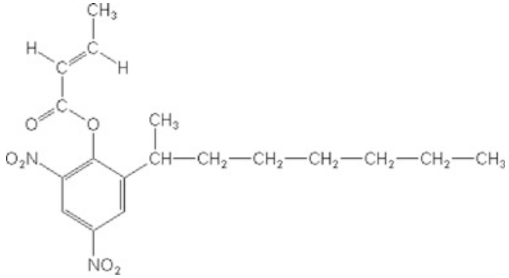
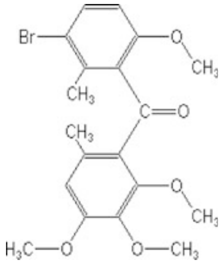
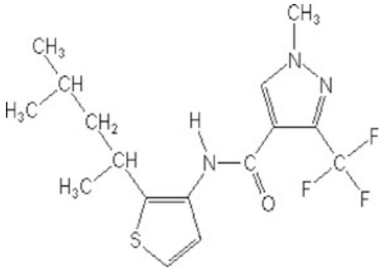
Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
20	Triticonazole 	Triazole	Inhibits sterol demthylation Used for the control of common bunt, loose smut and covered smut on barley, oats and wheat	1993	US 2010 EU 2009
21	Zoxamide 	Benzamide	Rainfast preventative action with residual properties, acts by inhibiting nuclear division Used for the control of various fungal inflections including blight in potatoes and tomatoes	2001	US 2012 EU 2013
Fungicides (likely to become generic post 2015)					
1	Ametoctradin 	Triazolopyrimidine	Selective, preventive, mitochondrial respiration inhibitor Uses: for the control of late blight and Downy mildew on potatoes and other crops including vines	2009	–
2	Benalaxyl-M 	Phenylamide	A systemic fungicide with protective, curative and eradication action Uses: to control Oomycetes particularly late blights on potatoes and tomatoes, and Downy mildew on a variety of crops	2003	–

3	Bixafen 	Pyrazole	Mode of action not known Uses: cereals for key disease control including strobilurin-resistant septoria in cereals	2006	–
4	Fenamidone 	Imidazole	Protective and curative action. Respiration inhibitor Uses: against Oomycete diseases such as downy mildew and certain leaf spot diseases	2001	–
5	Fluopicolide 	Benzamide	Protectant Uses: for the control of a range of diseases including downy mildew and blight	2006	–
6	Fluopyram 	Benzamide, pyramide	Mode of action not known Uses: for foliar application and as a seed treatment to control botrytis, powdery mildew and other diseases	2008	–
7	Fluoxastrobin 	Strobilurin	Systemic with protectant and curative properties Uses: for disease control of potatoes and a wide range of vegetables	2002	–

(continued)

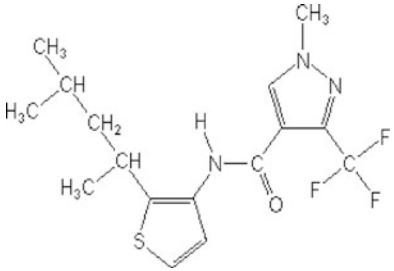
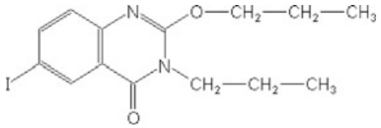
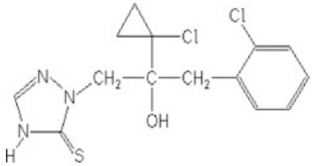
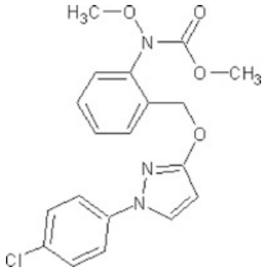
Annex 17.1 (continued)

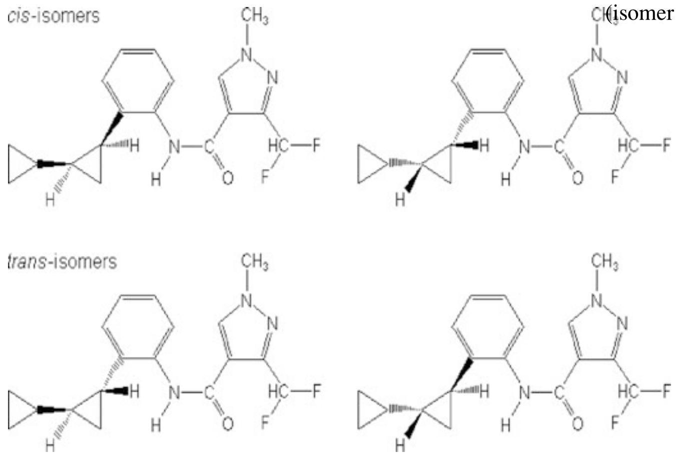
Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
8	Iminoctadine triacetate 	Guanidine	Protectant. Mode of action not known Uses: on fruit, trees and lawns to control a variety of pathogens including *Alternaria* and Gloeodes	2004	–
9	Isopyrazam 	Benonorbene	Protectant, foliar absorption, translocates, inhibits respiration and spore germination and germ tube elongation Uses: for the control of *Septoria tritici*, rusts and other diseases in cereals and some other crops	2008	–
10	Isotianil 	Thiazole	Mode of action not known Uses: to control rice blast	2007	–

11	Mandipropamid 	Mandelamide	Inhibits spore germination with preventative action Uses: to control Oomycete pathogens such as *Plasmopara viticola* on grapes and *Phytophthora infestans* on potatoes and tomatoes	2005	–
12	Meptyldinocap 	Dinitrophenol	Protectant, curative and eradicant activity by inhibiting spore germination Uses: active against powdery mildew mainly on grapes, cucurbits and strawberries	2007	–
13	Metrafenone 	Benzophenone	Protective and curative action. Interferes with hyphal morphogenesis Uses: to control eyespot and powdery mildew on cereals	2006	–
14	Orysastrobin 	Strobilurin	Systemic, broad spectrum with protective and curative properties. Interferes with the respiration process Uses: rice fungicide highly effective against *Magnaporthe oryzae, Pyricularia oryzae, Thanatephorus cucumeris* and *Rhizoctonia solani*	2006	–

(continued)

Annex 17.1 (continued)

Sr No	AI common name and structure	Chemical class	Mode of action/Biochemistry	Launch year	Year of patent expiry
15	Penthiopyrad 	Carboxamide	Broad spectrum, inhibits mitochondrial functions Uses: to control rust and *Rhizoctonia* diseases and also shows activity against grey mould, powdery mildew and apple scab	2003	–
16	Proquinazid 	Quinazolinone	Preventative and curative activity Uses: to control powdery mildew in a range of crops including cereals and grapes	2005	–
17	Prothioconazole 	Triazolinthione	Systemic with protective, curative and eradicative action. Long-lasting activity Uses: both as a seed treatment and foliar spray to treat a variety of diseases in cereals	2002	–
18	Pyrametostrobin 	Carbanilate	Systemic translaminar and protectant action having additional curative and eradicant properties. Respiration inhibitor Uses: active against powdery mildew and other diseases	–	–

19	Sedaxane 	Pyrazole anilide (isomer mix)	Broad spectrum Uses: for the control of pathogens such as Ascomycetes and Oomycetes species in crops such as fruit and wheat	2008	–

Note: SPC is UK Supplementary Protection Certificate Maximum Expiry Date

Disclaimer: In case of patent dates there may be other patents and patent application not identified. Authors do not give any guarantee to its accuracy or completeness

Annex 17.2 Regulatory authorities and legislation of some countries

Country	Regulatory authority	Legislation	Website
The European Union	27 member states each with its own regulatory body	Directive 1107/2009 Succeeding directive 91/414/EU	http://www.pesticideinfo.org/Detail_Country.jsp?Country=European%20Union
The United States	Environment Protection agency (EPA)	Federal Insecticide, Fungicide and Rodenticide Act (FIFRA)	www.epa.gov/pesticides/regulating/index.htm
Japan	Agricultural Chemicals Inspection Station (ACIS) Ministry of Agriculture, Forestry and Fisheries (MAFF)	Agricultural Chemicals Regulation Law No. 82, 1948, amended on 30 march 2007	www.acis.famic.go.jp/eng/shinsei/index.htm
Brazil	The national sanitary surveillance agency (ANVISA)	Law 9.782/1999	www.anvisa.gov.br
Australia	Australian Pesticides and Veterinary Medicines authority (APVMA)	Agricultural and Veterinary Chemical Act 1992	www.apvma.gov.au/registration/registering.shtml
China	Institute for the Control of Agrochemicals of the Ministry of Agriculture (ICAMA)	Regulation on Pesticide Administration (RPA) 1997 amended in 2002 and 2004	www.chinapesticide.gov.cn/en/en.asp
India	Central Insecticide Board & Registration Committee (CIB&RC)	Insecticide Act 1968	www.cibrc.nic.in

Petroleum and Its Products

18

Paul R. Robinson

Introduction

Petroleum makes the world go around. In the realm of global economics and politics, that statement is hard to dispute. Petroleum is uniquely versatile. When wisely handled, it is safe and clean. It is still abundant, and it can be stored for years at a time simply by leaving it in the ground. Under ambient conditions, it is a liquid and relatively noncompressible, so it can be carried across oceans in large tankers or pumped through pipelines hundreds of miles long. Its distilled products have high energy densities. Fifteen gallons of gasoline can move a 3,300-lb automobile across 350 flat miles at 65 miles per hour. (In metric units, those values are 57 L, 1,480 kg, 565 km, and 105 kph.) We transform petroleum into thousands of useful substances: fuels, lubricants, and chemicals with exceedingly different properties—from baby oil to pesticides, from adhesives to laxatives, from artificial sweeteners to sulfuric acid (Tables 18.1 and 18.2).

Crude oil and its coproduced cousin, natural gas, provide nearly 60 % of our energy (Fig. 18.1). Petroleum makes poor countries rich, and the economic health of rich countries depends on its cost and availability. Petroleum and its derivatives—especially fuels and explosives—have played a major role in all major wars since the start of the twentieth Century. And lest we forget: our unabating dependence on petroleum, natural gas, and other fossil hydrocarbons impacts the quality of our air, our water, and the earth itself.

Economics and politics. One cannot overstate the impact of petroleum on world economics and politics. The supply/demand picture is heavily skewed. As shown in Table 18.3, in 2008 the Middle East consumed far less oil than it produced, accounting for 7.7 % of global consumption but 30 % of global production; the 7.7 % is misleadingly high, because it includes the energy consumed to convert crude oil into exported products. Conversely, the United States and Western Europe consumed far more than they produced. The United States holds only 1.6 % of proven reserves, but it was the third largest producer (10.0 %) and by far the largest consumer (22.8 %). From January 2010 through March 2011, petroleum-related imports were responsible for 34–50 % of the total US trade deficit (Fig. 18.2).

Due to the need to import so much of such a valuable commodity—much of which is obtained from hostile and/or sensitive environs such as the North Sea, the North Slope of Alaska, and offshore California—the world's largest economies are vulnerable. Politics can be hostile, too. The Arab Oil Embargos of 1967 and 1973–1974 and the Iranian Revolution of 1979 caused prices to increase dramatically. The events correspond to price spikes in Fig. 18.3, which shows how petroleum prices have varied since 1861. Figure 18.3 also shows that prices in 2010 were at historical highs, both in present-day and inflation-corrected dollars. Both in the 1970s and recently, the higher prices led to reduced consumption (Fig. 18.4).

Consumption began to fall immediately, but it took a few years to reach "bottom." People started buying smaller cars, but not all at once. And while it's easy for people to drive less and install low-wattage light bulbs, it takes time to replace old equipment with newer, more-efficient equipment in power plants and factories. Unfortunately, the consumption "bottom" was transitory. When prices fell, consumption went back up.

Data selected from [3] show that 2009 was a year of milestones for the following reasons:

1. Global primary energy consumption fell by 1.1 %, the largest decline since 1980
2. OPEC oil production fell by 7.3 %, the largest decline since 1983
3. Natural gas production fell by 2.1 %, the first decline on record
4. Coal's share of world energy consumption rose to 29.4 %, the highest since 1970

P.R. Robinson (✉)
Shell Global Solutions, 3418 Clear Water Park Drive, Katy,
TX 77450, USA

J. A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_18, © Springer Science+Business Media New York 2012

Table 18.1 Selected petroleum products

Product	Uses, comments
Fuels	
Propane	Home heating, motor vehicles
n-Butane	Gasoline blend stock, portable lighters
i-Butane	Gasoline blend stock, production of C7 alkylate for gasoline
Naphtah/ gasoline	Spark-ignition engine fuel. Several different grades
Kerosene/jet	Home heating, illumination. Airplane fuel. Several grades
Gas oil/diesel	Diesel fuel. Four major grades including ULSD[a]
Fuel oil	Six major grades
Petroleum coke	Power plant fuel, fuel, and anodes for making steel
Lubricants	
>> 800 products—from "baby oil" to axle grease—all with stringent specifications	
Construction materials	
Asphalt	Paved roads, water-proofed buildings. Four major classes
Waxes	Coated paper packaging, water-proofing, candles
Miscellaneous	
Sulfur	Sulfuric acid and fertilizers
Needle coke	Anodes for making aluminum and high-grade steel
Graphite	Metallurgy, specialty fibers, paints, pencils

[a] ULSD stands for ultra-low-sulfur diesel

5. China and the United States accounted for similar percentages of total world energy consumption. Per Table 18.4: in 2009 China consumed 0.2% less than the United States—2,177 MTOE vs. 2,182 MTOE, respectively, where MTOE stands for million tonnes of oil equivalent. On a per capita basis, US consumption was 4.2 times higher than China's.

The first four milestones can be attributed to the record-high cost of oil and gas. As mentioned, people can and do conserve when energy prices go high enough.

Table 18.2 Selected petrochemicals and uses

Petrochemicals	Used to make
Alcohols	
Methanol	MTBE, acetic acid, formaldehyde . . .
Acetic acid	Vinyl acetate ⇒ plastics, latex paints. Cellulose acetate
Cellulose acetate	Yarns, sheet plastic, films, lacquers
Formaldehyde	Resins, butanediol, inks, nylon
Ethyl alcohol	Gasoline, solvents, personal care products, ethyl chloride, ethyl acetate . . .
Ethyl acetate	Solvent (e.g., nail polish remover)
Isopropyl alcohol	Solvents, personal care products
Higher alcohols	Solvents, plasticizers, detergents
Aromatics	
Benzene	Styrene, cumene, cyclohexane, maleic anhydride . . .
Toluene	Polyurethane, gasoline
Xylenes	Polyester fibers, resins, plasticizers, gasoline
Aromatics derivatives	
Cyclohexane	Nylon precursors
Cumene and phenol	Phenolic resins, epoxy, and polycarbonate resins
Olefins	
Ethylene	Polyethylene, ethylene dichloride, ethylene oxide
Ethylene dichloride	Polyvinyl chloride (PVC) ⇒ plastics
Ethylene oxide	Polystyrene, ethylene glycol
Ethylene glycol	Antifreeze, polyethylene terephthalate (PET) ⇒ plastic bottles
Propylene	Polypropylene, propylene oxide ⇒ propylene glycol
Propylene glycol	Antifreeze, polyesters, pharmaceuticals, hand sanitizers
Butenes	Maleic anhydride . . .
Butadiene	Complex polymers, including synthetic rubbers
Alpha olefins	Polymers, surfactants, synthetic lube oil and additives, fatty acids
Other	
Acrylointrile, acrylic acid	Polymers, including transparent polymers (Plexiglass, Lucite)
Aldehydes and ketones	Acetic acid, acetone, other solvents
Isobutane	Methylmethacrylate resins
Maleic anhydride	Polyesters, resins, plasticizers, dicarboxylic acids, Nylon precursors

Insert: primary energy sources vs. energy vectors. Periodically, news reporters tout zero-emission electric cars or hydrogen-powered vehicles that produce nothing but water. But when one considers the ultimate source of the electricity and hydrogen, such vehicles can be very dirty indeed.

Electricity is not a primary source of energy. It is an energy vector, generated from primary sources and carried through power lines to energy consumers.

Molecular hydrogen also can serve as an energy vector. (In the Sun, of course, where hydrogen undergoes nuclear fusion, it is a primary energy source.) On Earth, commercial hydrogen production plants employ the steam reforming of hydrocarbons or the electrolysis of water. Steam reforming coproduces CO_2, a significant "green-house gas." Electrolysis also produces CO_2 if the required electricity comes from coal, oil, or natural gas. Thanks to environmental laws, burning fossil fuels is cleaner than it used to be, but it still generates CO_2, sulfur compounds, nitrogen compounds, and in the case of coal, significant amounts of ash. If the electricity comes from hydroelectric or nuclear power plants, then electrolytic hydrogen is carbon-neutral. Compared to electricity, hydrogen is easier to store. As hydrogen gas, it is kept in salt domes, high-pressure tanks, and caverns such the ones near Borger and Sweeney, Texas. It is also stored as a hydride in special metal alloys. Hydrogen is transported from production plants or storage facilities through pipelines or in tube trailers.

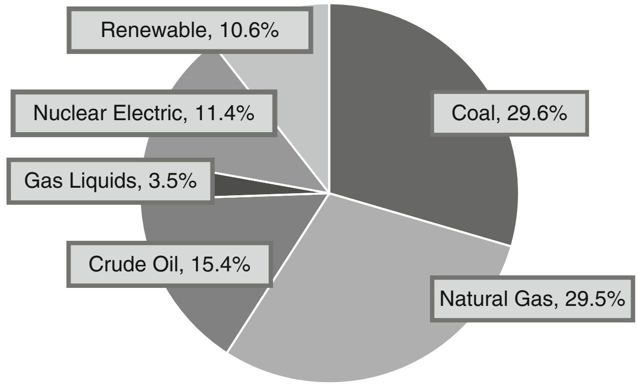


Fig. 18.1 Worldwide primary production of energy, 2007. Based on data from the U.S. Energy Information Administration (EIA) [1, 2]

Insert: conservation and production efficiency. When discussing the future of energy supply and demand, it is important to consider the potential for increased conservation and power production efficiency.

According to a report by the Rocky Mountain Institute [4], energy production efficiency in the United States increased 65–75 % since 1970. The ten top-performing states accounted for a large portion of the increase, creating a huge efficiency gap between the top-performing states and the rest of the nation; the data were normalized for differences in climate and the ratio of residential vs. industrial consumption. If the rest of the country achieved the same electric productivity as the Top Ten, the country would save about 1.2 million gigawatt-hours annually. This is equivalent to 30 % of the nation's total power consumption and 62 % of its coal-based power. The return on investment for efficiency improvement depends on several factors, especially economy of scale and the cost of capital, but the top-performing states achieved significant returns when interest rates were 3–5 times higher than in 2011. Power production from fossil fuels is responsible for 39 % of all CO_2 emissions in the United States, so a 30 % improvement in power production efficiency would lead to a huge reduction in CO_2 production [5].

Table 18.3 Worldwide petroleum reserves, production, and consumption in 2008

	Oil reserves		Oil production		Oil consumption	
	10^9 Barrels	Percent (%)	MBPD	Percent (%)	MBPD	Percent (%)
North America	212	15.9	15,032	17.6	23,896	28.0
Canada	179	13.4	3,332	3.9	2,260	2.6
Mexico	12	0.9	3,187	3.7	2,128	2.5
United States	21,	1.6	8,514	10.0	19,498	22.8
Central, S. America	110	8.2	7,411	8.7	6,167	7.2
Brazil	12	0.9	2,439	2.9	2,520	2.9
Venezuela	87	6.5	2,639	3.1	760	0.9
Europe	14	1.1	5,218	6.1	16,145	18.9
France	0.1	0.0	80	0.1	1,986	2.3
Germany	0.4	0.0	151	0.2	2,569	3.0
Italy	0.4	0.0	164	0.2	1,639	1.9
Norway	6.9	0.5	2,464	2.9	220	0.3
United Kingdom	3.6	7.4	12,535	14.7	4,333	5.1
Eurasia	99	7.4	12,535	14.7	4,333	5.1
Kazakhstan	30	2.3	1,431	1.7	239	0.3
Russia	60	4.5	9,794	11.5	2,900	3.4
Middle East	748	56.2	25,841	30.2	6,568	7.7
Iran	138	10.4	4,180	4.9	1,755	2.1
Iraq	115	8.6	2,386	2.8	638	0.7
Kuwait	104	7.8	2,729	3.2	325	0.4
Qatar	15	1.1	1,203	1.4	129	0.2
Saudi Arabia	267	20.0	10,783	12.6	2,297	2.7
United Arab Emirates	98	7.3	3,046	3.6	463	0.5
Africa	115	8.6	10,879	12.7	3,216	3.8
Algeria	12	0.9	2,181	2.6	299	0.3
Angola	9.0	0.7	2,014	2.4	64	0.1
Libya	41	3.1	1,875	2.2	273	0.3
Nigeria	36	2.7	2,169	2.5	286	0.3
Asia & Oceania	34	2.6	8,592	10.0	25,138	29.4
China	16	1.2	3,987	4.7	7,850	9.2
India	5.6	0.4	888	1.0	2,940	3.4
Japan	0.0	0.0	134	0.2	4,785	5.6
World	1,332	100	85,507	100	85,462	100

Based on data [1]

Fossil Hydrocarbons: Categories

Industry professionals divide fossil hydrocarbons into several categories. As with many things in the oil industry, the definitions are functional and the distinctions can be fuzzy, especially between heavy oil and bitumen. Natural gas is often lumped with liquid petroleum in articles such as this, because underground oil is always associated with gas, and because we use similar technology to discover and recover them.

Gases

Natural gas is fully discussed in Chap. 20. Here, it is sufficient to say the following:

- Natural gas contains mostly methane (CH_4)
- So-called *wet gas* contains significant amounts of butanes, pentanes, and heavier hydrocarbons, which can

Fig. 18.2 United States total trade deficit and the trade deficit due to petroleum. Based on data from the U.S. Energy Information Administration [1]

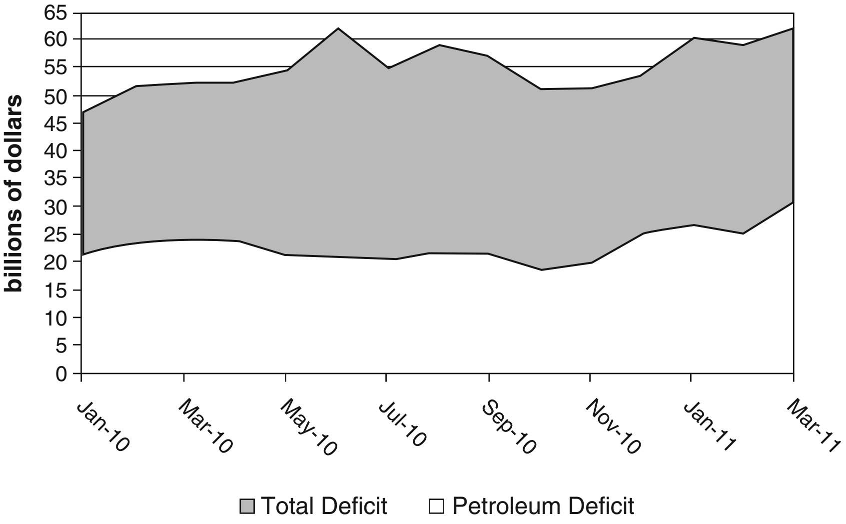
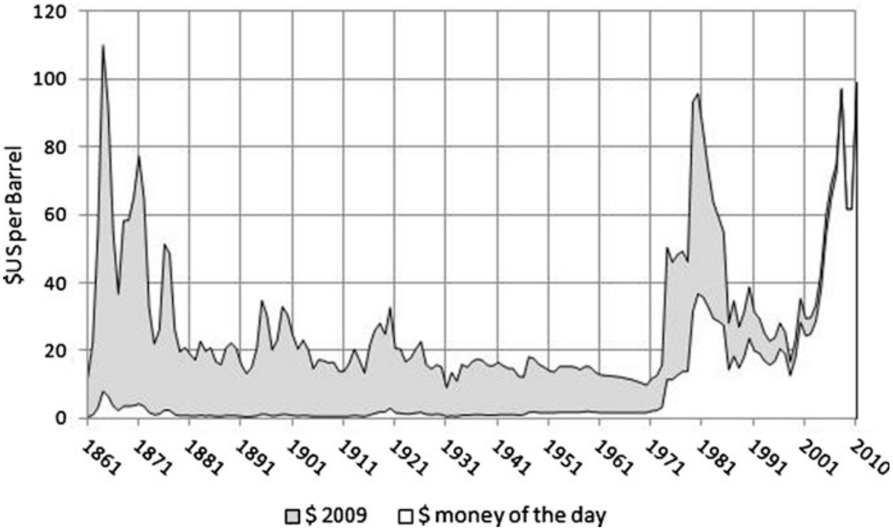


Fig. 18.3 Cost of petroleum since 1861. Based on data from [3]



be recovered as liquid condensate in natural gas processing plants

- *Dry gas* includes only traces of higher hydrocarbons
- *Sour gas* contains hydrogen sulfide, and *acid gas* contains carbon dioxide and/or hydrogen sulfide. Sour-gas processing plants coproduce elemental sulfur, which is used to make sulfuric acid and fertilizers

Some natural gas contains commercial quantities of inert gases—helium, neon, and/or argon. Almost all commercial helium comes from natural gas plants [6].

Liquids

Crude oil is the common name for liquid petroleum. People refer to crude oil as a single substance, but there are hundreds of different crudes with significantly different compositions. Crudes typically are named for their source country, reservoir, and/or some distinguishing physical or chemical property. Table 18.5 presents selected physical and chemical properties for ten crude oils.

In sweet crudes such as Tapis, the sulfur content is low. Sour crudes have more sulfur, which gives them a tart taste. (That's right: In the old days, prospectors did indeed characterize crude oil by tasting it.) Synthetic crude oil is produced from coal, kerogen, or natural bitumen. Processing costs are higher for conventional or synthetic crudes with high density and large amounts of sulfur, nitrogen, and trace contaminants. *Heavy oil* has a density greater than 1.0, so it sinks in water. The viscosities of heavy oil range from about 5,000 to 10,000 cP.

For the examples in Table 18.5, the sulfur content ranges from 0.14 wt.% to 5.3 wt.%, and the nitrogen content ranges from nil to 0.81 wt.%. The density (specific gravity) ranges from 0.798 to 1.014.

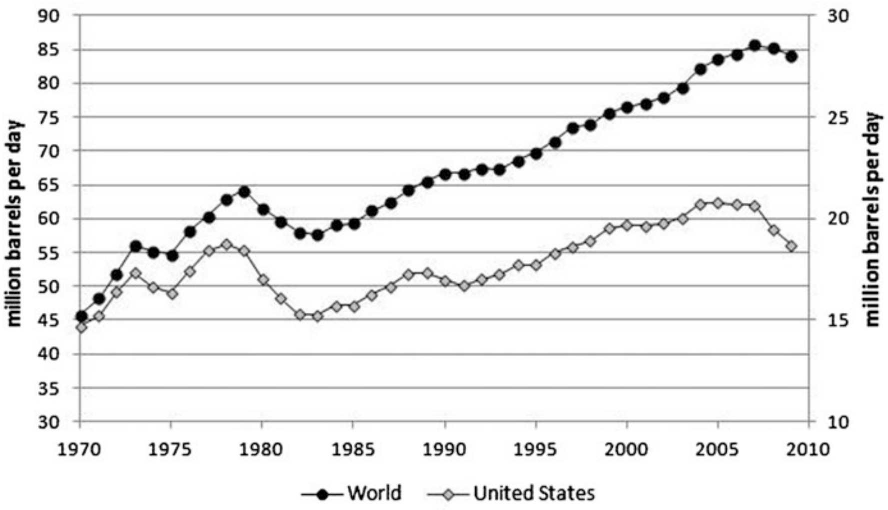


Fig. 18.4 Petroleum consumption since 1970 [3]

Table 18.4 Global changes by region in total energy consumption: 2009 vs. 2008

Region/country	Change 2009 vs. 2008 (%)	2009 Share of world total (%)
Total North America	−4.7	23.9
United States	−5.0	19.5
Total Europe/Eurasia	−6.0	24.8
Russian Federation	−6.4	5.7
Total Africa	−1.1	3.2
Total Asia Pacific	4.4	37.1
Japan	−8.0	4.2
China	8.7	19.5
India	6.6	4.2

Data selected from [3]

A commercial crude oil assay determines the chemical and physical properties of the whole crude and several distilled fractions. The fractions correspond to boiling ranges for common fuels. Using assay reports, refiners and oil traders determine the value of a crude by comparing its properties with refinery specifications, which include distillate properties and limits on trace contaminants. Limits on sulfur content can be set by sulfur-plant constraints or by the metallurgy of process equipment. Other common constraints include TAN (total acid number) and trace-metal content. Traders do purchase off-spec crudes, but usually for subsequent trading.

Table 18.6 shows an abbreviated template for a detailed assay report. Table 18.7 shows distillation yields for four common crudes. The naphtha content of Brent is twice as high as Ratawi, and its vacuum residue content is 60 % lower. Bonny Light yields the most middle distillate and the least amount of vacuum residue.

Table 18.5 Selected properties of ten crude oils [7]

Crude oil	API gravity[a]	Residue[b] (vol%)	Sulphur (wt%)	Nitrogen (wt%)
Alaska North Slope	27.1	53.7	1.2	0.2
Arabian Light	33.8	54.2	1.8	0.07
Arabian Heavy	28.0	46.6	2.8	0.15
Athabasca	8	50.8[c]	4.8	0.4
Brent (North Sea)	39	38.9	0.3	0.10
Boscan (Venezuela)	10.2	82.8	5.5	0.65
Kuwait export	31.4	49.5	2.6	0.21
Shengli (China)	24.7	72.5	1.0	
Tapis blend (Malaysia)	45.9	26.3	0.028	0.018
West Texas	40.2	36.4	0.3	0.08

[a] *API Gravity* is related to specific gravity by the formula: API = 141.5 ÷ (specific gravity @ 60°F)--131.5

[b] Unless otherwise stated, cut point = 343 °C–plus (650 °F–plus)

[c] Cutpoint = 525 °C (913 °F)

Bitumen

Like heavy oils, natural bitumens have densities greater than 1.0, but the viscosities are higher—more than >10,000 cP. Under ambient conditions, natural bitumen is a tar or a solid, but when heated it softens and flows. In practical terms, it is recovered as a solid but transported and processed as a liquid.

(It is important at this point to distinguish between natural bitumen and refined bitumens. The latter are specialty products with rather tight specifications. Refined bitumens are used primarily for paving and construction.)

Natural bitumen was probably the first form of petroleum used by man. In 3000 BCE, Mesopotamians were lining water canals, sealing boats, and building roads with natural

Table 18.6 Template for a detailed crude oil assay report

	Whole crude	Light naphtha	Medium naphtha	Heavy naphtha	Kero	AGO	LVGO	HVGO	VR	AR
True boiling point, °C	Initial	10	80	150	200	260	340	450	570	340
True boiling point, °C	Final	80	150	200	260	340	450	570	End	End
True boiling point, °F	Initial	55	175	300	400	500	650	850	1,050	650
True boiling point, °F	Final	175	300	400	500	650	850	1,050		
Yield of cut (wt% of Crude)		×	×	×	×	×	×	×	×	×
Yield of cut (vol% of crude)		×	×	×	×	×	×	×	×	×
Gravity, API	×	×	×	×	×	×	×	×	×	×
Specific gravity	×	×	×	×	×	×	×	×	×	×
Sulfur, wt%	×	×	×	×	×	×	×	×	×	×
Nitrogen, ppm	×		×	×	×	×	×	×	×	×
Viscosity @ 50 °C (122 °F), cST	×			×	×	×	×	×	×	×
Viscosity @ 135 °C (275 °F), cST	×			×	×	×	×	×	×	×
Freeze point, °C				×	×	×	×			
Freeze point, °F				×	×	×	×			
Pour point, °C	×			×	×	×	×	×	×	×
Pour point, °F	×			×	×	×	×	×	×	×
Smoke point, mm				×	×	×				
Aniline point, °C			×	×	×	×	×	×		
Aniline point, °F			×	×	×	×	×	×		
Cetane index, ASTM D976				×	×	×				
Diesel index			×	×	×	×	×	×		
Characterization factor (K)	×	×	×	×	×	×	×	×	×	×
Research octane number, Clear		×	×	×						
Motor octane number, Clear		×	×							
Paraffins, vol%		×	×	×	×	×	×			
Naphthenes, vol%		×	×	×	×	×	×	×		
Aromatics, vol%		×	×	×	×	×	×	×		
Heptane Asphaltenes, wt%	×								×	×
Micro carbon residue, wt%	×								×	×
Rams bottom carbon, wt%	×								×	×
Vanadium, ppm	×								×	×
Nickel, ppm	×								×	×
Iron, ppm	×								×	×

Table 18.7 Distillation yields for four selected crude oils [7]

Source field	Brent	Bonny Light	Green Canyon	Ratawi
Country	Norway	Nigeria	USA	Mid East
API gravity	38.3	35.4	30.1	24.6
Specific gravity	0.8333	0.8478	0.8752	0.9065
Sulfur, wt%	0.37	0.14	2.00	3.90
Yields, wt% feed				
Light ends	2.3	1.5	1.5	1.1
Light naphtha	6.3	3.9	2.8	2.8
Medium naphtha	14.4	14.4	8.5	8.0
Heavy naphtha	9.4	9.4	5.6	5.0
Kerosene	9.9	12.5	8.5	7.4
Atmospheric gas oil	15.1	21.6	14.1	10.6
Light VGO	17.6	20.7	18.3	17.2
Heavy VGO	12.7	10.5	14.6	15.0
Vacuum residue	12.3	5.5	26.1	32.9
Total naphtha	30.1	27.7	16.9	15.8
Total middle distillate	25.0	34.1	22.6	18.0
Naphtha plus distillate	55.1	61.8	39.5	33.8

bitumen. Egyptians were greasing chariot wheels with pitch and embalming mummies with asphalt [8, p. 2].

In tar sands, bitumen is associated with sand and clay, from which it can be recovered with hot water or steam. Tar sands (also known as oil sands) contain much of the world's recoverable oil. The largest deposits are in Alberta, Canada, where the proven reserves are 170 billion barrels; in addition, the province holds 1.4 billion barrels of conventional crude [9]. In the United States, tar sands are found primarily in Eastern Utah, mostly on public lands. These deposits contain 12–19 billion barrels of recoverable oil. In 2008, proven oil reserves in Venezuela totaled 87 billion barrels, mostly bitumen. In comparison, for that same year, Saudi Arabian reserves of conventional crude oil were 267 billion barrels [10].

Solids

Compared to liquids and gases, solids are harder to recover, transport, and refine. Liquids and gases can be pumped

Table 18.8 Typical composition of Green River oil shale [12]

Kerogen content: 15 wt.%[a]	
Kerogen composition, wt% of kerogen	
Carbon	80.5
Hydrogen	10.3
Nitrogen	2.4
Sulfur	1
Oxygen	5.8
Total	100
Minerals, wt% of mineral content	
Carbonates	48
Feldspars	21
Quartz	15
Clays	13
Analcite and pyrite	3
Total	100

[a] Equivalent to 25 gal oil per ton of rock

Table 18.9 US consumption of petroleum products: average for 1991–2011.

Product	Consumption (barrels/day)	Percent of total (%)
Gasoline	8,032	43.6
Jet fuel	1,576	8.6
Other middle distillates	3,440	18.7
Residual fuel oil	867	4.8
Other oils	4,501	24.4
Total consumption	18,416	100
Gasoline + jet + other middle distillates	13,048	70.8

Averages are based on data taken from [1]

through pipelines and into refineries with relative ease. Slurries of coal and water can be transported as liquids, but the water must be removed and eventually purified at considerable expense before the coal can be burned or gasified. Solid coal is consumed on a large scale to produce heat, steam, and electricity. These days, coal-powered transportation vehicles are rare. Coal-burning steam ships and railway locomotives are less efficient than their oil-powered counterparts. Typically, the specific energy of petroleum is 90 % greater than a ton of bituminous coal and 40 % greater than a ton of anthracite [11]. Even if for some reason a railroad or shipping company wanted to burn coal, doing so wouldn't be practical due to the present lack of coaling stations.

Kerogen is the solid organic matter in sedimentary rocks. Unlike bitumen, it doesn't flow even when heated. But at high-enough temperatures—e.g., 900 °F (480 °C)—it decomposes into gases, liquids, bitumen, and refractory coke. Huge amounts of kerogen are trapped in oil shale deposits. Fenton et al. [12] estimated that 1.3 trillion barrels of shale oil could be recovered from the world's oil shale reserves. Table 18.8 presents composition information on Green River oil shale from the western United States. About 91 % of the kerogen is hydrogen and carbon, but only 15 % of the sample is kerogen. Shale oil—synthetic crude from oil shale—tends to contain high amounts of arsenic, a severe poison for refinery catalysts. Usually, the arsenic is removed in existing hydrotreating units with special high-nickel chemisorption catalysts, which trap the arsenic by forming nickel arsenides.

Coal and coal technology are fully discussed in Chap. 19. Here we just say that coal is a black or brown combustible rock composed mostly of carbon, hydrocarbons, and ash. Generally, it is classified into four ranks—anthracite, bituminous, sub-bituminous, and lignite. Anthracite contains 86–97 % carbon and has a high heating value. Anthracite is relatively rare. Bituminous coal is far more common. It contains 45–86 % carbon and is burned to generate electricity. It is also used extensively in the steel and iron industries. Sub-bituminous coal contains 35–45 % carbon, and lignite contains 25–35 % carbon. Lignite is crumbly, has a high moisture content and a relatively low heating value.

Over the years, special circumstances have driven the large-scale conversion of coal into liquids, both directly and indirectly. Direct processes convert coal into various combinations of coal tar, oil, water vapor, gases, and char. The coal tar and oil can be refined into high-quality liquid fuels. Slurry-phase hydrocracking technology achieves very high direct conversion of coal into liquids. Several coal conversion processes are described in a recent book by Speight [13].

Developed in 1925, the Fischer-Tropsch (F-T) process is the main indirect route for converting coal into liquids. The coal is first gasified to make synthesis gas—a balanced mixture of CO and hydrogen. Over F-T catalysts, synthesis gas is converted into a full range of hydrocarbon products, including paraffins, alcohols, naphtha, gas oils, and synthetic crude oil. The F-T process was used extensively in Germany between 1934 and 1945. In South Africa, an improved version of the F-T process is used on a large scale to manufacture chemicals and fuels.

Synthesis gas is also derived from natural gas via steam-methane reforming. It can be converted into hydrogen and petrochemicals such as methanol. Worldwide, vast amounts of hydrogen are used to produce ammonia via the Haber-Bosch process.

Chemical Composition

The chemistry of the hydrocarbons in petroleum, bitumen, and kerogen is exceedingly complex. To one extent or another, they all contain water, inorganic salts, and dirt. After these are removed, most of the remaining compounds are hydrocarbons, some of which contain hetero atoms such

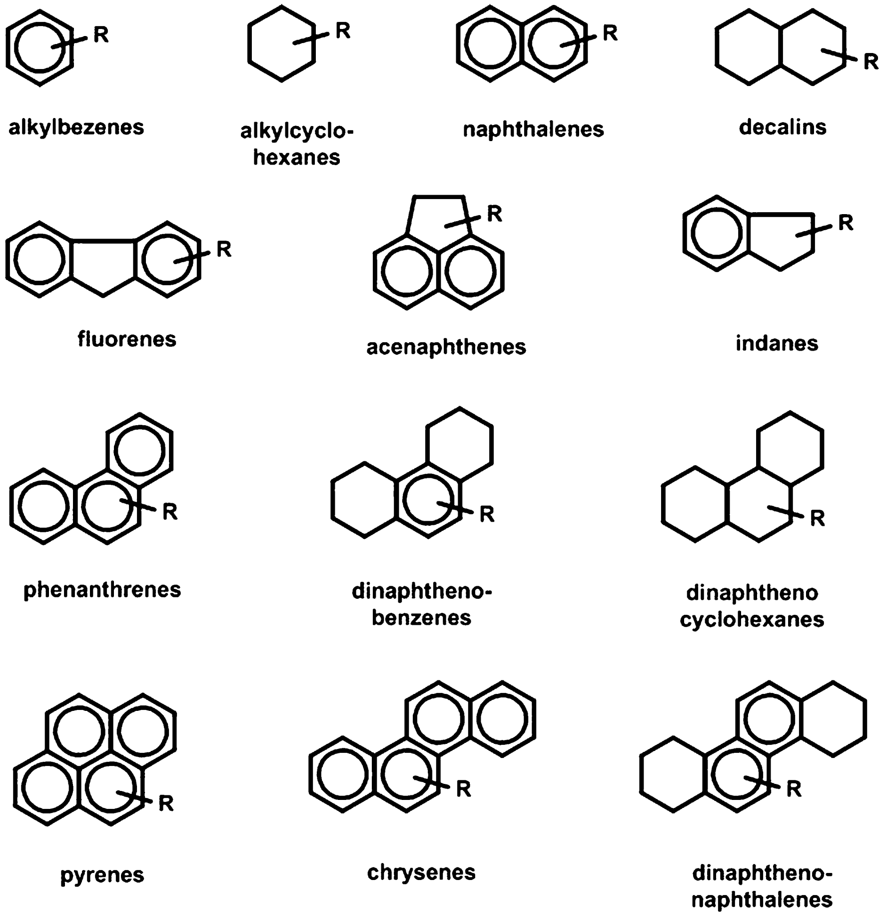


Fig. 18.5 Examples of hydrocarbon ring compounds

as sulfur, oxygen, and nitrogen, and trace metals. The non-hydrocarbon compounds can include different types of elemental sulfur. More than any other element, carbon binds to itself to form straight chains, branched chains, rings, and complex three-dimensional structures. The most complex molecules are biological—proteins, carbohydrates, fats, and nucleic acids. This is significant here because petroleum was formed from the remains of ancient microorganisms—primarily plankton and algae.

Figures 18.5, 18.6, and 18.7 show the complexity of petroleum. In addition to normal and isoparaffins (alkanes), it contains a large number of mono- and poly-ring hydrocarbons, to which paraffin chains can be attached (Fig. 18.5). Many compounds contain hetero atoms, especially sulfur, nitrogen, and/or oxygen (Fig. 18.6). Sulfur compounds also include aliphatic sulfides and disulfides.

Figure 18.7 shows plots of carbon number vs. atmospheric equivalent boiling point (AEBP) for different classes of compounds—aromatics, polynaphthenes, disulfides, benzothiophenes, etc. The indicated boiling points are for pure compounds. In mixtures, the boiling points are shifted by molecular interactions. Molecular interactions are greatest for polar compounds, such as those with hetero atoms. According to the Continuity Principle, physical and chemical properties for a given class of compound vary continuously with the number of carbon atoms in attached alkyl groups. Boiling points are crucial, because fractional distillation is the primary means by which petroleum is separated into useful products. Boiling ranges for typical cuts—naphtha, kerosene (including jet fuel), AGO (atmospheric gas oil, including diesel fuel), vacuum gas oils, and residue—appear at the top Fig. 18.7. In practice, due to imperfect separation in commercial distillation towers, the fractions overlap.

Curve A is for normal paraffins. Branched isomers with the same carbon number boil at lower temperatures. Fully saturated poly ring compounds fall on *Curve B*. Fully unsaturated poly aromatics fall on *Curve C*. Phenanthrene ($C_{14}H_{10}$), a three-ring poly aromatic compound, is found in the AGO boiling range. Adding seven hydrogen molecules (14 hydrogen atoms) converts phenanthrene into perhydrophenanthrene ($C_{14}H_{24}$), which ends up in the kerosene cut. *Curve D* shows how adding alkyl groups to pyrene makes makes *Curve D* parallel to *Curve A*. *Curves E, F*, and *G* represent compounds containing hetero atoms—sulfur, nitrogen, and oxygen.

Figure 18.8 also illustrates the Continuity Principle, in this case with a plot of density versus molecular weight for

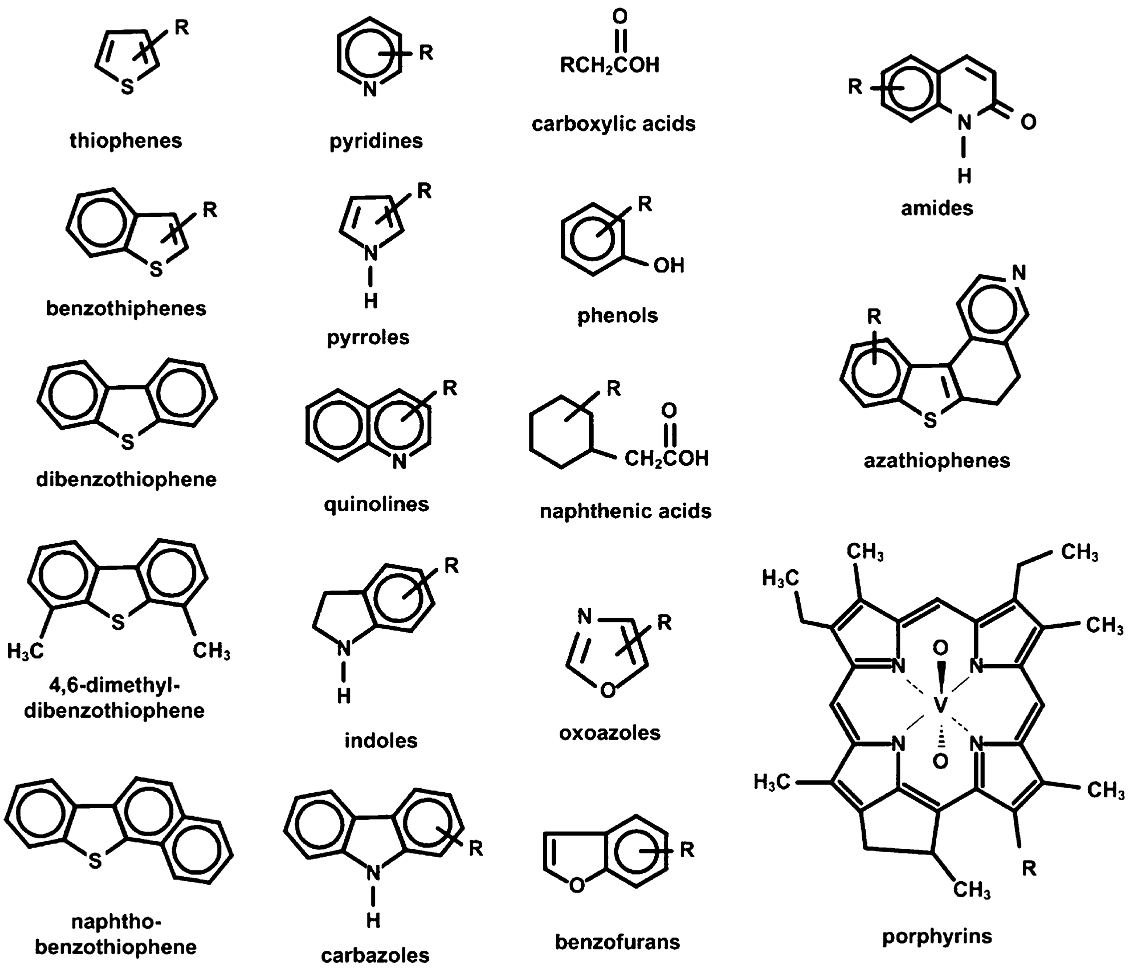


Fig. 18.6 Examples of hetero-atom compounds

five families of hydrocarbons. The "Z" number indicates hydrogen deficiency vs. paraffins. It comes from the subscript for hydrogen in the chemical formula C_nH_{2n+z}. For paraffins such as methane (CH_4) and octane (C_8H_{18}), $z = +2$, and the general formula is C_nH_{2n+2}. For benzene (C_6H_6) and alkylbenzenes, $z = -6$, and the general formula is C_nH_{n-6}.

Origin

Petroleum is complex due to its origin. According to Walters [17], eighteenth Century scientists concluded that coals were derived from plant remains. Mikhailo Lomonosov gets credit for suggesting that petroleum and bitumen were produced underground from coal at high pressure and temperature. In the mid-nineteenth Century, T.S. Hunt concluded that the organic matter in certain rocks comes from marine plants or animals.

Modern scientists believe petroleum came primarily from plankton and algae in the following way: the remains of dead organisms accumulated at the bottoms of ancient lakes or seas, along with sand, clay, salts, and other material. Over time, pressure, and heat transformed the deposits into sedimentary rock. Depending on the origin of the organic matter and the environment in which it aged, it became coal, kerogen, bitumen, petroleum, or natural gas.

Biomarkers. In 1936, Alfred Treibs linked the chlorophyll in plants to the porphyrins in petroleum (Fig. 18.9). Even after millions of years, porphyrins and other biomarkers retain much of their original structure. Biomarkers can be found in crude oil, rocks, sediments, and soil extracts.

Exploration

A book by Conaway [18] presents a good overview of how we look for and produce petroleum. Commercially significant petroleum systems include the following:

- Organic-rich source rock in which oil and gas were formed

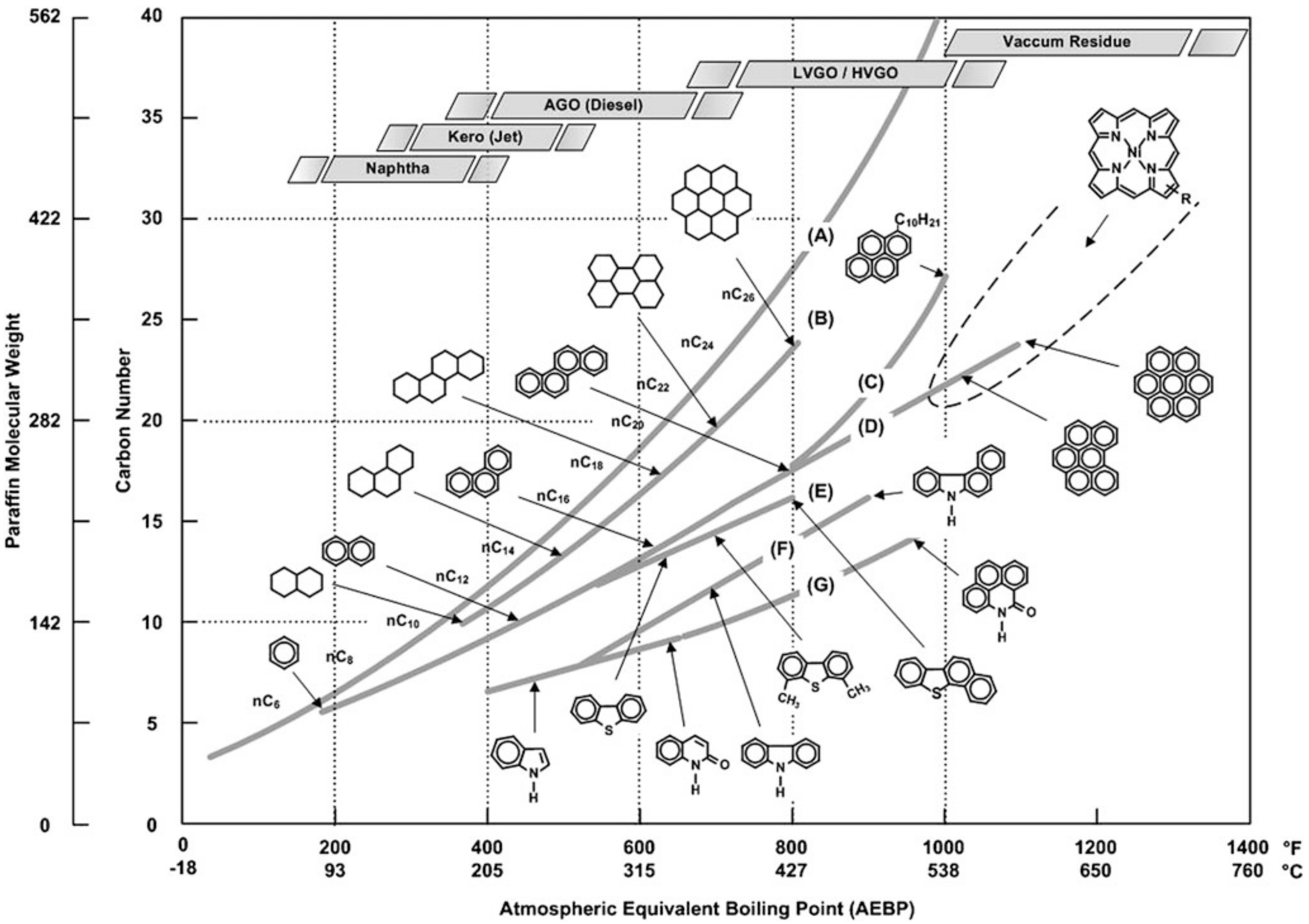


Fig. 18.7 Carbon number and atmospheric equivalent boiling points for different compounds [14, 15]. Shown at the *top* of the figure are the straight-run products that correspond to AEBPs of the compounds. Note how the saturation of polyaromatics such as chrysense (4 rings), phenanthrene (3 rings), and naphthalene (2 rings) can move them into different distillation fractions

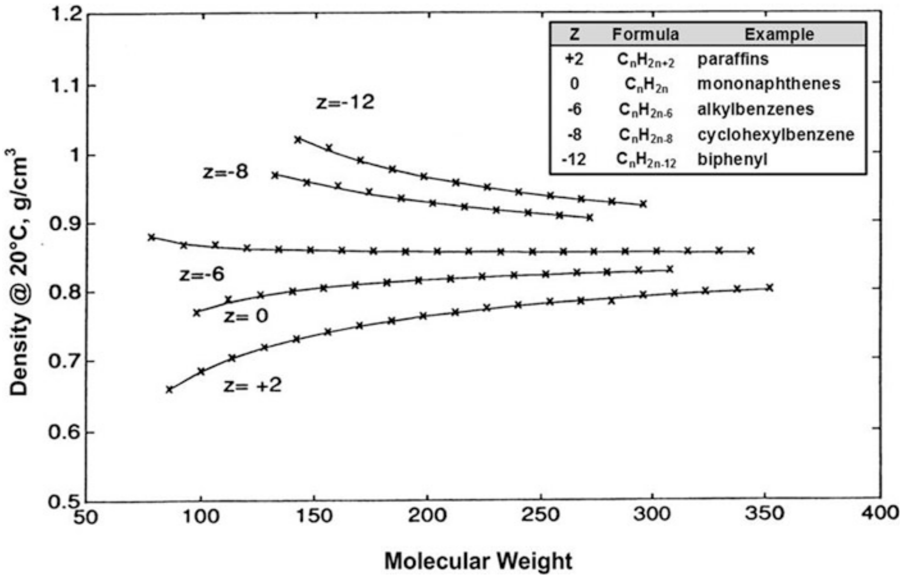


Fig. 18.8 Density vs. molecular weight for hydrocarbon families [16]

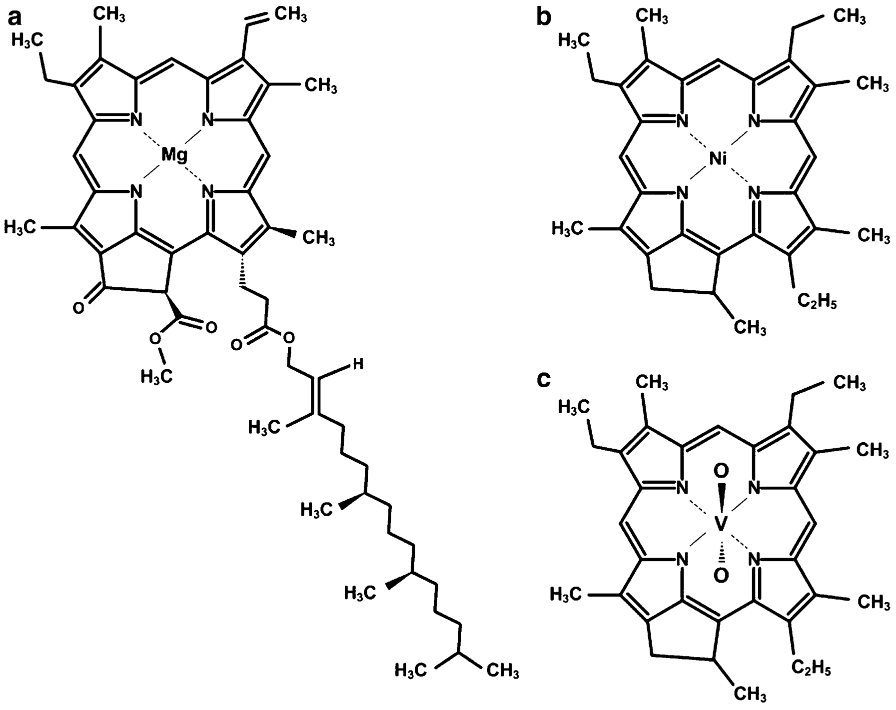


Fig. 18.9 Biomarkers: a comparison of Chlorophyll A with porphyrins found in petroleum

- Pathways comprised of permeable rock such as sandstone and limestone, which allow oil and gas to migrate. Often, the movement of oil and gas is promoted by the flow of underground water
- Porous, permeable reservoir rock capped by impermeable rock, which prevents fluids from migrating to the surface. Shale and salt domes are common caps

Searching for traps. When petroleum geologists look for oil and gas, they search for geological structures that might serves as traps. In so doing, they rely heavily on reflection seismology. During seismic exploration, explosives, or thumper devices send sound waves through the earth. Reflected sound waves are measured with hydrophones (in water) and geophones (on land). Different layers reflect sound in different ways. With the help of sophisticated software, geophysicists transform seismic data into 3-dimensional maps that show the structure of subsurface rock formations.

For petroleum, there are four major kinds of reservoir traps, three of which are illustrated in Fig. 18.10.

- Anticline traps are, in essence, inverted bowls. They can be symmetrical or asymmetrical. A steep anticline might be called a dome. Anticline traps hold most of the world's conventional crude oil
- Faults are formed at the boundaries of cracks in the earth's crust. The four major kinds of faults are thrust, lateral, normal, and reverse. Thrust and lateral faults are created by horizontal movement The San Andreas Fault in California is a well-known example of a lateral fault. To be more specific, the San Andreas is a right-lateral, strike-slip fault. Normal and reverse faults are created by vertical movement and are more likely to create fault traps for petroleum
- A salt-dome trap is created when a mass of underground salt is pushed up by underground pressure into a dome. The salt dome breaks through layers of rock and pushes them aside as it rises. The salt is impermeable, and if it abuts porous rock, it can serve as a reservoir cap
- In stratigraphic traps, the reservoir is capped by another reservoir or by layers of rock with lower porosity or permeability

Role of biomarkers in exploration. Biomarkers can indicate the relative amounts of oil and gas in a source rock, the age of a source rock, the environment in which organic matter was deposited, the maturity of a source rock, etc. Often, biomarker data are included in reservoir models of prospective oil fields.

Drilling and Production

Drilling. Key components of a drilling rig are shown in Fig. 18.11. To drill a well, a bit is attached to the end of a drill string, which is comprised of sections of steel pipe that are 30-ft (9 m) long. The drill string is lengthened as the well gets deeper by attaching additional sections of pipe to the top. Drill collars are thick-walled sections of drill pipe at the

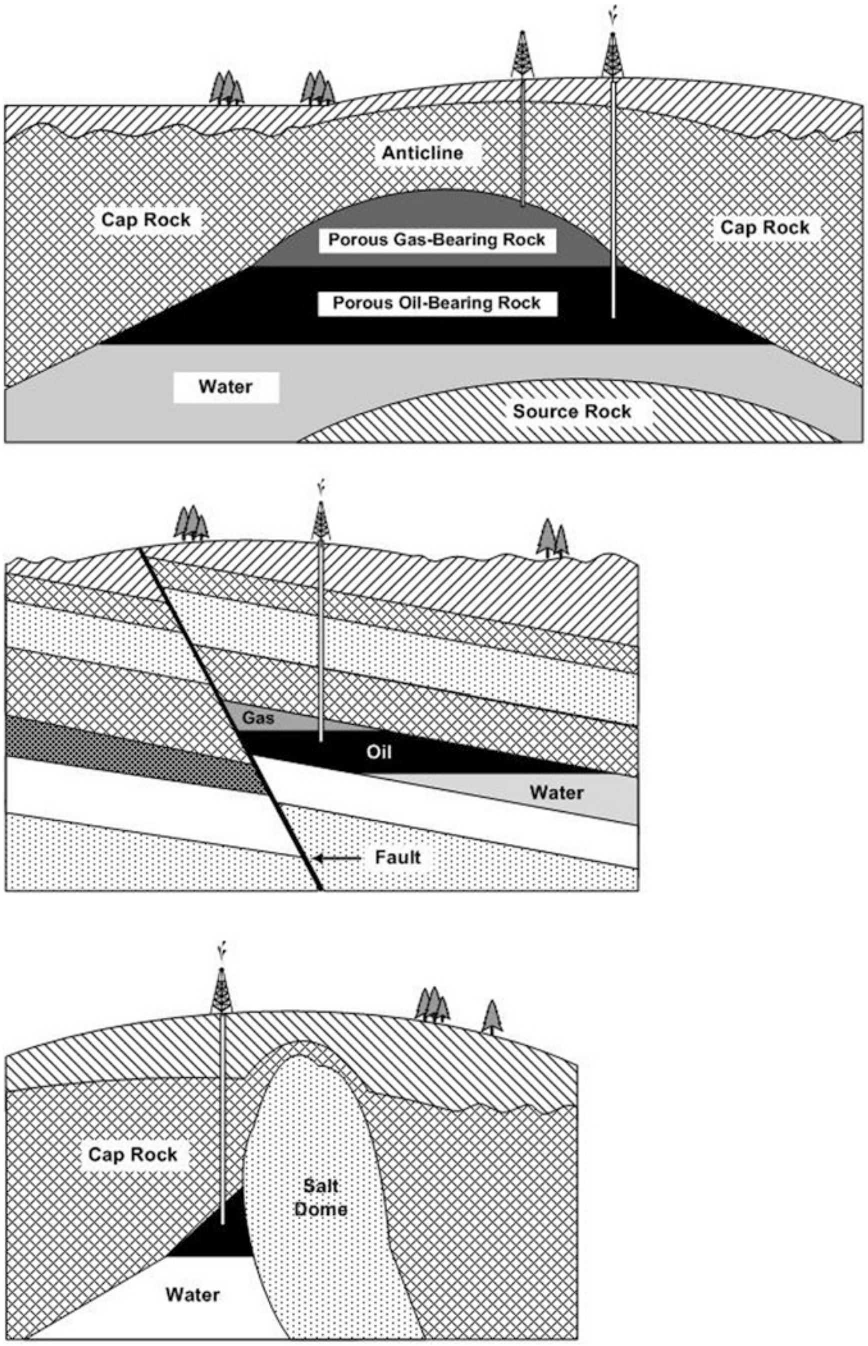


Fig. 18.10 Three types of hydrocarbon traps—anticline, fault, and salt dome. As with oil and gas, the water layers are not pools, but water-saturated porous rock

bottom of the drill string; collars apply extra weight to the bit. As the bit cuts through rock, mud is pumped down through the inside of the drill string, out the bit, and up through the space between the drill string and the well bore. Despite its dirty name, drilling mud is a designed blend of fluids, solids, and chemicals. It cools the drill bit and brings rock cuttings back to the surface. The cuttings are removed in shakers, and the fluid goes back to the mud pit, from which it is recycled.

The derrick must be tall enough to accommodate a 90-ft-long "triple" comprised of three sections of pipe. The derrick, pulleys, and hawser must be robust enough to support the lifting and manipulation of the entire drill string. The kelly is the top joint of a drill string. It has flat sides that fit inside a bushing on the rotary table, which turns the drill string and bit. Note that not all drilling rigs use a kelly system.

A surface blowout is a sudden increase in well pressure resulting in loss of containment and uncontrolled flow of oil and gas into the atmosphere. The blowout preventer (BOP) is a safety valve or other device beneath the floor of the drilling rig. When activated, it stops a blowout by sealing off the top of the well.

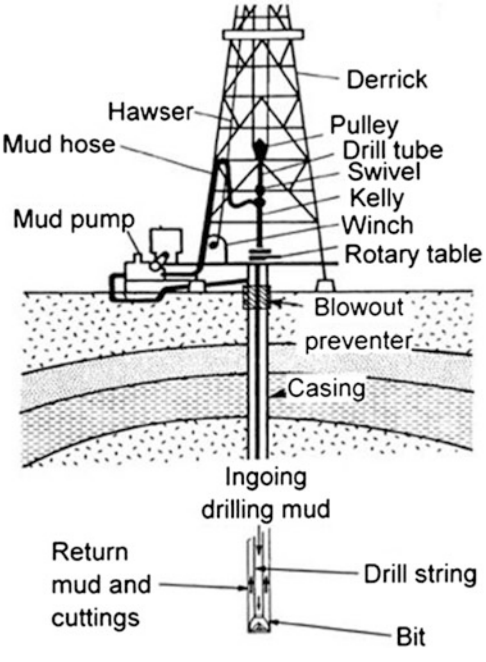


Fig. 18.11 Components of a typical drilling operation [19]

The density of drilling mud is crucial. If the mud density is too low, a well is susceptible to a surface blowout. If the mud density is too high, it can cause an underground blowout—the rupture of the reservoir underground—pushing drilling mud into another formation. Underground blowouts are the most common of all well control problems. Many surface blowouts begin as underground blowouts. Prompt, correct reaction to an underground event can prevent a dangerous and costly surface blowout [20].

Completion. Wells are completed by casing the well bore with steel pipe and cementing the casing into place. Casing prevents the well from collapsing. The outside diameters of casing pipe range from 4.5 to 16 in. (114–406 mm). Cementing is a key step. In a good cement job, the entire space between the casing and the well bore is filled with cement. Centralizers keep the casing from resting against the well bore, which could block the flow of cement, resulting in a poor cement job and increasing the risk of a blowout.

After cementing, perforation guns punch small holes through the casing into the reservoir rock, providing a path for the flow of oil and gas into the well. In open-hole completion, the last section of the well is uncased. Instead, the installation of a gravel pack stabilizes the casing and allows fluids to enter the well at the bottom. After perforation, special acid-containing fluids are pumped into the well to increase porosity and stimulate production. Usually, a smaller diameter tube is inserted into the casing above the production zone and packed into place. This provides an additional barrier to hydrocarbon leaks, raises the velocity at which oil flows under a given pressure, and shields the outer casing from corrosive well fluids.

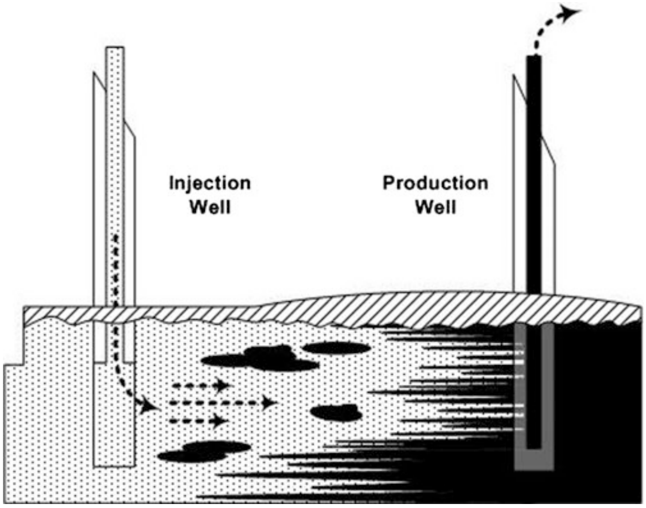


Fig. 18.12 Injection and production wells

Production. Liquid petroleum may flow on its own due to underground pressure, or it may be forced to the surface by the injection of treated water, natural gas, or CO_2. Figure 18.12 illustrates using an injection well to stimulate production.

Heavy oil and bitumen are recovered in several ways. They can be dug out with conventional mining techniques, or they can be liquefied by the injection of high-pressure steam. In so-called “huff and puff” operations, production is stopped, either naturally or on purpose. Hot steam is injected into the reservoir and left to soak, usually for a few days. During the soak, heat flows from the steam into the oil and surrounding rock, and the reservoir pressure increases. After the soak, the well is reopened and oil production resumes, at first by natural flow and then by artificial lift. Production decreases over time as the temperature and pressure of the reservoir go down. When the flow of oil drops below a certain limit, the process is repeated.

At surface facilities, notably in Venezuela and Canada, recovered bitumen is diluted with lighter hydrocarbons. The resulting “Dilbit” (diluted bitumen) flows under ambient conditions, so it can be transported conventionally in pipelines and oil tankers.

Kerogen is recovered from oil shale by several methods. From 1985 to 1990, Unocal recovered some 4.6 million barrels of synthetic crude oil from oil shale in a complex mining and upgrading venture at Parachute Creek, Colorado [21]. The plant yielded roughly 40 gal of oil per ton of rock. In the vertical-shaft retort, crushed kerogen-containing shale was pumped up from the bottom of the retort vessel. Hot

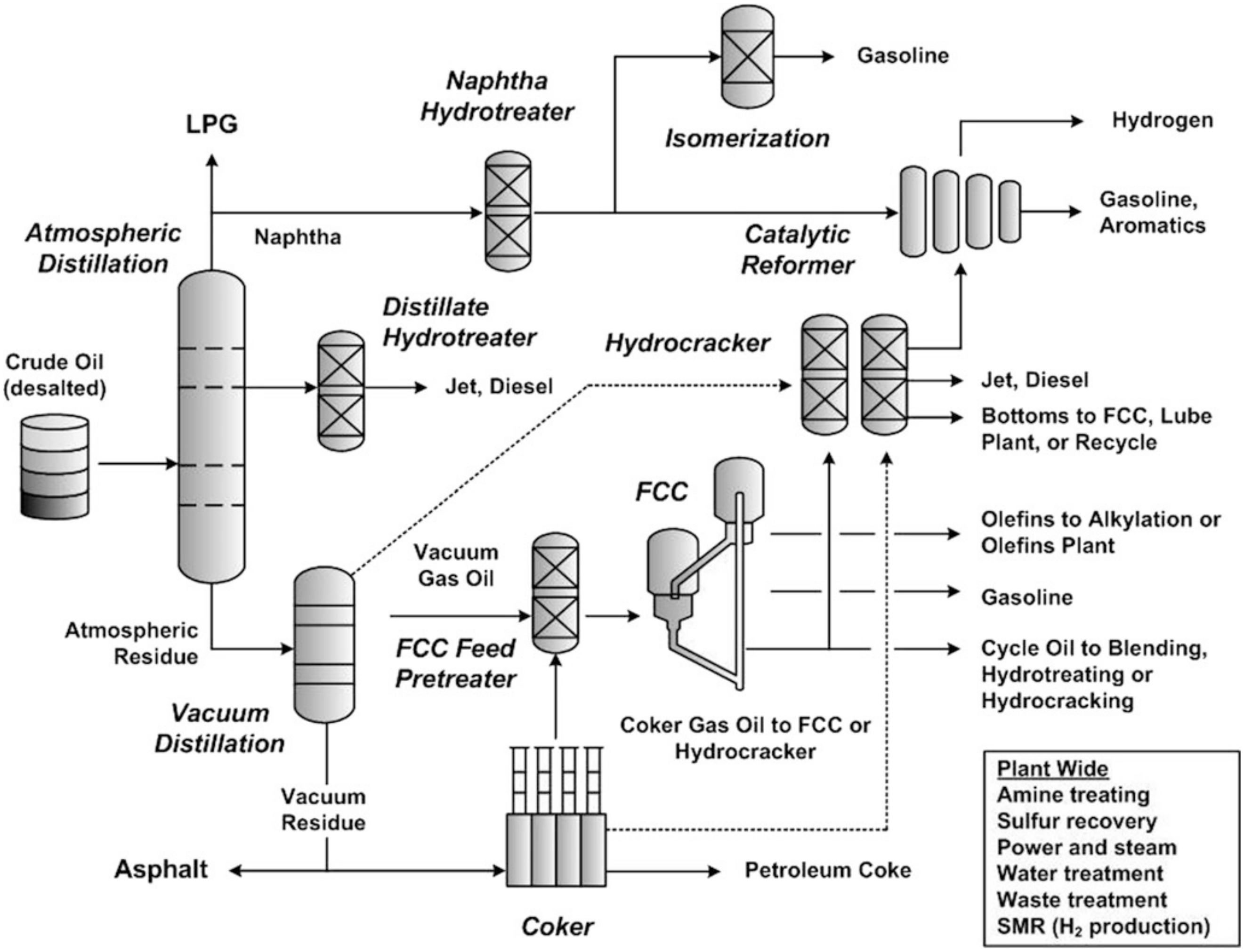


Fig. 18.13 Schematic diagram of a high-conversion petroleum refinery [15, Fig. 10.1]

recycle gas flowed counter-currently downward, decomposing the rock and releasing hydrocarbons. Condensed shale oil was removed from the retort at the bottom. Part of the hot gas was recycled. The rest either was used to produce heat and hydrogen, or recovered as product. The spent shale was removed from the top of the retort, cooled, and stored in pits or returned to the mine. In the reducing environment of the retort, sulfur and nitrogen were converted to H_2S and NH_3, which were recovered from product gases by conventional means. The plant yielded high-quality synthetic crude oil suitable for further refining in conventional facilities.

Other oil shale processes involve partial combustion, either underground, at the surface, or in shafts drilled horizontally into kerogen-rich formations. From 1972 to 1991, Occidental Petroleum developed an in situ process, in which explosives were used to create underground chambers of fractured oil shale. The oil shale was ignited with external fuel, and air and steam were injected to control combustion. The hot rock fractured and released shale oil, which was pumped to the surface from a separation sump and collecting well.

Bitumens derived from oil shale and many tar sands contain small-but-significant amounts of arsenic, which are severe poisons for catalysts in refineries.

Refining

Refining petroleum is like cooking crab: you boil it, crack it, and pull out the bits. In a refinery, a given molecule might be boiled and recondensed hundreds of times. Cracking big molecules into small ones is the key to economic success. Pulling out the bits—the removal and disposition of contaminants—improves the cracking processes and protects the environment.

Figure 18.13 shows a simplified schematic diagram for a “typical” North American refinery. The word “typical” is inside quotation marks, because no two refineries are identical. European and Asian facilities include different process units because they make less gasoline and more middle distillates. They are less likely to employ delayed coking and more likely to use visbreaking for upgrading residue. Plants that produce substantial amounts of lubricants and/or petrochemical monomers include additional sections.

A petroleum refinery transforms the complex tangle of molecules in crude oil into a wide array of products, all of which must meet tight specifications. Listed below are some basic refinery operations.

- Separation
  - Distillation
  - Solvent processes
- Hydrotreating
- Conversion
  - Thermal cracking
  - Thermal hydrocracking
  - Catalytic cracking
  - Catalytic hydrocracking
- Catalytic reforming and isomerization
- Recombination of light intermediates
  - Alkylation
  - Polymerization
- Product treating and blending
- Protecting the environment
  - Gas treating
  - Sulfur recovery
  - SO_x and NO_x abatement
  - Waste treatment
- Related operations
  - Asphalt production
  - Lubricant manufacturing
  - Hydrogen production
  - Hydrogen recovery and purification
  - Production of petrochemical monomers

The Need for Conversion

The demand for petroleum remained relatively flat throughout the last half of the nineteenth Century. In the 1860s the most valuable fraction distilled from crude oil was kerosene. In 1878, thanks to the inventions of Thomas Edison, electric lighting slowly but surely began to displace kerosene as a preferred illuminant. But the advent of the gasoline-powered automobile dramatically increased the demand for petroleum. In addition to increasing overall crude oil demand, the automobile increased demand for naphtha, from which gasoline is derived.

During 1991–2011, the United States consumed, on average, 70 % of its petroleum as gasoline and middle distillates. But as shown in Table 18.7, yields of naphtha plus middle distillates for four selected (relatively common) crudes range from 33.8 to 61.8 %. This leaves sizeable supply/demand gaps, which create a sizeable light-product/heavy-product price differential. A common measure of this differential is the "heating oil crack spread." Since 2010, crack spreads ranged between \$6 and \$25 per barrel. In January 2011, they exceeded \$24 per barrel [22]. At that differential, for a refinery that processes 200,000 barrels of oil per day, 30 % conversion is worth about \$1.4 million per day—about \$500 million per year.

In response to such huge incentives, refiners, and licensors have developed an array of processes for converting heavy products into light ones. Many of these are described below. But before discussing the hardware, let's look at the chemistry around which the hardware is built.

Petroleum Refining Chemistry

"Cracking" and "pulling out the bits" involve the chemical reactions listed below.

- Hydrotreating and saturation
  - Saturation of aromatics (ASAT) and olefins (OSAT)
  - Hydrodesulfurization (HDS)
  - Hydrodenitrogation (HDN)
  - Hydrodeoxygenation (HDO)
  - Hydrometalation (HDM)
- Cracking (making little ones out of big ones)
  - Thermal cracking
  - Catalytic cracking
  - Hydrocracking (catalytic and thermal)
- Dehydrogenation, isomerization (no change in carbon number)
- Alkylation, polymerization (making big ones out of little ones)
- Acid/base processes: amine treating, caustic scrubbing
- Mercaptan oxidation (Merox)

Hydrotreating reactions: saturation of olefins. On a macro scale, the saturation of olefins is rapid and essentially irreversible. It releases considerable heat.

Hydrotreating reactions: saturation and condensation of aromatics and polyaromatics. Figure 18.14 provides examples of the reversible saturation of aromatics and polyaromatics. Figure 18.15 [8] summarizes thermodynamic calculations for the competition between the saturation of naphthalene and the condensation of naphthalene with *o*-xylene to form chrysene. At high pressures and low temperatures, equilibrium favors saturation. At low pressures and high temperatures, equilibrium favors dehydrogenation. At high-enough temperatures, equilibrium favors condensation (Fig. 18.16).

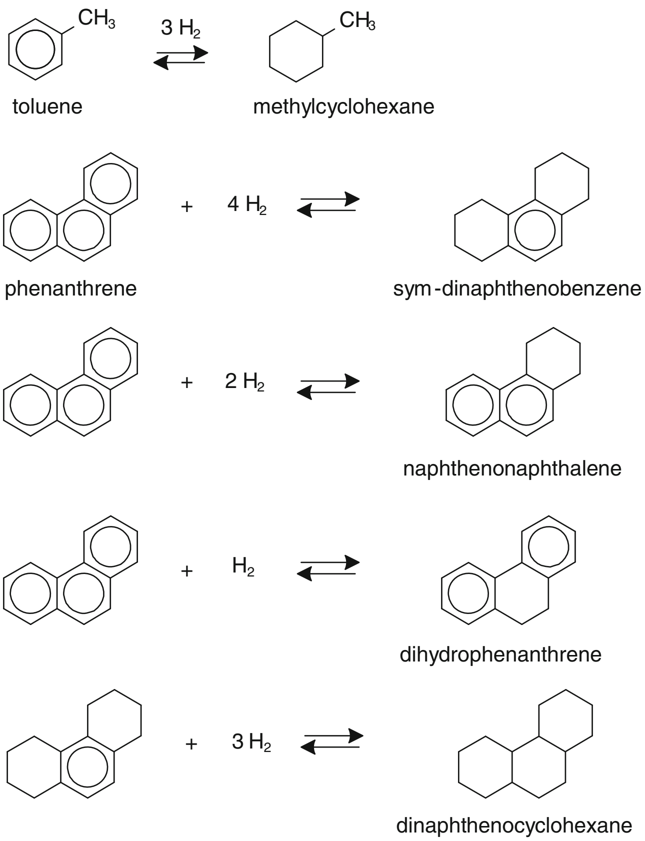


Fig. 18.14 Saturation of aromatics and polyaromatics

Figure 18.6 shows the so-called zig-zag mechanism for the production of large polyaromatics by adding 2-carbon and 4-carbon species to smaller polyaromatics. The condensation of large polyaromatics via the Scholl reaction can lead eventually to coke formation [25].

Figure 18.17 shows a mechanism for the one-at-time build up of rings on a nucleus of coke. The mechanism includes the following steps, all of which are to some extent reversible:

- Hydrogen abstraction by gas-phase radicals
- Addition of radicals to the coke surface
- Reaction of a gas-phase olefin with a radical on the surface
- Reaction of a gas-phase radical with an olefin group on the surface
- Cyclization
- Dehydrogenation

Hydrotreating and hydrocracking units operate in the crossover region between 315 and 425 °C (600 and 800 °F). Per Fig. 18.15, at temperatures below the crossover region, saturation is favored. Temperatures above crossover favor aromatics.

Hydrotreating reactions: HDS. Figure 18.18 shows representative hydrodesulfurization (HDS) reactions. Note that for sulfur removal from the first three reactants, the mechanism is direct. That is, the sulfur-containing molecule interacts with an active site on the catalyst, which removes the sulfur atom and replaces it with two hydrogen atoms. Additional hydrogen converts the sulfur atom into H_2S, which desorbs from the catalyst, leaving behind a regenerated active site.

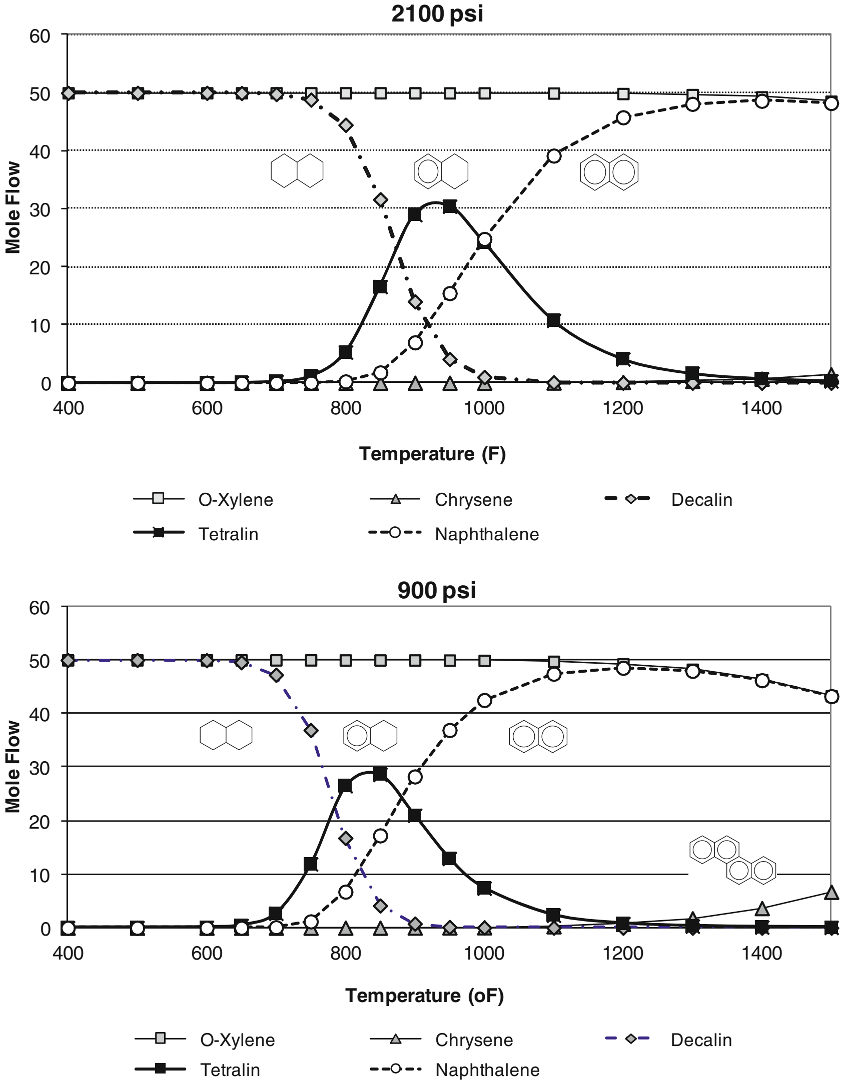
For the fourth reactant—4,6-dimethyldibenzothiophene (4,6-DMDBT)—the reaction proceeds via both a direct and an indirect route. The 4,6-DMDBT molecule is planar, and the two methyl groups are adjacent to the sulfur atom. This protects the sulfur atom from active sites on the catalyst, thereby inhibiting direct HDS. In the indirect route, saturating one of the aromatic rings that flank the sulfur atom converts the planar structure into a puckered arrangement with tetrahedral C–C bonds. This puckering rotates one of the inhibiting methyl groups away from the sulfur atom, giving it better access to the catalyst.

Making ultra-low-sulfur diesel (ULSD), in which the sulfur content is less than 10–15 wppm, requires severe hydrotreating, after which the only remaining sulfur compounds are the above-mentioned 4,6-DMDBT and other di- and trimethyl dibenzothiophenes. Because the removal of sulfur from these compounds is more facile after prior saturation, the crossover phenomenon affects the production of ULSD significantly, so much so that it governs the design of commercial units.

Hydrotreating reactions: HDN. Figures 18.19 and 18.20 show representative HDN reactions. Figure 18.19 illustrates the mechanism for the HDN of quinoline. As with sulphur removal from hindered DMDBTs, the crossover phenomenon is important for deep HDN, because nitrogen removal requires prior saturation of an aromatic ring adjacent to the nitrogen atom.

Hydrotreating reactions: HDO. In feeds to industrial hydroprocessing units, oxygen is contained in organic acids, ethers, peroxides, and other compounds. Some form by reacting with air during transportation and storage of crude oil, distillates, and cracked stocks generated by delayed coking and Fluid catalytic cracking (FCC) units (see below). Hydrodeoxygenation (HDO) proceeds rapidly, so rapidly that it can cause problems with excessive heat release. Also, oxygen compounds can form gums and polymers, which inhibit flow and increase pressure drop.

Hydrotreating reactions: HDM. Metals such as nickel and vanadium are present in high-boiling fractions, mostly in asphaltenes. Asphaltenes are mixtures of waxy solids with porphyrins (Fig. 18.9). Corrosion generates soluble iron, and entrained salt brings in alkali and alkaline earth metals,



°C	°F
204	400
260	500
316	600
371	700
427	800
482	900
538	1000
593	1100
649	1200
704	1300
760	1400

Fig. 18.15 Thermodynamic calculations illustrating the competition between the saturation and the condensation of polyaromatics [23]. Data for the graphs were generated by Aspen Plus for a six component system comprising naphthalene ($C_{10}H_8$), tetralin ($C_{10}H_{12}$), decalin ($C_{10}H_{18}$), *o*-xylene (C_8H_{10}), chrysene ($C_{18}H_{12}$), and hydrogen (not shown)

primarily sodium and calcium. Synthetic crudes from oil sands and oil shale tend to contain small-but-significant amounts of arsenic. Silicones are added to crude oil to facilitate flow though pipelines. Silicones are also used to control foaming in delayed coking units.

All of these metals poison refinery catalysts. Refiners remove them by hydrotreating over special wide-pore guard catalysts. Guard materials remove some contaminants by chemisorption. Other contaminants, such and Ni, V, and soluble Fe, are removed by HDM reactions. HDM converts the metals into sulfides, which adhere to the guard material.

Conversion. Table 18.10 illustrates a fundamental aspect of conversion. For a given class of hydrocarbons, "lighter" means a lower molecular weight, lower boiling point, lower density, and a higher hydrogen-to-carbon ratio (H/C). Methane, the lightest hydrocarbon, has an H/C ratio of 4.0. Benzopyrene has an H/C ratio of 0.6. The H/C ratio of common crude oils ranges from 1.5 to 2.0, while the H/C ratio for asphaltenes is about 1.15.

Some conversion processes—FCC, thermal cracking, and deasphalting—increase the H/C ratio by rejecting carbon. Hydrotreating and hydrocracking increase the H/C ratio by adding hydrogen. In this context, "rejecting carbon" does not mean that a few carbon atoms are removed from every large molecule. Rather, it means that heavy molecules are split ("cracked") into lighter molecules with higher H/C ratios and heavier molecules with lower H/C ratios.

Conversion: definition. Refiners define conversion as a change in boiling point. The goal of a conversion process is to crack material that boils above a particular cutpoint into

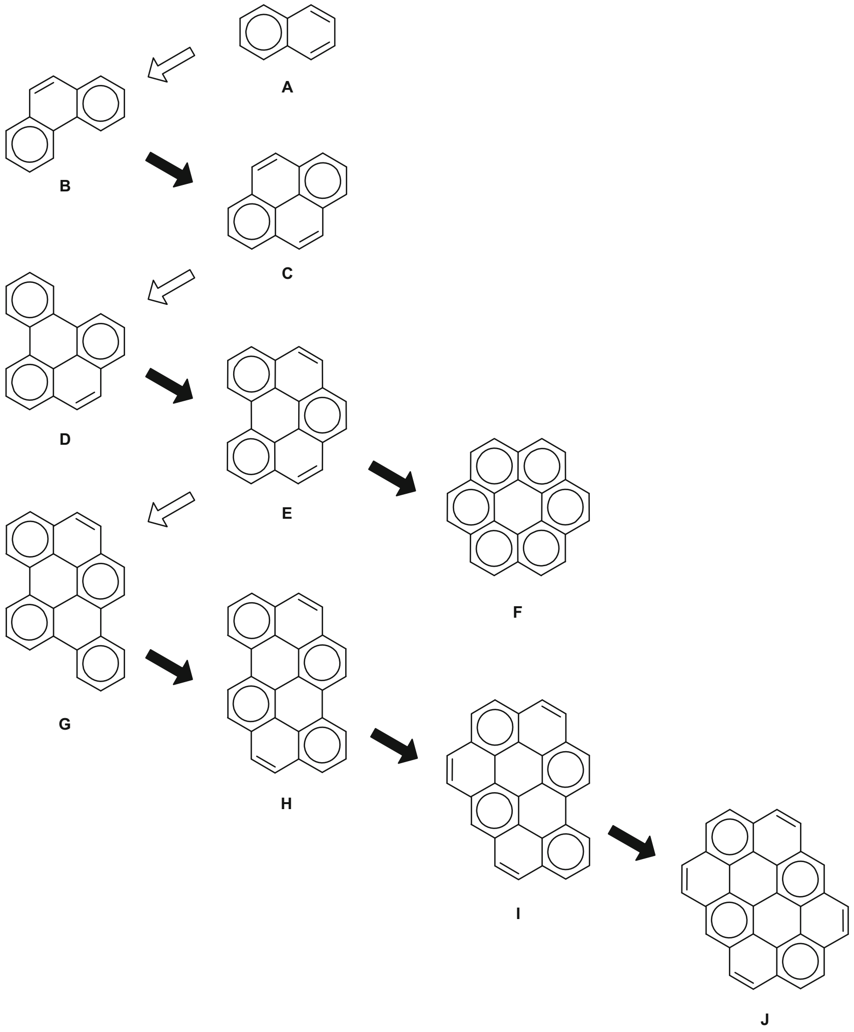


Fig. 18.16 Zig-zag mechanism for the condensation of polyaromatics by sequential addition of 2-carbon and 4-carbon units [24]. The isomers shown are (**a**) naphthalene, $C_{10}H_8$; (**b**) phenanthrene, $C_{14}H_{10}$; (**c**) pyrene, $C_{16}H_{10}$; (**d**) benzo[*e*]pyrene, $C_{20}H_{12}$; (**e**) benzo[*ghi*] perylene, $C_{22}H_{12}$; (**f**) coronene, $C_{24}H_{12}$; (**g**) dibenzo*[b,pqr]*perylene, $C_{26}H_{14}$; (**h**) benzo(pqr)naphtho(8,1,2-bcd)perylene, $C_{28}H_{14}$; (I) naphtho[2′.8′,2.4]coronene, $C_{30}H_{14}$; and (**j**) ovalene, $C_{32}H_{14}$. Note how the H/C ratio goes down as condensation increases, from 0.8 for naphthalene to 0.4375 for ovalene

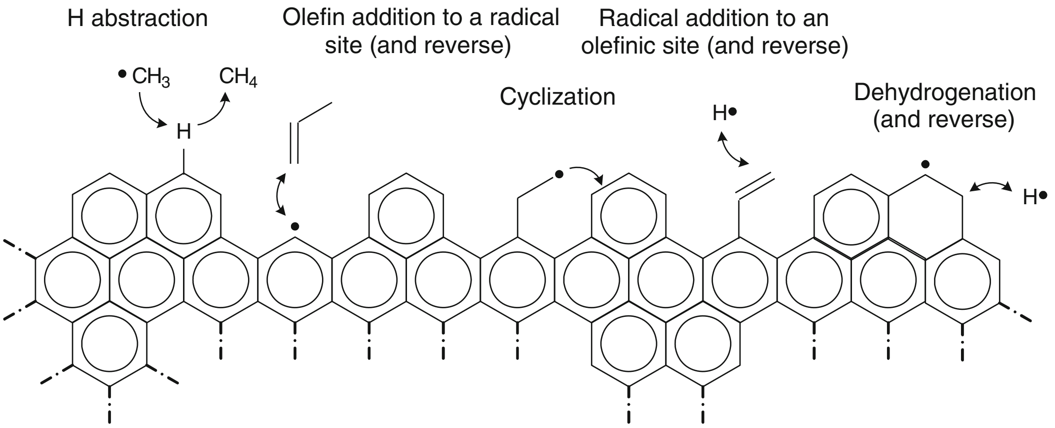


Fig. 18.17 Mechanism for the addition of rings to a nucleus of coke [26]

Fig. 18.18 Representative hydrodesulfurization (HDS) reactions

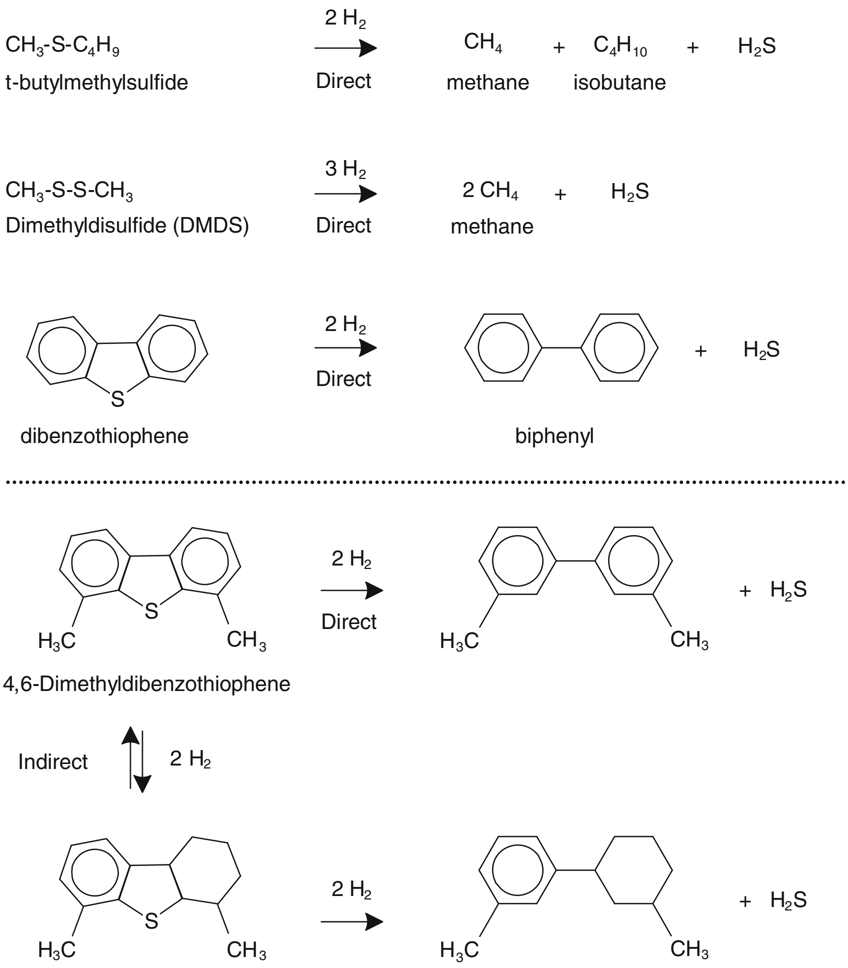
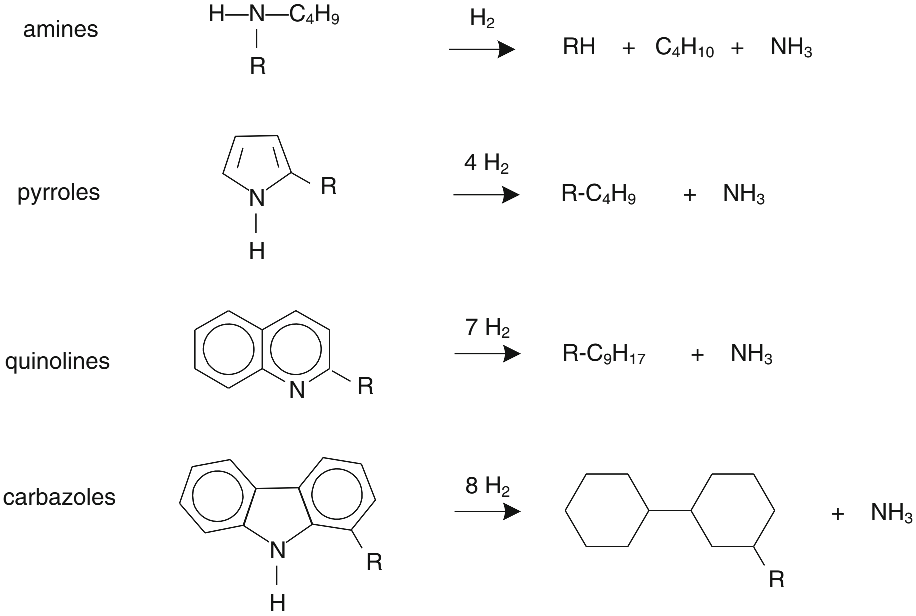


Fig. 18.19 Representative hydrodenitrogenation (HDN) reactions



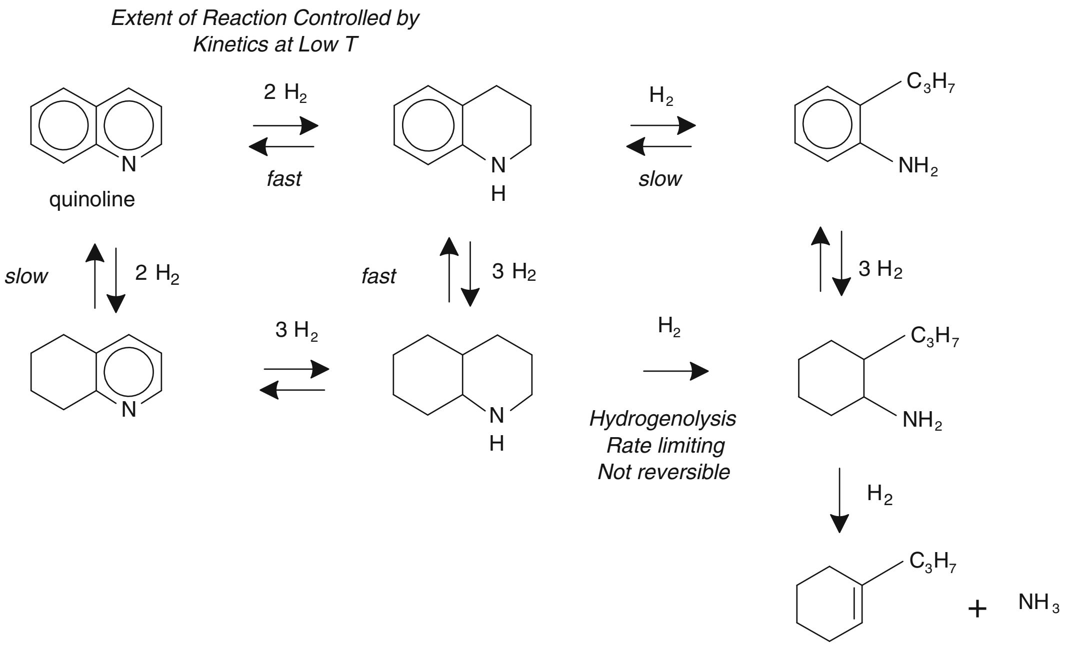


Fig. 18.20 Mechanism for the HDN of quinoline [27]

Table 18.10 Molecular weight, H/C, and boiling points for selected hydrocarbons

				Boiling point	
Compound	Molecular weight	Formula	H/C	°C	°F
Paraffins					
Methane	16.04	CH_4	4.0	−164	−263.2
Ethane	30.07	C_2H_6	3.0	−88.6	−127.5
Propane	44.10	C_3H_8	2.67	−42.1	−43.7
Butane (iso)	58.12	C_4H_{10}	2.50	−6.9	19.6
Octane (iso)	114.23	C_8H_{18}	2.25	99.2	210.6
Cetane (n)	226.44	$C_{16}H_{34}$	2.13	287	548.6
Aromatics					
Benzene	78.11	C_6H_6	1.0	80.1	176.2
Naphthalene	128.17	$C_{10}H_8$	0.8	218	424.4
Benzopyrene	252.32	$C_{20}H_{12}$	0.6	–	–

material that boils below that particular cutpoint. A typical cutpoint for naphtha-oriented conversion is 400 °F (204 °C), while a typical cutpoint for diesel-oriented conversion is 650 °F (343 °C). The equation for conversion is simple:

$$\text{Conversion} = P/\text{FF} * 100\%$$

where P = the amount of material in the product that boils below the chosen cutpoint and FF = the total amount of fresh feed. In "true conversion" calculations, the amount of P in the FF prior to conversion is subtracted from the numerator.

Conversion Chemistry

Thermal cracking. Figure 18.21 outlines the mechanism for the thermal cracking of hydrocarbons. This chain-reaction mechanism includes the following steps:

1. Chain reaction initiation: Free radicals form due to the direct thermal rupture of a carbon–carbon bond
2. Chain reaction propagation: A small radical reacts with a large hydrocarbon to give a small alkane and a large primary radical
3. Cracking: The large primary radical decomposes into ethylene and a primary radical with two fewer carbon atoms
4. Isomerization: A primary radical isomerizes into a secondary radical
5. Cracking: A secondary radical decomposes into propylene and a primary radical with three fewer carbon atoms
6. Radical addition: A radical reacts with an olefin to produce a larger radical. Addition reactions such as this eventually generate coke
7. Chain reaction termination: Two radicals react to produce an olefin and hydrogen. The olefin can undergo further reactions per Step 6
8. Chain reaction termination: Two radicals react to produce a large alkane
9. Chain reaction initiation: In this case, the initiator is external, e.g., an organic peroxide

The free-radical mechanism is consistent with the fact that thermal cracking produces significant amounts of hydrogen, methane, ethane, ethylene, and higher olefins.

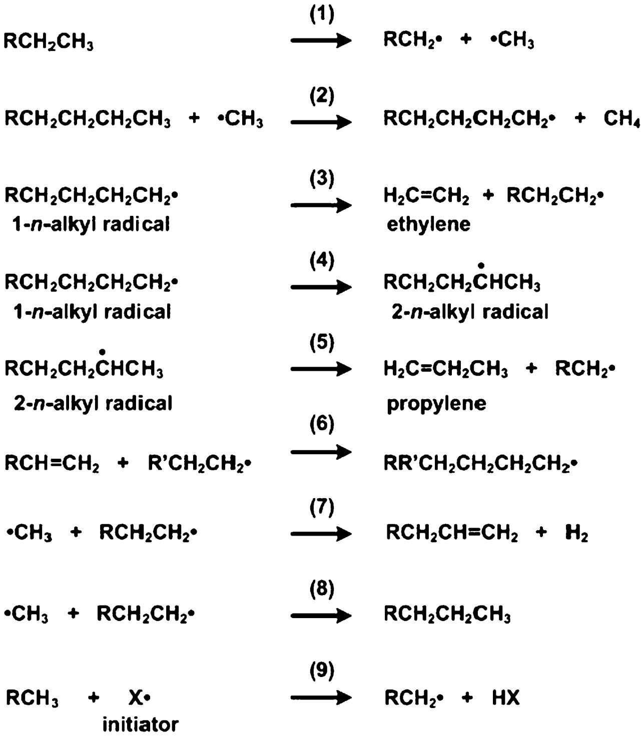
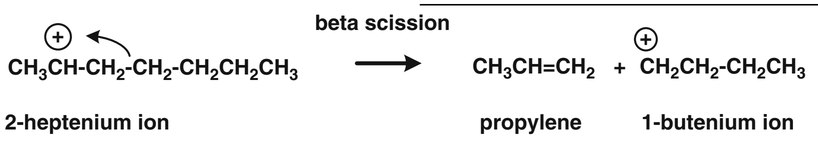


Fig. 18.21 Thermal cracking mechanism [28]

Real-world feeds contain sulfur, nitrogen, oxygen, and other hetero atoms. These form H_2S, ammonia, and water, along with a wide range of hetero-atom organic compounds. Consequently, products (including the water) from delayed coking units contain sulfur and nitrogen compounds.

Catalytic cracking. Figure 18.22 illustrates the mechanism for the catalytic cracking of paraffins. The catalysts contain strong acid sites, and the reactions involve carbocations (also known as carbenium ions or carbenium ions.)

Beta scission is illustrated below. The C–C bond that breaks during acid-catalyzed cracking is two carbons away from the positively charged carbon atom. Also, terminal (1°) carbocations ions are very reactive and quickly rearrange into 2° or 3° carbonations. Therefore, the smallest stable cracked fragment is a C_3 moiety.



The main steps in catalytic cracking are:

1. Dehydrogenation to form a heptene molecule. The reaction is catalyzed by metal sites on the catalyst. It is also reversible. At any point, an olefin can undergo hydrogenation to form a paraffin
2. Abstraction of a hydride ion from *n*-heptane by a Lewis acid site to make a carbocation
3. Addition of a proton to a heptene molecule to make a carbocation, which may or may not be the same as the one produced by Step 2. This reaction is reversible. At any point in the mechanism, a C_3 or higher carbocation can undergo proton elimination to form an olefin
4. Cracking via β scission. This means that cracking occurs two carbon atoms away from the carbon atom holding the positive charge
5. Rearrangement (alkyl shift). As an alternative to Step 4, carbocations can rearrange to form more-stable branched intermediates, which lead to branched products. This explains why the products of catalytic cracking are rich in branched isomers. (In a 3° carbocation, the carbon with the positive charge is attached to three other carbon atoms. In a 2° carbocation, the positive carbon is attached to two other carbons. In a 1° "terminal" carbocation, the positive carbon is attached to only one other carbon. The order of stability is: $3° > 2° > 1°$.)
6. Double-bond shift. After an olefin undergoes double-bond shift, the new olefin can react as before with an acid site per Step 3. It can also undergo alkylation by reacting with a carbocation (Fig. 18.21)
7. Isomerized carbocations can undergo β scission per Step 4

Catalytic cracking units also make aromatics and other heavy products via cyclization, alkylation, and polymerization of intermediate olefins. Figure 18.23 shows the kinds of ring-formation reactions that lead to the generation aromatics and polyaromatics during catalytic cracking. As shown in Fig. 18.16, polyaromatics can grow into larger polyaromatics, which eventually can form coke.

Catalytic hydrocracking. In many respects, the mechanism for catalytic hydrocracking is the same as for catalytic cracking. The main differences are:

- During hydrocracking, aromatic rings are saturated into naphthenic rings, which can then be cracked per the reactions in Fig. 18.24
- The presence of high-pressure hydrogen and metal sites favor the saturation of olefin intermediates. This inhibits ring-formation and ring-growth reactions

For the hydrocracking of fluorene ($C_{13}H_9$), Lapinas et al. [30] propose a mechanism that might apply generally to the hydrocracking of polyaromatics. In simplified form, Fig. 18.24 shows the sequence of reactions by which

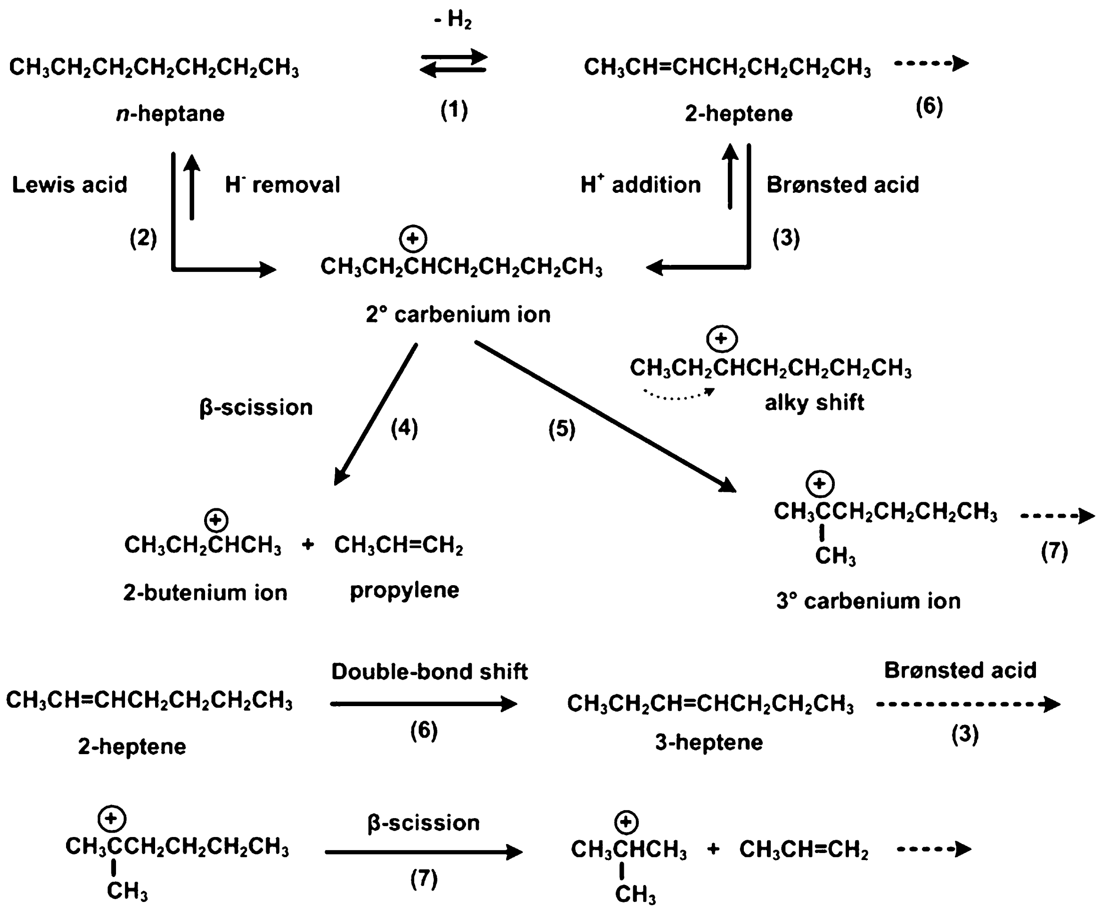


Fig. 18.22 Catalytic cracking of *n*-heptane [28]

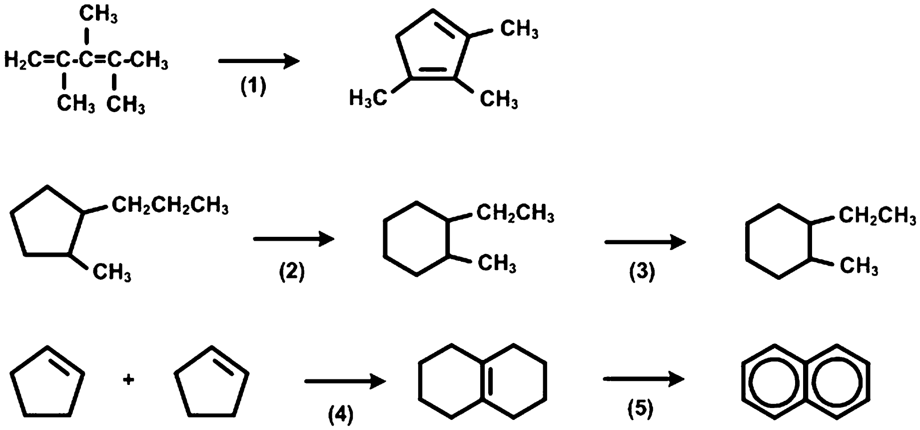


Fig. 18.23 Ring-formation reactions during catalytic cracking [29]

hydrocracking converts a heptyl ethyl naphthalene isomer into lighter compounds. The reactions include:

- Removal of part of a side chain by dealkylation
- Saturation of an aromatic ring
- Isomerization of the saturated ring
- Opening the saturated ring
- Paraffin hydrocracking

The first dealkylation reaction removes hexanes, not heptanes. The ethyl group is not removed, because (a) the production of methane by the β-scission of a two-carbon

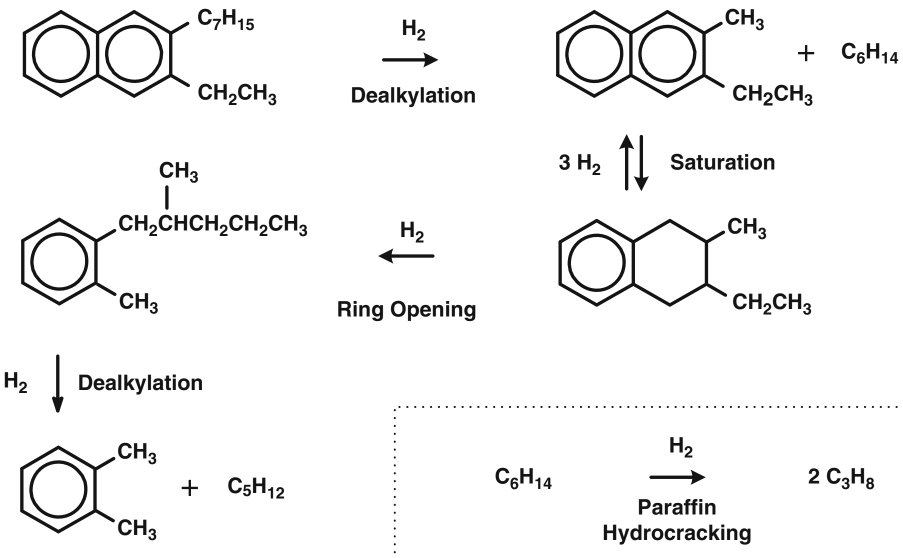


Fig. 18.24 Simplified reaction sequence for the hydrocracking of polyaromatics

C–C bond isn't possible, and (b) it is difficult to form the required intermediate from an aromatic ring stabilized by resonance energy.

Comparison of catalytic and thermal cracking. Based on [28], Table 18.11 provides a comparison between the catalytic and thermal cracking processes:

Alkylation chemistry. Figure 18.25 shows the mechanism for the alkylation reaction of 2-butene with isobutane. During the process, 2-butene can isomerize. Reactions 2a and 2c show how different isomers lead to different products. Reaction 4 illustrates how additional alkylation reactions could eventually produce sludge, coke, and other heavy products.

Catalytic reforming chemistry. Figure 18.26 shows the reactions involved in catalytic reforming, the purpose of which is to transform naphthenes and C_6 to C_{11} paraffins into aromatic compounds. The process yields considerable amounts of hydrogen, which is used in hydrotreaters, hydrocrackers, and other units requiring hydrogen. In contrast to hydrocracking, catalytic reforming operates at low pressure and high temperature, which favors production of aromatics. The conversion of cyclopentanes into cyclohexanes (Reaction 3) is a key step in the production of benzene and alkylbenzenes.

Isomerization chemistry. In refineries, isomerization units convert *n*-paraffins into isoparaffins. In C_4 isomerization, *n*-butane is converted to isobutane in a two-stage process, which employs HCl-promoted $AlCl_3$ for the first stage and a supported noble metal catalyst in the second stage. C_5/C_6 isomerization converts normal pentanes and hexanes and 1° hexanes into highly branched isomers.

Table 18.11 Comparison of product yields for catalytic and thermal cracking processes

Process	Type	Product characteristics
FCC	Catalytic cracking	C_1 and C_2: low Branched/normal paraffin ratio: high Light olefin yields: significant Heavy olefin yields: low Aromatics: higher than feed Coke formation: high Olefin reactivity: higher than paraffins Alkyl aromatics: scission next to the ring
Hydrocracking	Catalytic cracking in the presence of hydrogen	C_1 and C_2: low Branched/normal paraffin ratio: high Olefins: removed by pretreating Aromatics: lower than feed Coke formation: minimal Alkyl aromatics: scission next to the ring
Coking visbreaking	Thermal cracking	C_1 and C_2: high Branched/normal paraffin ratio: similar to feed Olefin production: high Aromatics: low Coke formation: high Olefin reactivity: similar to paraffins Alkyl aromatics: scission within the side chain

Based in part on Table A in [28]

During lube stock production, heavy *n*-paraffins are converted into isoparaffins with minimal cracking over silica-alumina phosphate (SAPO)-based catalysts or similar materials.

Fig. 18.25 Alkylation of 2-butene by isobutane [31]

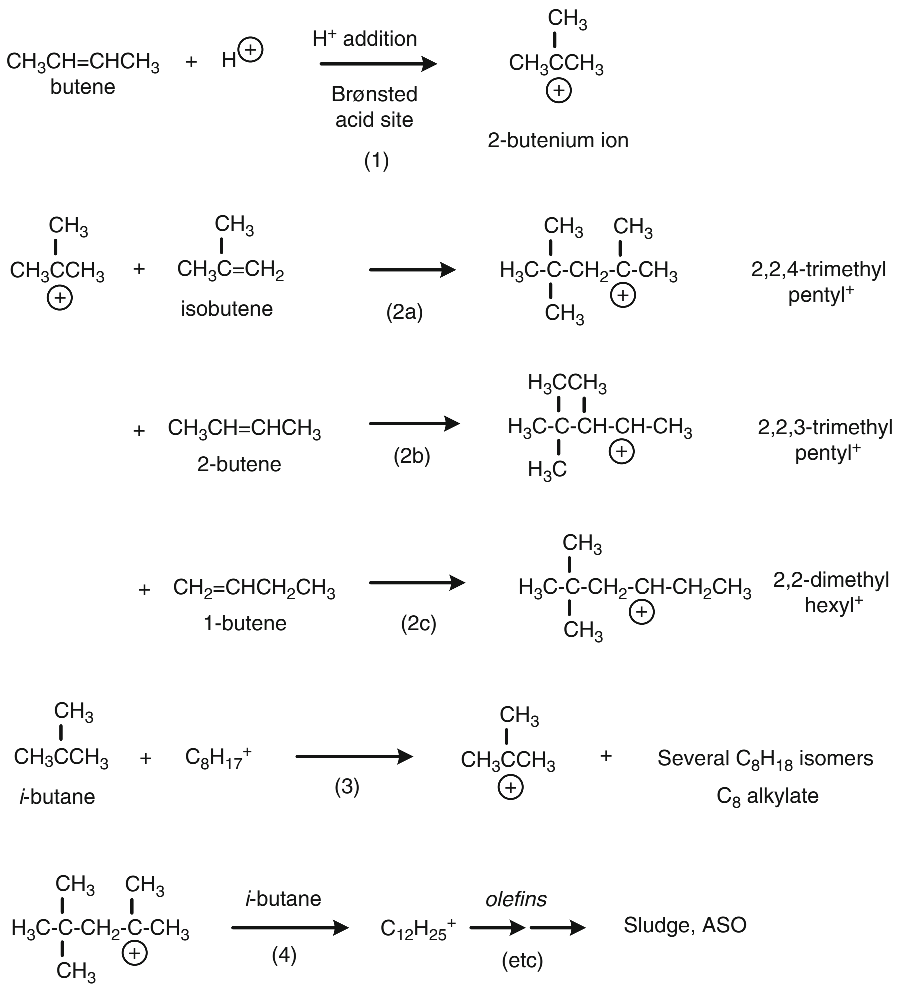
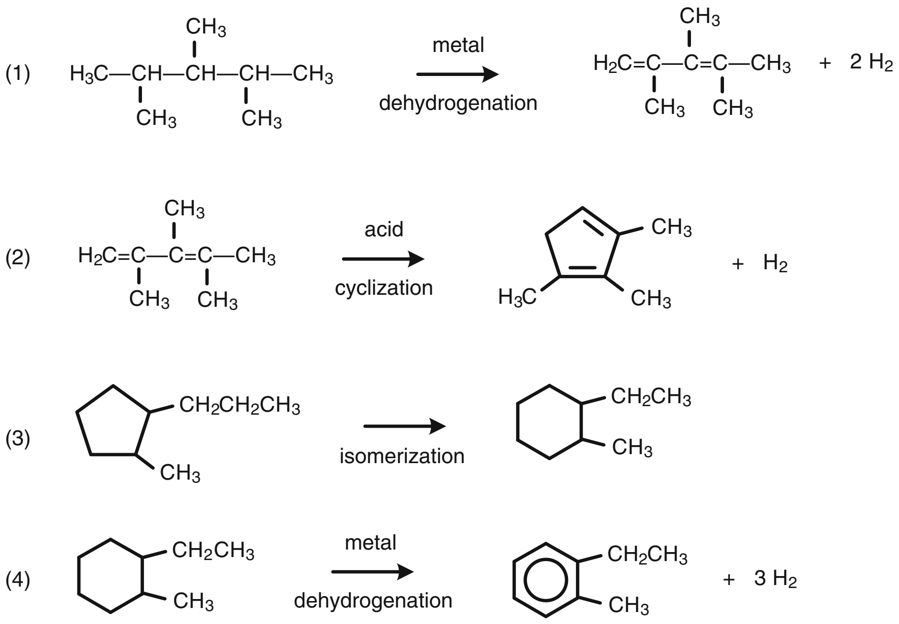


Fig. 18.26 Catalytic reforming reactions



Olefin isomerization converts straight-chain C_4–C_6 olefins into corresponding iso-olefins.

Steam-hydrocarbon reforming (SMR) chemistry. Steam-hydrocarbon reforming produces hydrogen through the reaction of steam with light hydrocarbons at very high temperatures—around 1,500 °F (816 °C). The hydrocarbon feed usually is methane, so even though other hydrocarbons are used, a generic term for the process is steam-methane reforming (SMR). The product of the initial reaction is a mixture of H_2, CO, CO_2, residual methane, and in some cases traces of other hydrocarbons. The initial product goes to one or more shift reactors, where the shift reaction between H_2O and CO yields CO_2 and additional hydrogen.

$$\text{Primary reaction}: CH_4 + H_2O \rightarrow CO + 3H_2$$

$$\text{Shift reaction}: CO + H_2O \rightarrow CO_2 + H_2$$

In older units, residual CO was removed by methanation over a nickel-based catalyst, and residual CO_2 was removed by adsorption with activated molten potassium carbonate (Benfield process). The product hydrogen contained up to 5 % methane.

$$\text{Methanation}: CO + H_2 \rightarrow CH_4$$

In newer units, pressure-swing adsorption (PSA) removes nearly all contaminates, yielding a product containing 99.99 % hydrogen.

Sulfur recovery (*Claus process*) *chemistry*. Claus-process sulfur recovery units burn hydrogen sulfide in just enough air to form a mixture of H_2S and SO_2 in a 2:1 molar ratio. In downstream beds of alumina catalyst, the H_2S reacts with the SO_2 to form elemental sulfur and water.

$$3H_2S + O_2 + S \rightarrow 2H_2S + SO_2 + H_2O$$

$$2H_2S + SO_2 \rightarrow 3S + 2H_2O$$

The product sulfur can include many different isomers of S_x.

Acid/base treating reactions. Several processes entail seemingly simple acid–base reactions. Amine treating removes acid gases—H_2S and CO_2—from fuel gas and off-gas streams. In some hydrotreaters and hydrocrackers, high-pressure amine units remove H_2S from the recycle gas. Caustic scrubbers are used in several ways, including removing the last traces of H_2S from the hydrogen used for processes that are highly sulfur sensitive.

Catalytic NOx removal chemistry. Selective catalytic reduction (SCR) removes nitrogen oxides (NOx) by reaction with ammonia to produce nitrogen via the following main reactions. Secondary reactions involve sulfur oxides and ammonium sulfates.

$$4NO + 4NH_3 + O_2 \rightarrow 4N_2 + 6H_2O$$

$$2NO_2 + 4NH_3 + O_2 \rightarrow 3N_2 + 6H_2O$$

$$NO + NO_2 + 2NH_3 \rightarrow 2N_2 + 3H_2O$$

Petroleum Refining Processes

Some passages in this section are abstracted from [8] and are used with kind permission from the publisher: Springer Science+Business Media, Inc. The *2011 Refining Process Handbook* published by *Hydrocarbon Processing* is an excellent source of additional information on individual processes.

Crude oil pretreatment. Crude oil comes from the ground mixed with a variety of substances: gases, water, salt, and dirt. These must be removed before the crude can be transported effectively and refined without undue fouling and corrosion. Some cleanup occurs in oil fields. In refineries, crude desalting units provide subsequent cleanup.

A field separator is often no more than a large covered vessel that provides enough residence time for gravity separation into four phases: gases, crude oil, water, and solids. Generally, the crude oil floats on the water. The water is withdrawn from the bottom and is disposed of at the well site. Gases are withdrawn from the top and piped to a natural gas processing plant or reinjected into the reservoir to maintain well pressure. Crude oil is pumped either to a refinery through a pipeline or to storage to await transportation by other methods.

At the refinery, the crude is treated with hot water in one or more desalters. Desalters employ either chemical or electrostatic precipitators to remove dissolved salts and collect remaining solids. In chemical desalting, water and surfactants are added to the crude, heated to dissolve salts and other impurities, then sent to a settling tank, where the water and oil separate. In electrostatic desalting, chemicals are replaced with a strong electrostatic charge, which drives the separation of water from oil.

If the crude isn't desalted, residual solids can clog downstream equipment and deposit on heat-transfer surfaces, thereby reducing heat-transfer efficiency. Salts can induce corrosion in major equipment and deactivate catalysts.

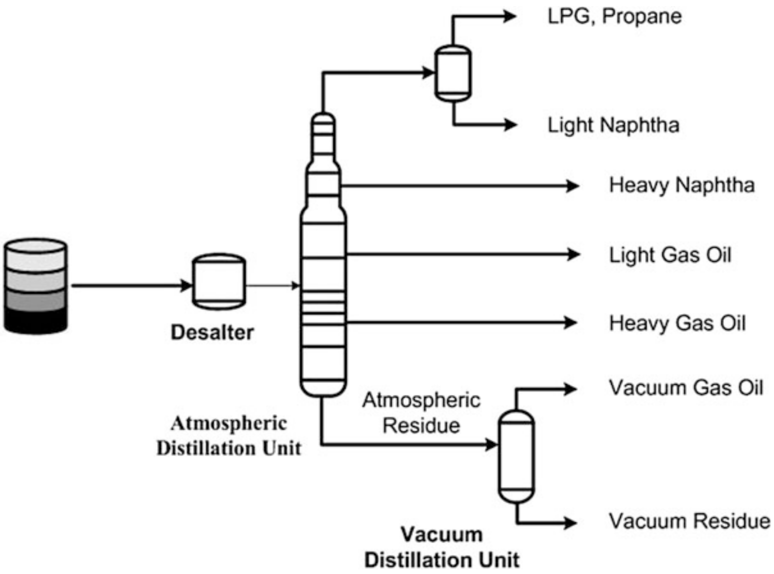


Fig. 18.27 Two-stage crude oil distillation unit [8], Fig. 9, with kind permission from Springer Science+Business Media, Inc

Figure 18.27 presents an overview of a typical two-stage crude oil distillation unit (CDU), and Fig. 18.28 shows an atmospheric distillation tower with bubble-cap internals.

Before reaching the atmospheric tower, desalted oil goes through a network of preheat exchangers to a fired heater, which brings the temperature up to 657–725 °F (347–385 °C). If the oil gets much hotter that this, it starts to crack, generating carbon. The carbon would deposit inside the pipes and equipment through which the oil flows.

The hot crude enters the tower just above the bottom. Steam is added to enhance separation; it does so largely by decreasing the vapor pressure of hydrocarbons in the column. When it enters the tower, most of the oil vaporizes. Unvaporized oil drops to the bottom of the tower, where it is drawn off. The vapors rise through the distillation trays, which contain perforations, bubble caps, downcomers, and/or modifications thereof. Each tray permits vapors from below to bubble up through the relatively cool condensed liquid on top of the tray. This vapor/liquid contacting knocks heavy material out of the vapor. Condensed liquid flows down through a pipe (downcomer) to the hotter tray below, where the higher temperature causes re-evaporation. A given molecule evaporates and condenses many times before finally leaving the tower.

Products are collected from the top, bottom, and side of the column. Side-draw products are taken from trays where the temperature corresponds to the cutpoint for a desired product. Some of the side-draws can be returned to the tower as a pump-around or pump-back stream to control tower temperatures, improve efficiency, and further enhance separation. Reflux plays a major role in controlling temperature at the top of the tower, where gases and light naphtha are drawn off overhead.

The bottom stream from the main fractionator is called atmospheric bottoms, atmospheric residue, or long resid. It goes to a second fired heater, where the typical outlet temperature is about 750–775 °F (400–413 °C). From the second heater, the atmospheric residue goes to a vacuum tower. Steam ejectors reduce the absolute pressure to about 7.0 psia (0.5 bara). Under vacuum, hydrocarbons vaporize at lower temperatures. Thus, molecules with normal boiling points above 650 °F (343 °C) can be vaporized at lower temperatures, where they are less likely to undergo thermal cracking.

Products from the crude distillation unit go to other process units, as shown in Table 18.12.

Solvent refining. Solvent refining processes include solvent deasphalting, solvent extraction, and deoiling/dewaxing processes.

Solvent deasphalting takes advantage of the fact that aromatic compounds are insoluble in paraffins. Propane and *n*-pentane are used to precipitate asphaltenes from residual oils. The deasphalted oil (DAO) is sent to hydrotreaters, FCC units, hydrocrackers, or fuel-oil blending. In FCC units, DAO is easier to process than the corresponding straight-run residue. This is because the asphaltenes in straight-run residue easily form coke and often contain catalyst poisons such as nickel and vanadium; the asphaltene content of DAO is (by definition) almost zero. In hydrocrackers, DAO is harder to process than straight-run VGO and FCC cycle oils, because although DAO no longer contains asphaltenes, it still has a very high endpoint.

In traditional propane deasphalting, residual oil and propane are pumped to an extraction tower at 150–250 °F (65–120 °C) and 350–600 psig (2,514–4,240 kPa). Separation occurs in a tower, which may include a rotating disc

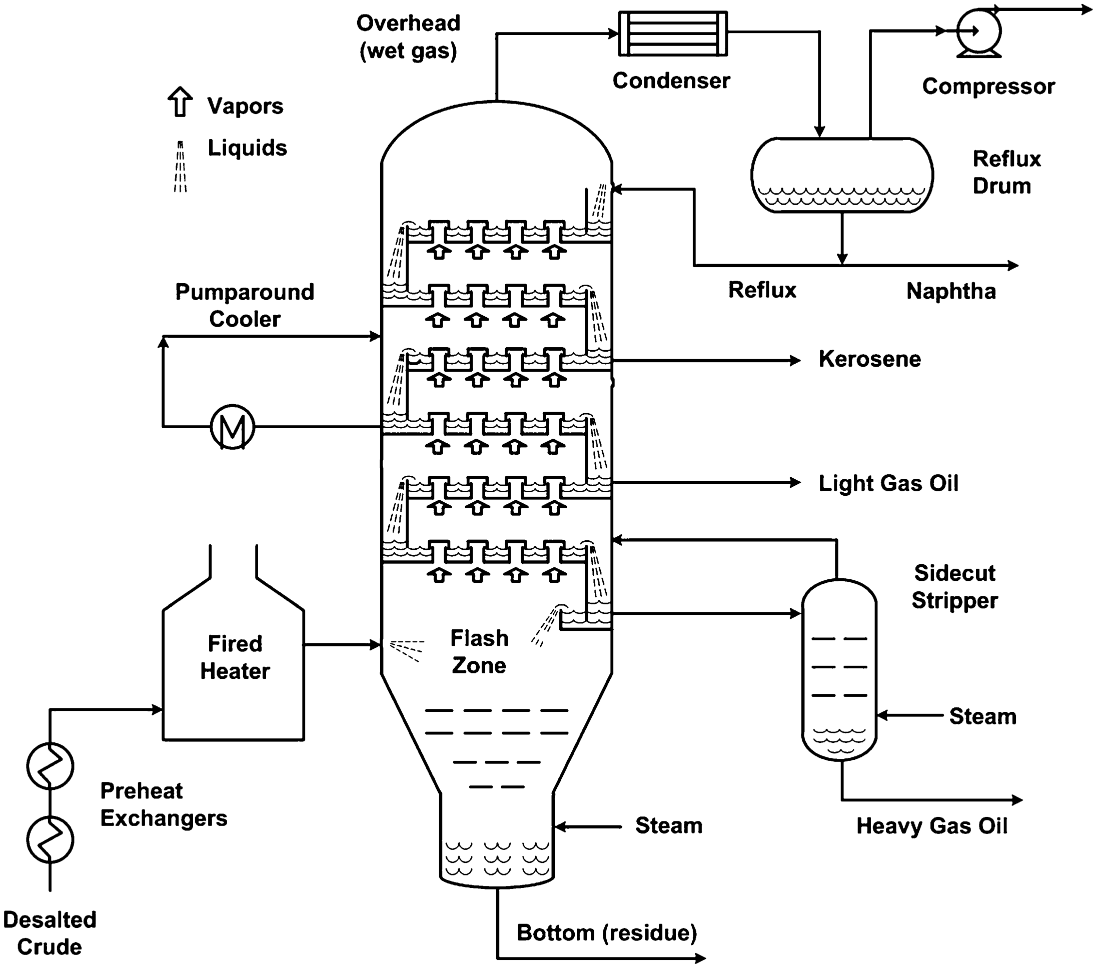


Fig. 18.28 Atmospheric distillation with bubble-cap trays [8], Fig. 10, with kind permission from Springer Science+Business Media, Inc

Table 18.12 Destinations for straight-run distillates

	Approx. boiling range fraction			
	°C	°F	Next destination	Ultimate product(s) or subsequent destination
LPG	−40 to 0	−40 to 31	Sweetener	Propane fuel
Light naphtha	39–85	80–185	Hydrotreater	Gasoline
Heavy naphtha	85–200	185–390	Cat. Reformer	Gasoline, aromatics
Kerosene	170–270	340–515	Hydrotreater	Jet fuel, No. 1 diesel
Gas oil	180–340	350–650	Hydrotreater	Heating Oil, No. 2 diesel
Atmos. resid	340+	650+	Visbreaker Resid hydrotreater Ebullated-bed hydrocracker	FCC or hydrocracker feed, low-viscosity resid Resid FCC Naphtha, gas oils, FCC
Vacuum gas oil	340–566	650–1,050	FCC Hydrotreater Hydrocracker Solvent refining	Gasoline, LCO, gases including C_3/C_4 olefins Fuel oil, FCC, lubes Naphtha, jet, diesel, FCC, olefins, lubes DAO, asphalt
Vacuum Resid	+540	+1,000	Coker Solvent refining Slurry-phase hydrocracker	Coke, coker gas oil, coker naphtha, gases DAO, asphalt Traditional hydrotreater or hydrocracker

Resid is an abbreviation for residue, and 340+ (etc.) means everything that boils above 340 °C (etc.)

contactor. Liquid products are evaporated and steam stripped to recover the propane solvent, which is recycled.

An advanced version of solvent deasphalting is "residuum oil supercritical extraction," or ROSE. The ROSE™ Process was developed by the Kerr-McGee Corporation and now is offered for license by KBR. In this process, residue and solvent are mixed and heated to above the critical temperature of the solvent. Liquid yields are higher under supercritical conditions, because the lighter part of the oil becomes more soluble. In addition to giving higher yields, the process is more energy efficient and has lower operating costs due to improved solvent recovery. The ROSE process can employ three different solvents, the choice of which depends upon process objectives:

Propane	Preparation of lube base stocks
Butane	Asphalt production
Pentane	Maximum recovery of liquid

Solvent extraction is used to remove aromatics and other impurities from lube base stocks and grease stocks. The feedstock is dried, then contacted with the solvent in a counter-current or rotating disk extraction unit. The solvent is separated from the product stream by heating, evaporation, or fractionation. Remaining traces of solvent are removed from the raffinate by steam stripping or flashing. Electrostatic precipitators can enhance separation of inorganic compounds. The solvent is then regenerated and recycled. Today, phenol, furfural, and cresylic acid are widely used as solvents. Liquid sulfur dioxide, chlorinated ethers, and nitrobenzene also have been used.

Solvent dewaxing removes wax (normal paraffins) from deasphalted lube base stocks. The main process steps include mixing the feedstock with the solvent, chilling the mixture to crystallize wax, and recovering the solvent. Commonly used solvents include toluene and methyl ethyl ketone (MEK). Methyl-isobutyl ketone (MIBK) is used in a *wax deoiling* process to prepare food-grade wax, which is used to coat the paper used for milk cartons and other such products.

Visbreaking. Visbreaking is a mild form of thermal cracking that achieves about 15 % conversion of atmospheric residue into gas oils and naphtha. At the same time, a low-viscosity residual fuel is produced. The two main types of visbreaking are "short-contact" and "soaker." In short-contact visbreaking, the feed is heated to about 900 °F (480 °C) and sent to a soaking zone reactor at 140–300 psig (1,067–2,170 kPa). The elevated pressure allows cracking to occur while restricting coke formation. To avoid overcracking, the residence time in the soaking zone is short—several minutes compared to several hours in a delayed coker. The hot oil is quenched with cold gas oil to inhibit further cracking and sent to a vacuum tower for product separation. Soaker visbreaking keeps the hot oil at elevated temperature for a longer time to increase the yield of middle distillates. Low-viscosity visbreaker gas oil can be sent to an FCC unit or hydrocracker for further processing, or used as heavy fuel oil.

Delayed coking. Delayed coking is a cyclic process that employs several coke drums. The drums operate on staggered 18 to 24-h cycles. Each cycle includes preheating the drum, filling it with hot oil, allowing coke and liquid products to form, cooling the drum, and decoking. As it is pumped into a drum, vacuum residue feed is heated to about 900–970 °F (487–520 °C). Thermal cracking begins immediately, generating coke and cracked products. Coke accumulates in the drum while the vapors go to a product fractionator. Meanwhile, hot feed keeps flowing into the drum until it is filled with solid coke. At the end of a cycle, the top and bottom heads of the drum are removed. A rotating cutting tool uses high-pressure jets of water to drill a hole through the center of the coke from top to bottom. In addition to cutting the hole, the water also cools the coke, forming steam as it does so. The cutter is then raised, step by step, cutting the coke into lumps, which fall out the bottom of the drum.

Light products include coker naphtha, light coker gas oil (LCGO), and heavy coker gas oil (HCGO). All of these require further processing due to their high content of olefins, which makes then unstable and poorly suited for direct blending into finished products. The coker naphtha and LCGO are hydrotreated. The HCGO can go either to an FCC unit or a hydrocracker.

Coke can account for up to 30 wt.% of the product. It can be shipped by rail, truck, or conveyor belt to a calciner, which converts *green coke* fresh from the unit into various grades of petroleum coke. Green coke can also be used as-is for fuel.

Sponge coke is named for its sponge-like appearance. It is produced from feeds that have low-to-moderate asphaltene concentrations. If sponge coke meets certain specifications, it can be used to make carbon anodes for the aluminum industry. Otherwise, it serves as a fuel. Sponge coke must be calcined before it can be formed into anodes. Fuel coke sometimes is used without prior calcination.

Needle coke, named for its needle-like structure, is a high-value product made from feeds that contain nil asphaltenes, such as hydrotreated FCC decant oils. Needle coke is converted into graphite electrodes for the electric-arc furnaces used in the steel industry.

Shot coke is undesirable because it tends to be unstable. It forms when the concentration of feedstock asphaltenes and/or coke-drum temperatures are too high. A block of shot coke is a cluster of discrete mini-balls 0.1–0.2 in. (2–5 mm) in diameter. The clusters can be as large as

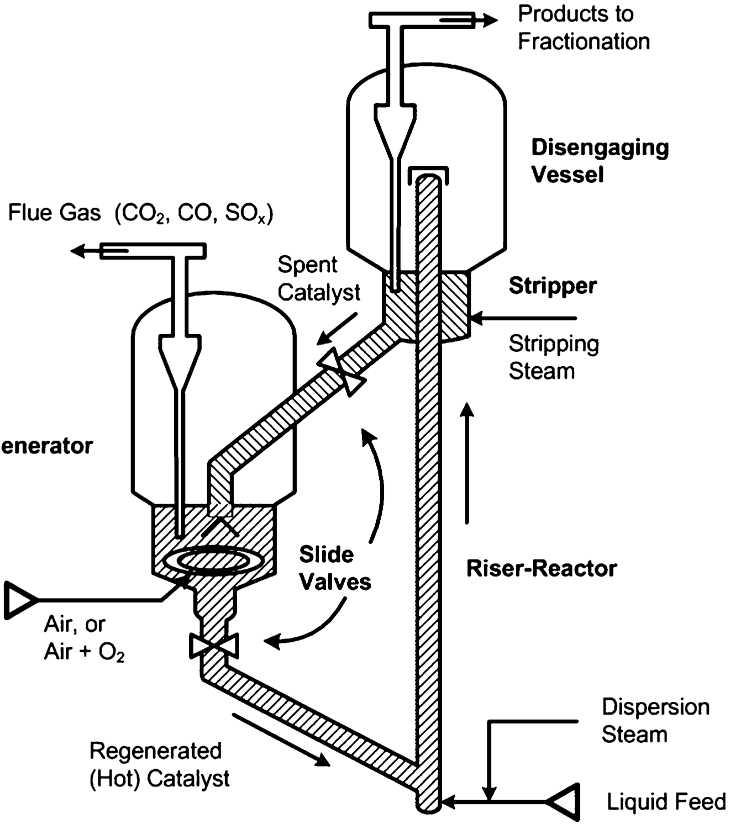


Fig. 18.29 FCC process flow [8], Fig. 13, with kind permission from Springer Science+Business Media, Inc

10 in. (25 cm) across. If a cluster breaks apart when the coke drum is opened, it can spray a volley of hot mini-balls in every direction. Adding aromatic feeds, such as FCC decant oil, can eliminate shot coke formation. Other methods of eliminating shot coke—decreasing temperature, increasing drum pressure, increasing the amount of product recycle—decrease liquid yields, which is not desired.

Specialty carbon products made from petroleum include recarburizer coke, which is used to make special steels, and titanium dioxide coke, which is used as a reducing agent in the titanium dioxide pigment industry.

Fluid coking. Fluid coking, also called continuous coking, is a moving-bed process for which the operating temperature is higher than the temperatures used for delayed coking. In continuous coking, hot recycled coke particles are combined with liquid feed in a radial reactor at about 50 psig (446 kPa). Vapors are taken from the reactor, quenched to stop any further reaction, and fractionated. The coke goes to a surge drum, then to a classifier, where the larger particles are removed as product. The smaller coke particles are recycled to a preheater, where they are mixed with fresh feed. Coking occurs both in the reactor and in the surge drum. Installation costs for fluid coking are somewhat higher than for delayed coking, but feeds can be heavier and heat losses are lower. Fluid coking makes more fuel gas than delayed coking.

Fluid catalytic cracking (*FCC*). FCC produces more than half the world's gasoline.

FCC process flow. A typical FCC unit (Fig. 18.29) comprises three major sections—riser/reactor, regenerator, and fractionation. In the riser/reaction section, preheated oil is mixed with hot, regenerated catalyst. The mixture acts as a fluid because the catalyst particles are about the size of sifted flour. The hot catalyst vaporizes the oil, and the vaporized oil carries the catalyst up the riser/reactor. The cracking reaction is very fast. It produces light gases, high-octane gasoline, and heavier products called light cycle oil (LCO), heavy cycle oil (HCO), slurry oil, and decant oil. It also leaves a layer of coke on the catalyst particles, rendering them inactive.

At the top of the riser, the riser outlet temperature (ROT) can reach 900–1,020 °F (482–549 °C). The ROT determines conversion and affects product selectivity, so FCC operators control it as tightly as possible. Higher temperatures favor production of olefin-rich light gases, especially propylene, at the expense of gasoline; many FCC units maximize propylene for use in nearby olefin plants. Moderate temperatures favor gasoline production. Lower temperatures decrease gasoline yields and increase heavier products—LCO and HCO.

In the disengaging section, steam helps separate the now-deactivated catalyst from the reaction products. The catalyst goes to the regenerator, where the coke is burned away by fluidized combustion in the presence of air or oxygen-enriched air. The regenerated catalyst is hot, with temperatures up to 1,350 °F (732 °C). It is returned to the riser/reactor, where the cycle begins again.

In a 60,000 barrels-per-day unit processing a typical mixture of vacuum gas oils, the total catalyst in the unit (the "inventory") is 400–500 tons. To maintain activity, about 0.5–1 wt.% of the inventory is replaced each day. If the feed to the unit contains significant amounts of residue, the replacement rate is higher. The discharged catalyst is cooled and shipped either to a land fill for disposal or to another refiner, which may have a particular use for "conditioned" FCC catalyst.

FCC feed pretreating. These days, most refiners pretreat FCC feeds in a fixed-bed hydrotreater. The hydrotreater removes trace metal contaminants such as nickel and vanadium. Nickel increases coke formation and decreases liquid yields. Vanadium reduces conversion, decreases liquid yields, and destroys the catalyst. In addition to removing Ni and V, the pretreater decreases concentrations of sulfur, nitrogen, and aromatics. In the FCC regenerator, sulfur on the coked catalyst is converted to sulfur oxides (SOx) in the flue gas. Clean air regulations restrict SOx emissions, which cause acid rain. Therefore, removing sulfur from the FCC feed—thereby reducing SOx formation—is highly beneficial. Removing nitrogen is beneficial, too, because basic feed nitrogen suppresses the activity of highly acidic FCC catalysts. Pretreating also saturates aromatics. As we have seen, saturating aromatics makes them more crackable, so pretreating increases FCC conversion, often by more than 10 vol.%.

FCC heat balance. FCC units must be heat-balanced, or they won't run. The burning of coke in the regenerator provides all of the heat required by the process. In fact, FCC units are significant sources of high-quality steam for other refinery units. Table 18.13 gives a representative breakdown of FCC heat requirements.

Table 18.13 Representative FCC heat balance

Factor	Portion of total
Heat up and vaporize fresh feed	40–50 %
Heat recycled oil	0–10 %
Heat of reaction (endothermic)	15–30 %
Heat steam	2–8 %
Heat losses	2–5 %
Heat air to regenerator temperature	15–25 %
Heat coke from the reactor to regenerator temperature	1–2 %
Total heat duty	500–1,000 Btu/lb 1160–2,325 kJ/kg

Residue FCC. Some FCC units process significant amounts of residue. These units use catalyst coolers (e.g., steam coils) in the regenerator or a second regeneration zone to remove excess heat from the unit. This is because residue generates substantially more coke than conventional FCC feeds, and excess heat is generated when the extra coke is burned off the catalyst. The trace-metal content of residues can be very high. Trace metals destroy FCC catalysts, so removing them—usually with hydrotreating—is essential.

Hydroprocessing. Hydrotreating and hydrocracking are similar processes. Their flow schemes are similar. Both use high-pressure hydrogen to catalytically remove contaminants from petroleum fractions. Both employ the same kinds of equipment—pumps, compressors, furnaces, heat exchangers, and product recovery hardware. In most cases, fixed-bed hot-wall reactors are used. Feeds and products for typical hydrotreaters and hydrocrackers are shown in Table 18.14.

Hydrotreating. Hydrotreating is used to convert sulfur-containing hydrocarbons into low-sulfur liquids and hydrogen sulfide. It also removes nitrogen, oxygen, and other trace contaminants. Hydrotreater feeds range from naphtha to vacuum residues. Generally, each fraction is treated separately. Materials with higher boiling points require more severe treatment conditions. For example, naphtha hydrotreating can be done at 200–500 psia and at 500–650 °F with a hydrogen consumption of 10–50 scf/bbl of feed. Conversely, a residue hydrotreating process might operate at 2,000 psia and at 650–800 °F, with a hydrogen consumption of 600–1,200 scf/bbl.

Figure 18.30 presents a process flow scheme for a two-reactor fixed-bed hydrotreater. Units designed for light feeds usually are simpler, with only one reactor and sometimes only one catalyst bed. Reaction conditions depend on feed quality and process objectives. Oil and hydrogen-rich gas flow down through reactors loaded with catalysts. Makeup gas comes in to replace consumed hydrogen. Gas flow can be once-through in naphtha hydrotreaters, but in distillate and VGO hydrotreaters, unconsumed hydrogen is recycled.

Hydrotreating produces both H_2S and NH_3. Under reaction conditions, these remain in the gas phase. But at lower temperatures, they combine to form solid ammonium bisulfide (NH_4SH). Ammonia also reacts with chlorides to form NH_4Cl; chloride can come with makeup gas, feed, or wash water. These salts can deposit in air coolers and heat exchangers, blocking flow and—even worse—inducing corrosion. Fortunately, they are water-soluble, so they can be controlled by injecting wash water into the reactor effluent.

FCC gasoline post-treating. Conventional hydrotreating does a good job of removing sulfur from FCC gasoline.

Table 18.14 Feeds and products for hydroprocessing units

Feeds	Products from hydrotreating	Products from hydrocracking
Naphtha	Catalytic reformer feed	LPG
Straight-run light gas oil	Kerosene, jet fuel	Naphtha
Straight-run heavy gas oil	Diesel fuel	Naphtha
Atmospheric residue	Lube base stock, low-sulfur fuel oil, RFCC[a] feed	Naphtha, middle distillates, FCC feed
Vacuum gas oil	FCC feed, lube base stock	Naphtha, middle distillates, FCC feed, lube base stock, $C =$ plant feed
Vacuum residue	RFCC[a] feed	See note[b]
FCC light cycle oil	Diesel blend stocks, fuel oil	Naphtha
FCC heavy cycle oil	Blend stock for fuel oil	Naphtha, middle distillates
Visbreaker gas oil	Diesel blend stocks, fuel oil	Naphtha, middle distillates
Coker gas oil	FCC feed	Naphtha, middle distillates, FCC feed, lube base stock, $C =$ plant feed
Deasphalted oil	Lube base stock, FCC feed	Naphtha, middle distillates, FCC feed, lube base stock

[a] RFCC = "residue FCC unit" or "reduced crude FCC unit," which are specially designed to process feeds that contain high concentrations carbon-forming compounds

[b] Traditional fixed-bed hydrocrackers cannot process vacuum residue. However, ebullated-bed and slurry-phase hydrocrackers can. Products from the latter include naphtha, middle distillates, and FCC feed

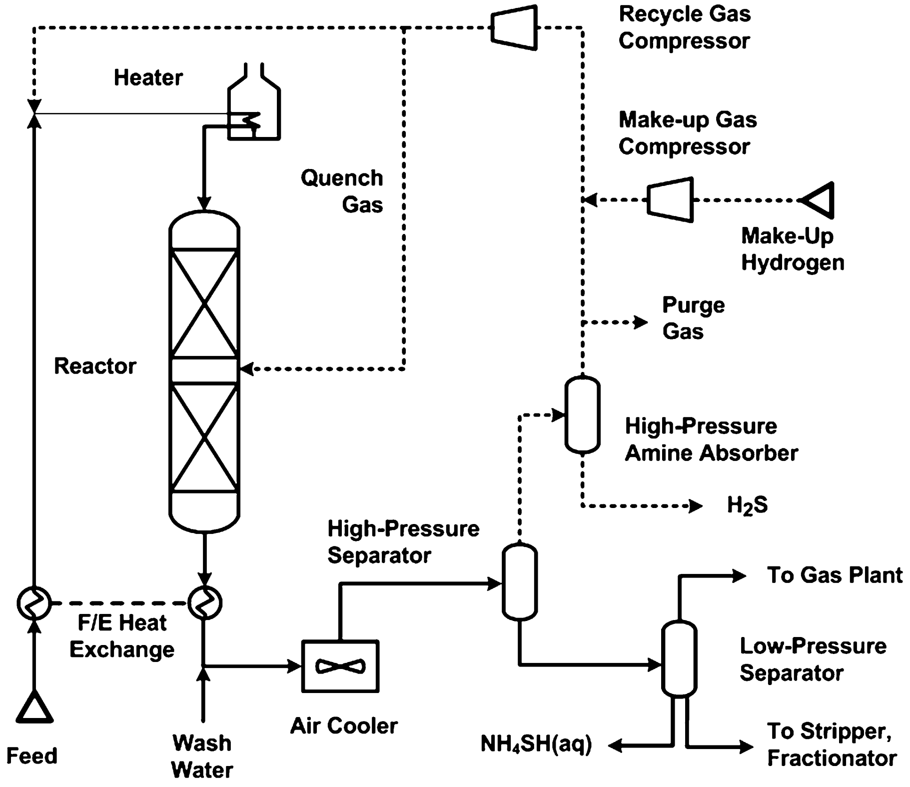


Fig. 18.30 Fixed-bed hydrotreating process flow [15], Fig. 10.20, with kind permission from Woodhead Publishing

Unfortunately, it also does a good job of reducing octane by saturating C_6–C_{10} olefins. Several processes remove sulfur at minimum octane loss. Licensors include Axens (IFP), CDTECH, ExxonMobil, and UOP. The ConocoPhillips S Zorb process uses selective adsorption to remove sulfur from FCC gasoline. The feed is combined with a small amount of hydrogen, heated, and injected into an expanded fluid-bed reactor, where a proprietary sorbent removes sulfur from the feed. A disengaging zone in the reactor removes suspended sorbent from the vapor, which exits the reactor as a low-sulfur stock suitable for gasoline blending. The sorbent is withdrawn continuously from the reactor and sent to the regenerator section, where the sulfur is removed as SO_2 and sent to a sulfur recovery unit. The clean sorbent is

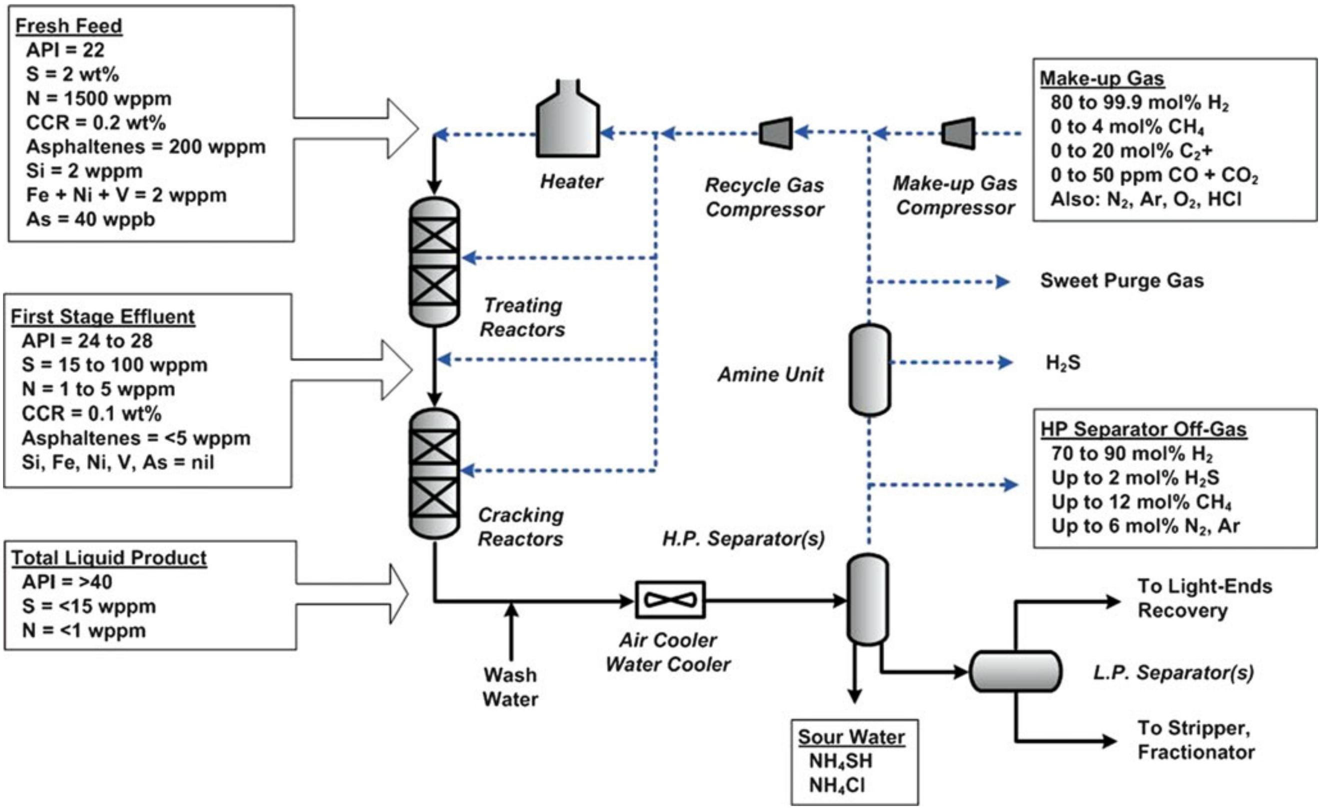


Fig. 18.31 Once-through hydrocracking process flow [15], Fig. 10.21, with kind permission from Woodhead Publishing

reconditioned and returned to the reactor. The rate of sorbent circulation is controlled to help maintain the desired sulfur concentration in the product.

Fixed-bed hydrocracking. As with other conversion processes, the purpose of hydrocracking is to break carbon–carbon bonds, thereby converting heavy hydrocarbons into lighter ones. Figure 18.31 shows a two-reactor once-through hydrocracker with some typical properties for feed and product streams. Figure 18.32 shows a two-stage hydrocracker, in which unconverted oil from the first two reactors goes to a third reactor for additional conversion. With respect to equipment and process flow, fixed-bed hydrocrackers are similar to fixed-bed hydrotreaters. Hydrotreating is an integral part of hydrocracking, because hydrotreating is used to remove organic nitrogen from the feed. Organic nitrogen poisons acidic cracking sites, so its removal is essential.

As shown in Table 18.15, a fixed-bed hydrocracking unit can have significant product flexibility, producing either large amounts of C_4-plus naphtha or large amounts of middle distillates. In petroleum refining, this kind of flexibility is unique.

Ebullated-bed hydrocracking. In contrast to fixed-bed hydrocrackers, ebullated-bed (e-bed) units can process large amounts of residual oils. Catalyst life does not limit these units, because fresh catalyst is added and spent catalyst is removed continuously. In e-bed units (Fig. 18.33), hydrogen-rich recycle gas is bubbled up through a mixture of oil and catalyst particles. This provides three-phase turbulent mixing, which is needed to ensure a uniform temperature distribution. At the top of the reactor, catalyst is disengaged from the process fluids, which are separated in downstream flash drums. Most of the catalyst is returned to the reactor. Some is withdrawn and replaced with fresh catalyst. When compared to fixed-bed processes e-bed technology offers the following advantages:

- The ability to achieve more than 50 wt.% conversion of atmospheric residue
- Ample free space between catalyst particles, which allows entrained solids to pass through the reactor without accumulation, plugging, or build up of pressure drop
- Better liquid-product quality than delayed coking

Disadvantages versus fixed-bed processes include high catalyst attrition, which leads to high rates of catalyst consumption; higher installation costs due to larger reactor volume and higher operating temperatures; and sediment formation. Recent improvements include second-generation catalysts with lower attrition; catalyst rejuvenation, which allows the reuse of spent catalysts; improved reactor design leading to higher single-train feed rates; and two-reactor layouts with inter-stage separation.

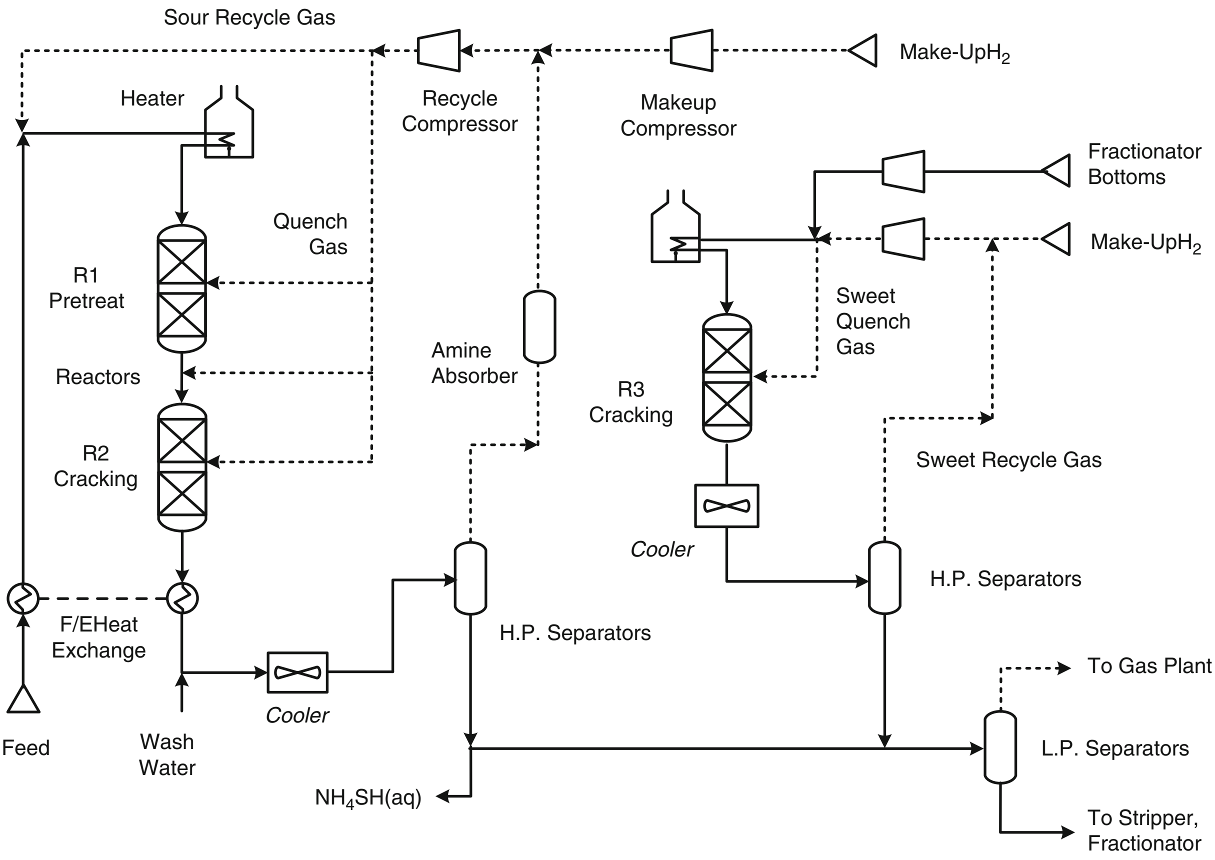


Fig. 18.32 Two-stage hydrocracking process flow [15], Fig. 10.22, with kind permission from Woodhead Publishing

Slurry-phase hydrocracking. Slurry-phase hydrocracking (Fig. 18.34) achieves 95 wt.% conversion of vacuum residue or coal tar. Process conditions are severe—more than 840 °F (450 °C) and 2,000–3,000 psig (14,000–20,800 kPa). The catalysts are nonacidic, comprising finely divided carbon or powered iron sulfate. Inside the reactor, the liquid/powder mixture behaves as a single phase due to the small size of the additive particles. The additives prevent bulk coking by providing highly dispersed nucleation sites for "micro coking" per [26]. The additive isn't recovered. Instead, it ends up in a pitch fraction, which comprises <5 wt.% of the vacuum residue or coal tar feed.

The advantages of slurry-phase processes include:

- The ability to achieve more than 95 wt.% conversion of vacuum residue and other feedstocks, including coal [33]
- Good product quality in two-stage designs, which incorporate a fixed-bed hydrotreating/hydrocracking reactor
- Feedstock flexibility
- Low-cost, micron-size disposable catalysts
- For a given volume of residue feed, lower reactor volume than e-bed processes

Disadvantages include the need to dispose of the coproduced pitch.

Catalytic reforming. The three major process flows for catalytic reforming are:

- Semi-regenerative
- Cyclic
- Continuous catalyst regeneration (CCR)

Figure 18.35 shows a fixed-bed semi-regen reformer. Catalyst cycles last from 6 to 12 months. A cycle ends when the unit is unable to meet its process objectives—typically octane and overall C_5-plus yields. At the end of a cycle, the entire unit is brought down and coke is burned off the catalyst. Desulfurized naphtha is mixed with hydrogen, heated to >900 °F (>480 °C) and passed through a series of fixed-bed reactors. At mentioned the feed is spiked with an organic chloride, which converts to hydrogen chloride in the reactors. This provides the required catalyst acidity and helps minimize catalyst coking.

The major chemical reactions—dehydrogenation and dehydrocyclization—are endothermic and the reactors are adiabatic. Consequently, the temperature drops as reactants flow through a reactor. Between reactors, fired heaters bring the process fluids back to desired reactor inlet temperatures (RIT). The last reactor effluent is cooled and sent to a separator, from which hydrogen-rich gas is removed and

Table 18.15 Hydrocracking product flexibility

Feed	Straight-run vacuum gas oil		
Boiling range, °C	340–550		
Boiling range, °F	644–1,020		
API gravity	22.0		
Specific gravity	0.9218		
Nitrogen, wppm	950		
Sulfur, wt%	2.5		
Primary product objective	Naphtha	Kerosene	Gas oil
Weighted average reactor temp, °C	base	−6	−12
Weighted average reactor temp, °F	base	−11	−22
Product yields, vol % fresh feed			
Butanes	11	8	7
Light naphtha	25	18	16
Heavy naphtha	90	29	21
Kerosene or gas oil	–	69	77
Total C_4-plus	126	124	121
Chemical H_2 consumption			
Nm^3/m^3	345	315	292
Scf/bbl	2,050	1,870	1,730
Product qualities			
Light naphtha (C_5-82 °C)			
RON Clear	79	79	80
Heavy naphtha			
P/N/A	45/50/5	44/52/4	–
RON clear	41	63	67
End point, °C (°F)	216 (421)	121 (250)	118 (244)
Kerosene			
Flash point, °C (°F)	–	38 (100)	–
Freeze point, °C (°F)	–	−48 (−54)	–
Smoke point, mm	–	34	–
FIA aromatics, vol%	–	7	–
End point, °C (°F)	–	282 (540)	–
Gas oil			
Cloud point, °C (°F)	–	–	−15 (5)
API gravity	–	–	44
Cetane number	–	–	55
Flash point, °C (°F)	–	–	52 (126)
End point, °C (°F)	–	–	349 (660)

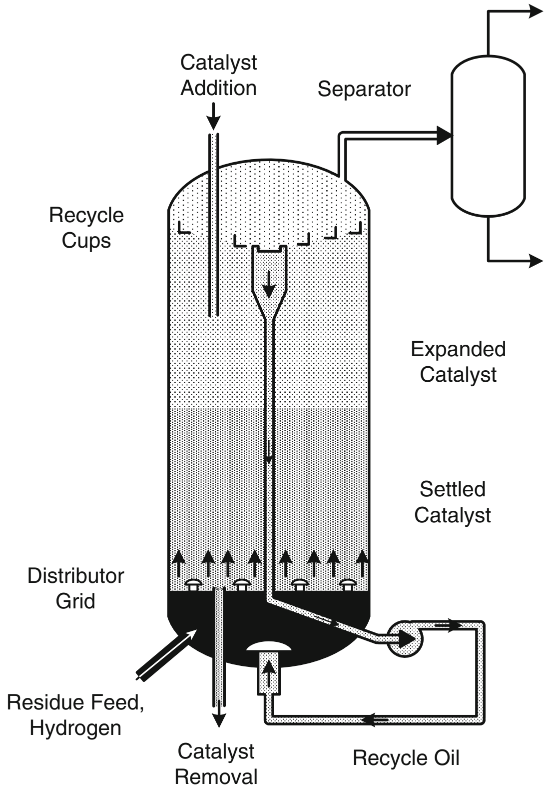


Fig. 18.33 Ebullated-bed hydrocracking process flow [32], Fig. 21, with kind permission from Springer Science+Business Media, Inc

recycled to the reactors. The liquid product flows to a stabilizer column, where entrained gases are removed, before going to the gasoline blender or aromatics plant.

A cyclic reformer has more reactors, and catalyst cycles are shorter—20 to 40 h. Shutdowns are staggered so that only one reactor is down at a given time.

In the CCR Platforming process (Fig. 18.36) hydrotreated feed mixes with recycle hydrogen and goes to a series of adiabatic, radial-flow reactors arranged in a vertical stack. Catalyst flows down the stack, while the reaction fluids flow radially through the catalyst beds. Heaters are used between reactors to reheat the reaction fluids to the required temperature. Flue gas from the fired heaters is typically used to generate steam. A CCR can operate at very low pressure (100 psig, 791 kPa). This improves yields of aromatics and hydrogen, but it accelerates catalyst deactivation by increasing the rate of coke formation. But faster coke formation is okay in a CCR reformer, because the catalyst always is being regenerated.

Isomerization. As mentioned, isomerization units convert *n*-paraffins into isoparaffins and straight-chain olefins into iso-olefins. Butane isomerization provides isobutane for subsequent alkylation. Pentane/hexane isomerization improves the octane of light naphtha streams.

In butane isomerization, the feed contains *n*-butane or, more typically, mixed butanes. The catalyst is highly sensitive to water, so the feed must be thoroughly dried. In the low-temperature first stage, the catalyst comprises aluminum chloride promoted by HCl. Hydrogen gas is added to inhibit olefin formation. In the high-temperature second stage, the catalyst contains a noble metal such as platinum. Temperatures range from 230 to 340 °F (110–170 °C) and pressure ranges from 200 to 300 psi (14–20 bar). The reactor effluent goes to a flash drum, from which hydrogen is recovered and recycled. HCl is removed in a stripper column. The liquids go to a fractionator (deisobutanizer), which separates unconverted *n*-butane from the isobutane product. The *n*-butane is mixed with fresh feed and recycled.

In the Shell Hysomer process for pentane/hexane isomerization, the feed is combined with hydrogen-rich gas, heated to 445–545 °F (230–285 °C) and routed to the

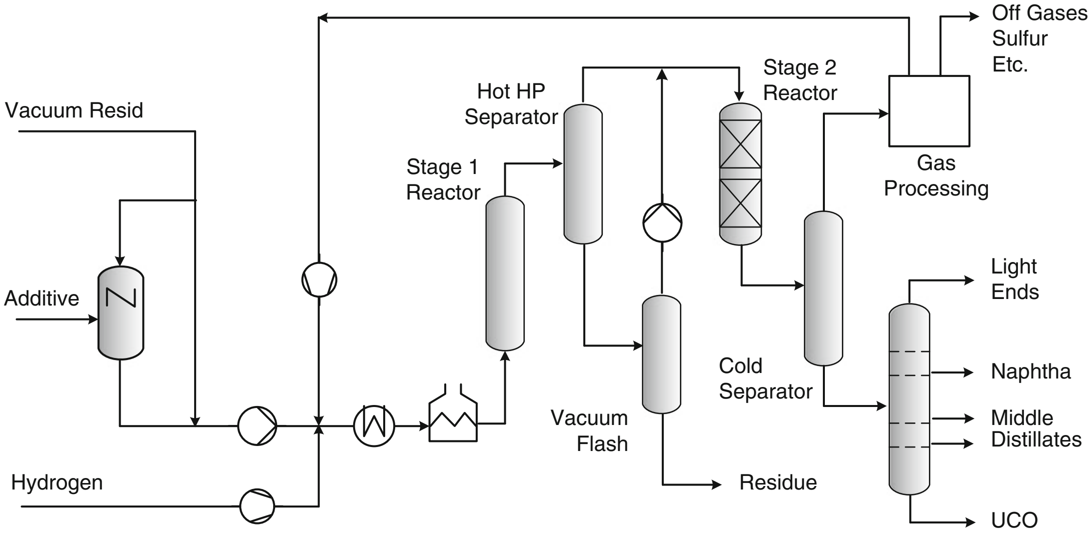


Fig. 18.34 Two-stage slurry-phase hydrocracking process flow. Based on drawings supplied by KBR Technology. Used with kind permission from KBR, Inc

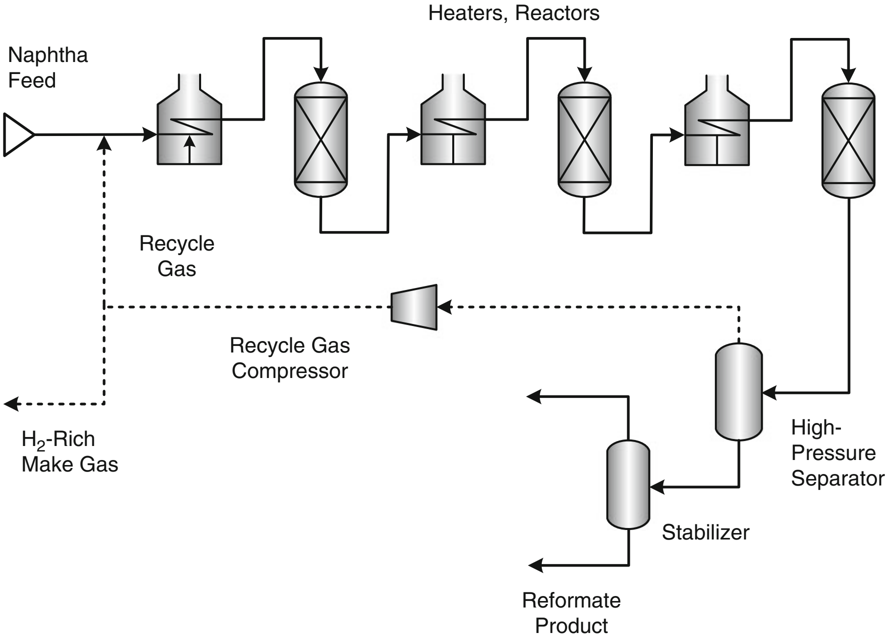


Fig. 18.35 Semi-regen catalytic reforming process flow [8], Fig. 18, with permission from Springer Science+Business Media, Inc

Hysomer reactor at 190–440 psi (13–30 bar). As with fixed-bed hydrotreating and hydrocracking, the process fluids flow down through the catalyst bed, where a part of the *n*-paraffins are converted into branched paraffins. The catalyst is comprised of a noble metal on a zeolite-containing support. The reactions are exothermic, and temperature rise is controlled by injecting relatively cold quench gas. The reactor effluent is cooled and sent to a flash drum, which separates hydrogen from the liquid product. The hydrogen is recycled. The liquid is fractionated, and the *n*-paraffins are recycled. The net conversion of *n*-paraffins into branched products can be as high as 97 %, and the octane can be boosted by 8–10 numbers.

Often, the heat-exchanger and fractionation systems of isomerization units are integrated with those of other process units, such as catalytic reformers. In the total isomerization process (TIP), C5/C6 isomerization is integrated with molecular sieve separation, which provides complete conversion to *n*-paraffins.

Fig. 18.36 CCR catalytic reforming process flow [8], Fig. 19, with kind permission from Springer Science+Business Media, Inc

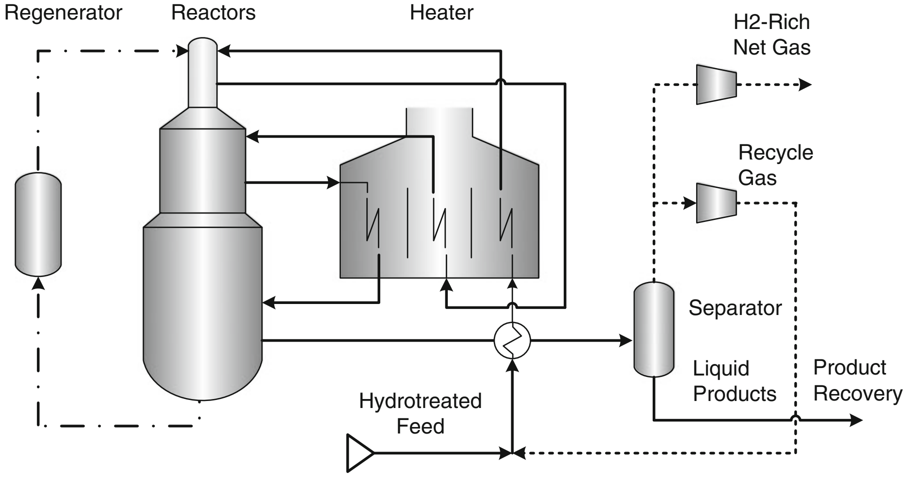
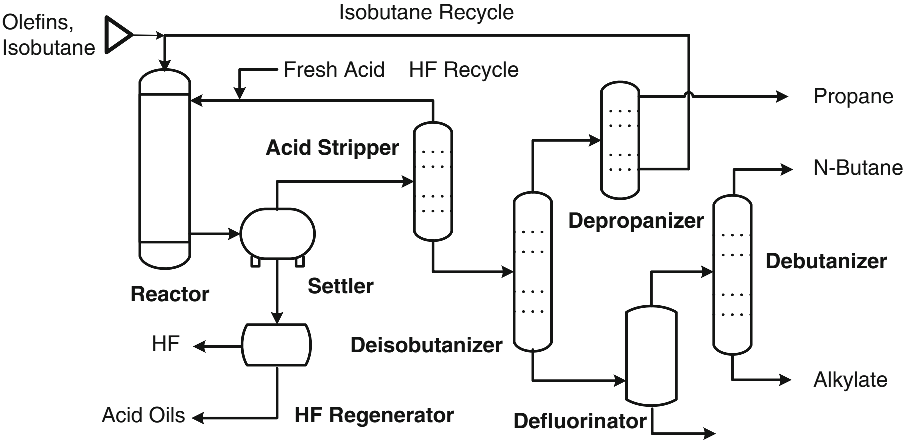


Fig. 18.37 HF Alkylation process flow [8], Fig. 24, with kind permission from Springer Science+Business Media, Inc



The CDTECH® ISOMPLUS process achieves near-equilibrium conversion of *n*-butenes into isobutylene, and n-pentenes into isoamylene over a zeolite catalyst.

Alkylation. During alkylation, isobutane reacts with light olefins such propylene, butylenes, and amylenes in the presence of strong acids to form branched chain hydrocarbons. These hydrocarbons, often referred to as alkylate, have high-octane values and high vapor pressures, making them an excellent contributor to the gasoline blending pool.

Alkylation employs either hydrofluoric acid or sulfuric acid. The process usually runs at low temperatures to avoid polymerization of the olefins. Temperatures for HF catalyzed reactions are approximately 100 °F (38 °C). For sulfuric acid they are approximately 50 °F (10 °C). Figure 18.37 describes HF alkylation.

Petroleum Refining Catalysts

Most of the processes we have discussed require catalysts. Except for alkylation units, where the acids are in liquid form, most refinery catalysts are solids supported on high-strength solids such as γ-alumina, silica, or alumino silicates.

Solid acids, especially synthetic zeolites, are used extensively in catalysts for FCC and hydrocracking. Zeolites occur naturally, but the ones used for refinery catalysis are synthetic. Synthetic zeolites are also used as drying agents, ion-exchangers, and molecular sieves for gas separation. Their microporosity provides them with high surface area, which contributes to their superb catalytic activity. They also possess a mesopore structure that affects the kinds of molecules they crack and the products formed by the cracking.

Related materials are used in other processes. For example, metal-promoted silica-alumina phosphates such as SAPO-11 are used to isomerize normal paraffins in catalytic dewaxing (CDW) units.

Figure 18.38 shows six zeolite structures. In cracking catalysts, HY zeolite is the most common, but beta and ZSM-5 are used as well. The building blocks of A, X, and other zeolites are tetrahedral units of Si and Al oxides. In the ultra-stable Y (USY) zeolites employed in cracking catalysts, the Si/Al ratio is >10.

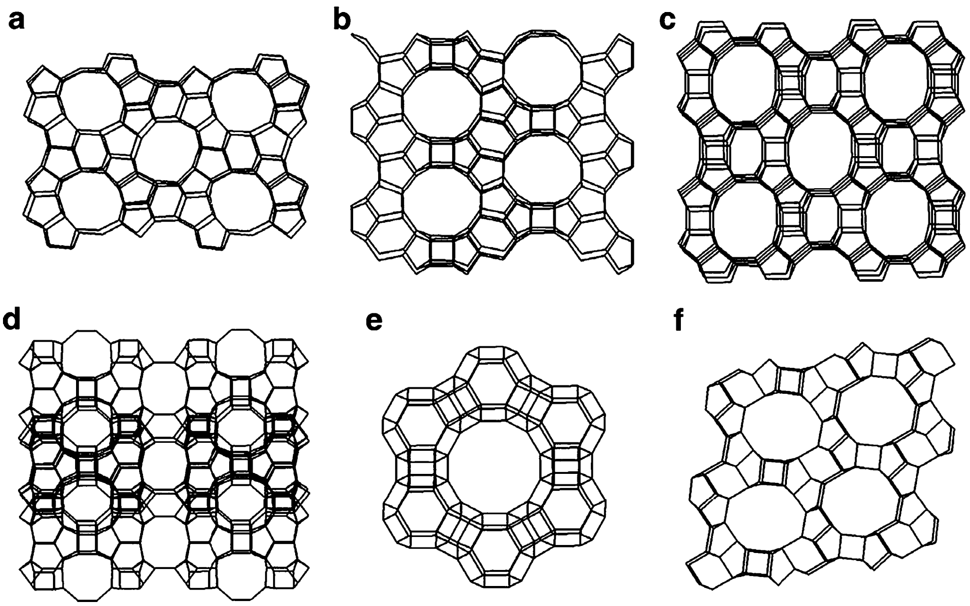


Fig. 18.38 Structures of zeolites ZSM-5 (**a**), mordenite (**b**), beta (**c**), MCM-22 (**d**), zeolite Y (**e**), and zeolite L (**f**). This image is Fig. 1 in [34]

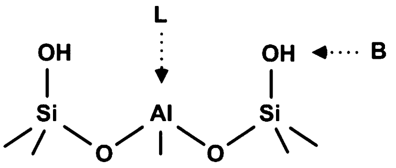


Fig. 18.39 Brønsted (B) and Lewis acid sites (L) in zeolites. Adapted from Fig. 6 in [35]

The acidity of zeolites comes from their structure. Figure 18.39 shows how one can picture these materials as a silica (SiO_2) superstructure, in which every so often an aluminum atom replaces a silicon atom. The silicon atoms have a valence of +4, and each one binds to four oxygen atoms. The oxygen atoms in hydroxyl groups are Brønsted acids (proton donators). Replacing Si (+4) with Al (+3) creates electron-rich Lewis acid sites, which are associated with positive counter-ion such as Na^+, K^+, $NH_4{}^+$, or H^+. The counter ions can be swapped via ion exchange. For example, when Na-Y zeolite is exchanged with an ammonium salt, the Na^+ ion is replaced by $NH_4{}^+$. When NH_4-Y is heated to the right temperature, the ammonium ion decomposes, releasing NH_3 (gas) and leaving behind highly acid HY zeolite.

ZSM-5 is a shape-selective zeolite made by including a soluble organic template in the mix of raw materials. (Templates for this kind of synthesis include quarternary ammonium salts.) ZSM-5 enhances distillate yields in FCC units and CDW in hydroprocessing units, where due to its unique pore structure, it selectively cracks waxy n-paraffins into lighter molecules.

Amorphous silica-alumina (ASA) catalysts are also used for hydrocracking. ASA catalysts are less active but better for producing middle distillates. Figure 18.40 provides a conceptual comparison of amorphous ASA catalysts with crystalline zeolite catalysts. The pore diameters of HY zeolite catalysts are uniform and relatively small—around 7.5 Ångstroms, give or take. Mesopores have larger diameters, which admit mid-sized molecules such as those found in vacuum gas oils. ASA catalysts include small, medium, and large pores. The width of ASA pores can exceed 100 Å. The larger pores can accommodate larger molecules, which explains why ASA catalysts do a better job of cracking feeds with very high endpoints. The pore diameter for ZSM-5 is 6.3 Å.

Catalyst Manufacturing

The following steps are used to prepare solid refinery catalysts:

- Precipitation
- Filtration (or centrifugation), washing, and drying
- Forming
- Calcining
- Impregnation and/or co-mulling
- Activation

Other steps, such as kneading, mulling, grinding, and/or sieving also may be used.

In the precipitation step, two solutions are combined to form a desired solid. For example, mixing a solution of aluminum nitrate [$Al(NO_3)_3$] with a solution of sodium aluminate [$Na_2Al_2O_4$] yields a gel of aluminum hydroxide [$Al(OH)_3$]. As the gel ages, tiny crystals grow larger and a pore structure starts to develop.

Filtration, washing, and drying remove salts and other impurities. In our example, after the precipitation of Al

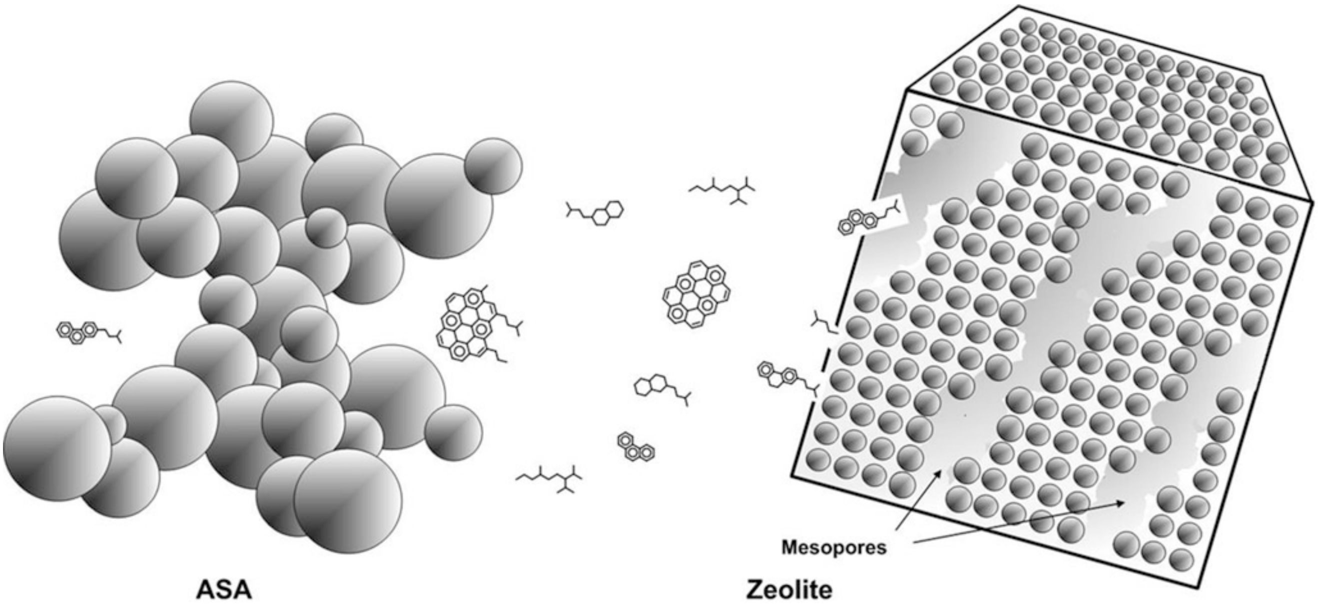


Fig. 18.40 Comparison of amorphous alumina silica (ASA) and zeolite cracking catalyst structures

$(OH)_3$ from aluminum nitrate and sodium aluminate, the coproduced sodium nitrate is washed away with water. Adding ammonium hydroxide expedites sodium removal. Subsequent drying removes excess water and initiates the transformation of $Al(OH)_3$ into alumina [Al_2O_3].

Catalysts can be shaped like wagon wheels, spheres, cylinders, hollow cylinders, lobed extrudates, pellets, etc. (Fig. 18.41). A cross-section of a lobed extrudate can look like a 3-leaf or 4-leaf clover without the stem. Compared to cylindrical extrudates, shaped extrudates have a higher surface-to-volume ratio, and the average distance from the outside of a particle to the center is shorter. To make extrudates, a paste is formed and forced through a die. The resulting spaghetti-like strands are dried and broken into short pieces with a length/diameter ratio of 2:4; for main-bed hydrotreating catalysts, diameters range from 1.3 to 4.8 mm. The particles are calcined, which hardens them and removes additional water and volatile molecules such as ammonia.

Spherical catalysts are made by (a) spray-drying slurries of catalyst precursors, (b) spraying liquid onto powders in a tilted rotating pan, or (c) dripping a silica-alumina slurry into hot oil. Pellets are made by compressing powders in a dye. FCC catalysts are made by spray drying.

Impregnation distributes active metals within the pores of a catalyst support. Like sponges, calcined supports are especially porous. Far more than 99 % of the surface area is inside the pores. When the pores are exposed to aqueous solutions containing active metals, capillary action pulls the aqueous phase into them. After drying, the catalyst might be soaked in another solution to increase the loading of the same (or a different) active metal. Catalysts can also be made by co-mulling active metal oxides with the support. Co-mulling tends to cost less because it requires fewer steps. It also produces materials with different activities—sometimes higher, sometimes lower—than impregnation.

Catalyst Deactivation

Over time, refinery catalysts deactivate and must be replaced. The major causes of deactivation are feed contaminants (trace metals, particulates, etc.) and catalyst coking; the latter is discussed above in some detail. In FCC, continuous catalyst replacement (CCR) processes, and ebullated-bed hydrocracking, aged catalysts are continuously removed and replaced with fresh. But in fixed-bed units, catalyst replacement requires a shutdown. Large hydrocrackers can process >40,000 barrels of feed per day. With a typical net margin of \$15–20 per barrel for hydrocracking alone, every day of down time costs \$600,000 to \$800,000. When the hydrocracker comes down, refiners may have to shut down (or reduce feed rates in) other units, such as the catalytic reformer, which gets feed from the hydrocracker (see Fig. 18.13). Therefore, the refinery-wide cost of hydrocracker down time can be far greater than \$1 million per day.

For a fixed-bed hydrotreating or hydrocracking unit, catalyst cycles range from 12 to 36 months. At the start of a cycle, average reactor temperatures are low, ranging from 550 to 660 °F (288–349 °C). As the cycle proceeds, the catalyst deactivates, so operators must raise temperatures

Fig. 18.41 Catalyst loading scheme showing size/shape grading. Photo provided by Criterion Catalysts & Technologies, LLP

to maintain HDS, HDN, or conversion. A catalyst cycle ends for one of the following reasons:

- The temperature required to achieve the unit's main process objectives hits a metallurgical limit
- The main process objective can be met only at reduced feed rate
- Pressure drop across the unit reaches the upper limit
- Selectivity decreases. If the production of light gases exceeds the capacity of one or more towers in the downstream gas plant, operators must decrease feed rate or reduce conversion. Both options are expensive
- Feed upsets. For a fixed-bed hydrotreater, a slug of residue can poison part of the catalyst with trace contaminants or foul it with particulates, asphaltenes and/or refractory carbon
- Equipment failure. Hardware problems occur most frequently in rotating equipment—pumps and compressors

Petroleum Products

Product Specifications

Products from modern refineries comprise several individual streams, which are blended to meet desired specifications. Product specifications developed by ASTM International are widely used throughout the world. (Until 2001, ASTM was an acronym for the American Society for Testing and Materials.) Table 18.16 lists some common refinery products along with their ASTM designations.

Analogous institutions in other countries, such as British Standards Institution (BSI), serve similar functions. In addition to setting specifications, these institutions develop and publish test methods used to analyze a wide variety of materials. ASTM and its international cousins cooperate both with each other and with government regulators. For example, recent low-sulfur gasoline and diesel directives from the U.S. Environmental Protection Agency are incorporated into D975-04 and D4814-03a, respectively.

Other widely used tests and specifications are defined by licensors. For example, UOP's *Laboratory Test Methods*, distributed by ASTM, defines several hundred procedures for analyzing catalysts, chemicals, and fuels.

Additives are essential components of finished fuels. They increase stability, improve flow properties, and enhance performance. Cetane-improvers are routinely added to diesel fuel, and additives that prevent intake-valve

Table 18.16 ASTM specification numbers for hydrocarbon fuels

Product	ASTM specification	Description
Gasoline	D4814	Standard specification for automotive spark-ignition engine fuel
Jet	D1655	Standard specification for aviation turbine fuels
Kerosene	D3699	Standard specification for kerosene
Diesel	D975	Standard specification for diesel fuel oils
Fuel oil	D396	Standard specification for fuel oils

deposits are now required in all grades of gasoline in the United States.

Gasoline. In the mid- to late-twentieth Century, making gasoline was a relatively simple task. If a mixture of components met specifications for volatility and octane, it could be shipped to retail outlets and sold as-was. If the octane was low, the problem could be fixed by adding a little tetraethyl lead. Butanes could be added or left out as needed to adjust volatility. In 2010, due to environmental regulations, refiners must also meet restrictions on Reid Vapor Pressure (RVP), sulfur content, and oxygen content. Thanks to such restrictions, especially the limit on RVP, the air is much more breathable in large American and European cities.

Octane numbers for hydrocarbons. In a spark-ignition engine, some compounds start to burn before they reach the spark plug. This premature ignition causes knocking, which reduces the power of the engine, increases engine wear, and can cause serious damage. Octane number is a measure of the propensity of fuels to knock in gasoline engines. It is based on a scale in which the octane number of n-heptane is zero and the octane number of isooctane (2,2,4-trimethylpentane) is 100. When a fuel is tested in a standard single-cylinder engine, mixtures of isooctane and n-heptane are used as standards. ASTM D2699 and ASTM D2700 describe methods for measuring research octane number (RON) and motor octane number (MON), respectively.

In North America, the pump octane of gasoline is the average of RON and MON: $(R + M)/2$. This is the number displayed on pumps at filling stations. Typical grades are "regular" with a pump octane of 87, "mid-grade" with a pump octane of 89, and "premium" with a pump octane of 91–93. In some locales, customers can dial in any octane they want between 87 and 93.

Several refinery streams have the right vapor pressure, boiling range, sulfur content, and octane to end up in the gasoline pool. Table 18.17 shows properties for blend stocks from which gasoline might be made. The raffinate comes from an aromatics extraction unit, and the pyrolysis gasoline comes from a nearby ethylene plant.

Table 18.17 Gasoline blend stock composition and properties: examples

Component	Density (kg/m^3)	RVP (bar)	Boiling range °C	°F	RON	MON
Butanes	0.575	3.6	−12 to −0.5	10.4–31	97	95
Straight-run Gasoline	0.64	1.15	27–80	81–176	80.2	76.6
Reformate	0.815	0.08	78–197	172–387	100.5	89.5
Raffinate	0.685	0.3	65–112	149–234	60.3	54.5
Heavy FCC Gasoline	0.76	0.2	43–185	109–365	90.5	79.5
Light FCC Gasoline	0.66	1.3	25–89	77–192	94.5	81.5
Pyrolysis gasoline	0.845	0.3	47–180	117–356	101.5	86.5
Alkylate	0.705	0.3	39–195	102–383	98	93.5
MTBE	0.746	0.5	48–62	118–144	115	97

Reformulated gasoline (*RFG*). In 1970, gasoline blending became more complex. The U.S. Clean Air Act required the phase-out of tetraethyl lead, so refiners had to find other ways to provide octane. In 1990, the Clean Air Act was amended. It empowered EPA to impose emissions limits on automobiles and to require reformulated gasoline (RFG). RFG was implemented in several phases. The Phase I program started in 1995 and mandated RFG for ten large metropolitan areas. Several other cities and four entire States joined the program voluntarily. In the year 2000, about 35 % of the gasoline in the United States was reformulated.

Tier 1 reformulated gasoline regulations required a minimum amount of chemically bound oxygen, imposed upper limits on benzene and RVP (RVP), and ordered a 15 % reduction in volatile organic compounds (VOC) and air toxics. VOC react with atmospheric NOx to produce ground-level ozone. Air toxics include 1,3-butadiene, acetaldehyde, benzene, and formaldehyde.

The regulations for Tier 2, which took force in January 2000, were based on the EPA Complex Model, which estimates exhaust emissions for a region based on geography, time of year, mix of vehicle types, and—most important to refiners—fuel properties. As of 2006, the limit on sulfur in the gasoline produced by most refineries was 30 wppm.

Initially, the oxygen for RFG could be supplied as ethanol or C_5–C_7 ethers. The ethers have excellent blending octanes and low vapor pressures. But due to leaks from filling station storage tanks, methyl-t-butyl ether (MTBE) was detected in ground water samples in New York City, Lake Tahoe, and Santa Monica, California. In 1999, the Governor of California issued an executive order requiring the phase-out of MTBE as a gasoline component. That same year, the California Air Resources Board (CARB) adopted

Table 18.18 Additives used in gasoline

Additive type	Function
Oxygenates anti-oxidation	Decrease emissions
	Minimize oxidation and gum formation during storage
Metal passivation	Deactivate trace metals that can accelerate oxidation
Corrosion inhibition	Minimize rust throughout the gasoline supply chain
Anti-icing	Minimize ice in carburetors during cold weather
IVD control (detergent)	Control deposition of carbon on intake valves
CCD control	Control deposition of carbon in combustion chambers

California Phase 3 RFG standards, which took effect in stages starting in 2002. The standards include a ban on MTBE and a tighter cap on sulfur content—less than 15 wppm.

In many other countries, including Finland, MTBE is still considered a premium blend stock, as safe as other components when filling station storage tanks are properly maintained.

Gasoline additives. Table 18.18 lists the kinds of additives used to prepare finished gasoline. Additive packages vary from season-to-season, region-to-region, and retailer to retailer. After-market additives contain similar types of ingredients, and usually are more concentrated. They are packaged to be added by consumers to their own vehicles.

Low-sulfur gasoline and ultra-low-sulfur diesel. In recent years, the U.S. Environmental Protection Agency (EPA) and the European Parliament promulgated clean-fuel regulations that are lowering the sulfur content of gasoline and diesel fuel. New sulfur-content standards for on-road diesel fuel range from 10 wppm in EU countries to 15 wppm in the United States.

Diesel fuel. The most important property of diesel fuel is cetane number, which indicates how the fuel will perform in diesel engines. Other important diesel-fuel properties include flash point, cloud point, pour point, kinematic viscosity, lubricity—and of course sulfur. Cloud point and pour point indicate the temperature at which the fuel tends to thicken and then gel in cold weather. Viscosity measures the tendency of a fluid to flow. In a diesel engine, viscosity indicates how well a fuel atomizes in spray injectors. It also measures its quality as a lubricant for the fuel system. Lubricity measures the fuel's ability to reduce friction between solid surfaces in relative motion. It indicates how the engine will perform when loaded. Table 18.19 lists cetane numbers for selected pure compounds. As with octane, blended cetane numbers can differ significantly from those for pure compounds.

Table 18.19 Cetane numbers for selected pure compounds

Compound	Type	Carbons	Formula	Cetane No.
n-Decane	Paraffin	10	$C_{10}H_{22}$	76
Decalin	Naphthene	10	$C_{10}H_{18}$	48
α-Methylnaphthalene	Aromatic	11	$C_{11}H_{10}$	0[a]
n-Pentylbenzene	Aromatic	11	$C_{11}H_{16}$	8
3-Ethyldecane	Paraffin (iso)	12	$C_{12}H_{26}$	48
4,5-Diethyloctane	Paraffin (iso)	12	$C_{12}H_{26}$	20
3-Cyclohexylhexane	Naphthene	12	$C_{12}H_{24}$	36
Biphenyl	Aromatic	12	$C_{12}H_{10}$	21
α-Butylnaphthalene	Aromatic	14	$C_{14}H_{16}$	6
n-Pentadecane	Paraffin	15	$C_{15}H_{32}$	95
n-Nonylbenzene	Aromatic	15	$C_{15}H_{24}$	50
n-Hexadecane (cetane)	Paraffin	16	$C_{16}H_{34}$	100[a]
2-Methyl-3-cyclohexylnonane	Naphthene	16	$C_{16}H_{34}$	70
Heptamethylnonane	Paraffin (iso)	16	$C_{16}H_{34}$	15[a]
8-Propylpentadecane	Paraffin (iso)	18	$C_{18}H_{38}$	48
7,8-Diethyltetradecane	Paraffin (iso)	18	$C_{18}H_{38}$	67
2-Octylnaphthalene	Aromatic	18	$C_{18}H_{24}$	18
n-Eicosane	Paraffin	20	$C_{20}H_{42}$	110
9,10-Dimethyloctane	Paraffin (iso)	20	$C_{20}H_{42}$	59
2-Cyclohexyltetradecane	Naphthene	20	$C_{20}H_{40}$	57

[a] Used as standards for ASTM D976

Diesel additives. Table 18.20 lists some of the additives used in diesel fuel and the reasons they are used.

Turbine fuel. Kerosene, jet fuel, and turbine fuel have similar boiling ranges. The key product properties are flash point, freezing point, sulfur content, and smoke point. The flash point is the lowest temperature at which a liquid gives off enough vapor to ignite when an ignition source is present. The freezing point is especially important for jet aircraft, which fly at high altitudes where the outside temperature is very low. Sulfur content is a measure of corrosiveness. The measurement of smoke point goes back to the days when the primary use for kerosene was to fuel lamps. To get more light from a kerosene lamp, you could turn a little knob to adjust the wick. But if the flame got too high, it gave off smoke. Even today, per ASTM D1322, smoke point is the maximum height of flame that can be achieved with calibrated wick-fed lamp, using a wick "… of woven solid circular cotton of ordinary quality."

Table 18.20 Diesel fuel additives

Additive type	Function
Anti-oxidation	Minimize oxidation and gum formation during storage
Cetane improvement	Increase cetane number
Dispersion	Improve behavior in fuel injectors
Anti-icing	Minimize ice formation during cold weather
Detergent	Control deposition of carbon in the engine
Metal passivation	Deactivate trace metals that can accelerate oxidation
Corrosion inhibition	Minimize rust throughout the diesel fuel supply chain
Cold-flow improvement	Improve flow characteristics in cold weather

Table 18.21 Turbine fuel specifications

Specification	Jet A	Jet B	JP-4	JP-5	JP-8
Flash point, °C (min)	38	–	–	60	38
Freeze point, °C (max)	−40 (Jet A) −47 (Jet A-1)	−50	−58	−46	−47
API gravity	37–51	45–57	45–57	36–48	37–51
Distillation, °C					
10 % max	205	–		205	205
20 % max	–	145	145	–	–
50 % max	–	190	190	–	–
90 % max	–	245	245	–	–
EP	300	–	270	290	300
Sulfur, wt% max	0.3	0.3	0.4	0.4	0.3
Aromatics, vol% max	22	22	25	25	25
Olefins, vol% max	–	–	5	5	5

The smoke point of a test fuel is compared to reference blends. A standard 40 %/60 % (volume/volume) mixture of toluene with 2,2,4-trimethylpentane has a smoke point of 14.7, while pure 2,2,4-trimethylpentane has a smoke point of 42.8. Clearly, isoparaffins have better smoke points than aromatics. Table 18.21 lists specifications for common turbine fuels.

Lube base stocks. Refiners prepare lube base stocks from residual oils by removing asphaltenes, aromatics, and waxes. Lube base stocks are hydrofinished, blended with other distillate streams for viscosity adjustment, and compounded with additives to produce finished lubricants. In the past, solvent-based technology was used to prepare lube base stocks. Propane deasphalting was used to remove asphaltenes. Furfural and related substances were used to extract aromatics, and MEK, or MIBK were used to remove wax. With the advent of CDW, some or all of these solvent-based methods can be replaced with hydroprocessing. CDW was developed by Mobil in the 1980s. The process employs ZSM-5, which selectively converts waxy *n*-paraffins into lighter hydrocarbons. The Isodewaxing Process, commercialized in 1993 by Chevron, reduces wax catalytically by isomerising *n*-paraffins into isoparaffins. Isodewaxing also removes sulfur and nitrogen, and it saturates aromatics. Products have a high viscosity index (VI), low pour point, and excellent response to additives.

Waxes. In a traditional lube plant, the raffinate from the solvent extraction unit contains a considerable amount of wax. To recover the wax, the raffinate is mixed with a solvent, usually propane, and cooled in a series of heat exchangers. Further cooling is provided by the evaporation of propane in the chiller and filter feed tanks. The wax forms crystals, which are continuously removed, filtered, and washed with cold solvent. The solvent is recovered by flashing and steam stripping. The wax is purified by heating with hot solvent, after which it is re-chilled, re-filtered, and given a final wash.

Making food-grade wax requires wax deoiling with solvents such as methyl-isobutyl ketone (MIBK).

Greases. Greases are made by blending salts of long-chained fatty acids into lubricating oils at 400–600 °F (204–315 °C). Antioxidants are added to provide stability. Some greases are batch-produced, while others are made continuously. The characteristics of a grease depend to a great extent on the counter-ion (calcium, sodium, aluminum, lithium, etc.) in the fatty-acid salt.

Asphalt. Asphalt can be produced directly from vacuum residue or by solvent deasphalting. Vacuum residue is used to make *road-tar asphalt*. To drive off light ends, it is heated to about 750 °F (400 °C) and charged to a column where a vacuum is applied to prevent thermal cracking.

In road-paving, the petroleum residue serves as a binder for aggregate, which can include stone, sand, or gravel. The aggregate comprises about 95 % of the final mixture. Polymers are added to the binder to improve strength and durability. The recommended material for paving highways in the United States is Superpave32 hot-mix asphalt. Superpave was developed in 1987–1993 during a US$50 million research project sponsored by the Federal Highway Administration.

Roofing asphalt. Roofing asphalt is produced by bubbling air through liquid asphalt at 260 °C (500 °F) for 1–10 h. During this "blowing" process, organic sulfur is converted to H_2S and SO_2. Catalytic salts such as ferric chloride ($FeCl_3$) may be used to adjust product properties and increase the rates of the blowing reactions, which are exothermic. To provide cooling, water is sprayed into the top of the blowing vessel, creating a blanket of steam that captures sulfur-containing gases, light hydrocarbons, and other gaseous contaminants. These are recovered downstream. Cooling water may also be sprayed on the outside of the vessel. The length of the blow depends on desired product properties, such as softening temperature and penetration rate. A typical plant blows 4–6 batches of asphalt per 24-h day. There are two primary substrates for roofing asphalt—organic (paper felt) and fiberglass. The production of felt-based roofing shingles consists of:

- Saturating the paper felt with asphalt
- Coating the saturated felt with filled asphalt
- Pressing granules of sand, talc, or mica into the coating
- Cooling with water, drying, cutting and trimming, and packaging

If fiberglass is used as the base instead of paper felt, the saturation step is eliminated.

Insert: the future for petroleum: ethanol in gasoline. In the petroleum industry, few topics are more controversial than ethanol in gasoline. Expert opinions provided during US Congressional hearings confirm the following: For every PhD, there is an equal and opposite PhD. Some facts are indisputable.

- In 2009, the United States and Brazil produced 87.8 % of the world's fuel ethanol, making 10.6 billion gallons and 6.6 billion gallons, respectively [36]
- In the United States, the advantages of putting bio-derived fuels into transportation fuels include cleaner air and less dependence on imported oil. Indeed, replacing 10–15 % of the gasoline in the United States with ethanol would be significant. Most sources say that corn can give 1.1–1.7 times more energy than is required for growing and processing [37]. This energy includes, of course, the production of ethanol. The unfermented "distillers grain" is high in fats and proteins and serves as an excellent animal feed. In contrast, sugar cane gives an energy gain of 8:1
- Thanks to ethanol from sugar cane, Brazil is now energy self-sufficient. Among large Western countries, it is the first to have a sustainable biofuel economy [38]. The eventual production of significant amounts of ethanol from sugar in the United States is unlikely, because it would require substantial investments in infrastructure. In China, the government is investing heavily in obtaining Biofuels from inedible vegetables (not including broccoli). Such fuels include ethanol [39] made from sweet sorghum, sweet potatoes, and cassava

(continued)

- Using corn ethanol for fuel increases food prices, increases consumption of fertilizer and pesticides, and puts pressure on water resources
- Ethanol increases the vapor pressure of gasoline. Therefore, to meet vapor-pressure specifications, refiners must remove butanes and even pentanes from the gasoline blend pool, creating periodic gluts of these compounds in certain areas. The gluts, of course, depress prices for these commodities
- In October 2011, the US government was paying an ethanol subsidy of 45 cents per gallon to oil refining companies and imposing a 54 cents per gallon tariff on imported ethanol. On June 16, 2011 the US Senate voted 73–27 to eliminate the subsidy. The vote did not become law due to White House opposition [40]
- Efficient production of ethanol from cellulose could answer many of the objections to ethanol from corn. Possible substrates are abundant. They include corn stover, switchgrass, miscanthus, and wood chips. The most promising routes involve multistage biological processes such as enzymatic hydrolysis and microbial fermentation. Research in this area is lively [41]

Protecting the Environment

Oil spills. In April 2010, the petroleum industry solidified its reputation for making deadly mistakes that cause titanic damage to the environment. The *Deepwater Horizon* blowout killed 11 people, injured 17 others, and released up to 4.9 million barrels of oil into the Gulf of Mexico. For additional details, please consult the Report to the President prepared by the National Commission on the BP Deepwater Horizon Oil Spill [42].

Previous oil spills include the following [43]:

> On March 16, 1978, the super tanker Amoco Cadiz ran aground on the Portsall Rocks three miles off the coast of Brittany. The wreck broke up completely before any of the remaining oil could be pumped out, so the entire cargo—more than 1.6 million barrels of crude oil—spilled into the sea. The resulting slick was 18 miles wide and 80 miles long. It polluted 200 miles of coastline, including the beaches of 76 Breton communities. The Amoco Cadiz incident caused more loss of marine life than any previous oil spill. Nearly 20,000 dead birds were recovered. About 9,000 tons of oysters died. Fish developed skin ulcerations and tumors. Years later, echinoderms and small crustaceans had disappeared from many areas, but other species had recovered. Evidence of oiled beach sediments can still be seen in sheltered areas, and layers of subsurface oil remain under many impacted beaches.

In March 1989, the super tanker *Exxon Valdez* ran aground in Prince William Sound in Alaska, spilling more than 250,000 barrels of oil into the Sound. Eventually, 33,000 birds and 1,000 otters died because of the spill. To clean up the mess, eleven thousand workers treated 1,200 miles (1,900 km) of shoreline, using 82 aircraft, 1,400 vessels, and 80 miles (128 km) of oil-containing booms. In response to the disaster, the US Congress passed the Oil Pollution Act of 1990. The Act streamlined and strengthened the ability of the U.S. Environmental Protection Agency (EPA) to prevent and react to oil spills. A trust fund, financed by a tax on oil, was established to pay for cleanups when a responsible party cannot afford to do so. The Act requires oil storage facilities and vessels to submit plans that tell how they intend to respond to large oil discharges. EPA published regulations for above-ground storage facilities, and the US Coast Guard published regulations for oil tankers. The Act also requires the development of area contingency plans to prepare for oil spills on a regional scale.

On January 25–27, 1991, during the occupation of Kuwait, Iraqis pumped 4–6 million barrels of oil—16 to 25 times more than the amount spilled by the Exxon Valdez—into the Arabian Gulf. On January 27, allied bombers stopped the spill by destroying the pumping stations. Ad Daffi Bay and Abu Ali Island experienced the greatest pollution. Despite the ongoing war, the clean up of the oil spill proceeded rapidly. Kuwaiti crude is rich in light ends, and water in the Arabian Gulf water is relatively warm. For these reasons, about half of the spilled oil evaporated, leaving behind a thick emulsion which eventually solidified and sank to the bottom of the sea. Another 1.5 million barrels were recovered by skimming. But on February 23–27, 1991, retreating Iraqi soldiers damaged three large refineries and blew up 732 Kuwaiti oil wells, starting fires on 650 of them. Up to six million barrels per day were lost between February 23 and November 8, 1991. Crews from 34 countries assembled to fight the oil-well fires. Initially, experts said the fires would rage for several years. But due to the development of innovative fire-fighting technology, the job took less than 8 months.

The oil-well fires burned more than 600 million barrels.

Air quality. In the 1970s and 1980s, environmental laws compelled refineries to reduce emissions of SO*x*, NO*x*, CO_2, and hydrocarbons. In the atmosphere, SOx reacts with water vapor to make sulfurous and sulfuric acids, which return to earth as acid rain. Volatile hydrocarbons react with NOx to make ozone. CO_2 is a major "greenhouse" gas. To reduce these pollutants, the industry tightened its operation by:

- Reducing fugitive hydrocarbon emissions from valves and fittings
- Removing sulfur from refinery streams and finished products
- Adding tail-gas units to sulfur recovery plants
- Reducing the production of NOx in fired heaters
- Scrubbing SOx and NOx from flue gases
- Reducing the production of CO_2 by increasing energy efficiency

Sulfur recovery. Conversion processes, hydrotreaters, and sweetening units remove chemically bound sulfur from petroleum fractions. The sulfur can end up as SOx, H_2S, NH_4SH, or NaSH.

Fuel-oil fired heaters and the regenerators of FCC units are major sources of refinery SOx and NOx emissions. The most obvious way to reduce SOx emissions from a heater is to use low-sulfur fuels. Unfortunately, although that solution requires no investment, it is probably the most expensive due to the relatively high cost of buying low-sulfur fuel oil and/or hydrotreating high-sulfur fuel oil.

A large fraction of the sulfur in the feed to an FCC unit ends up in coke on the catalyst. SOx are formed in the regenerator when the coke is burned away. Therefore, removing sulfur from the feed, usually by hydrotreating, decreases SOx emissions. FCC feed pretreating has other substantial benefits. Removing basic nitrogen decreases deactivation of acid sites on the FCC catalyst, which allows the FCC to reach a given conversion at lower temperatures. The saturation of aromatics in the feed pretreater provides the biggest benefit, because it converts hard-to-crack aromatics into easier-to-crack naphthenes. This alone can justify the installation of an FCC feed pretreater.

FCC SOx transfer additives. SOx transfer additives, first developed by Davison Chemical, react with SOx in the FCC regenerator to form sulfates. When the sulfated additive circulates to the riser/reactor section, the sulfate is reduced to H_2S, which is recovered by amine absorption and sent to the sulfur plant. In some units, these additives reduce FCC SOx emissions by more than 70 %. Consequently, if a pretreater or post-treater still must be installed, its size can be reduced. The chemistry of SOx transfer is summarized below:

FCC regenerator (oxidizing environment)

$$\text{Coke on catalyst (solid)} + O_2 + H_2O \rightarrow CO_2 + SO_2,\ SO_3,\ HySO_x\ \text{(gases)}$$

$$SO_2,\ SO_3,\ HySO_x\ \text{(gas)} + \text{MO (solid)} + O_2 \rightarrow MSO_4\text{(solid)}$$

FCC Riser-Reactor (reducing Environment)

$$\mathrm{MSO_4\,(solid) + 5H_2 \rightarrow MO\,(solid) + H_2S\,(gas) + 4H_2O\,(gas)}$$

Flue-gas scrubbing is a refiner's last chance to keep NOx and SOx out of the air. In wet flue-gas desulfurization, gas streams containing SOx react with an aqueous slurry containing calcium hydroxide $Ca(OH)_2$ and calcium carbonate $CaCO_3$. Reaction products include calcium sulfite ($CaSO_3$) and calcium sulfate ($CaSO_4$), which precipitate from the solution. NOx removal is more difficult. Wet flue-gas scrubbing removes about 20 % of the NOx from a typical FCC flue gas. To remove the rest, chemical reducing agents are used. As mentioned above, in the SCR process, anhydrous ammonia is injected into the flue gas as it passes through a bed of catalyst at 500–950 °F (260–510 °C). The chemical reaction between NOx and ammonia produces N_2 and H_2O.

When sulfur-containing feeds pass through hydrotreaters or conversion units, most of the sulfur is converted into H_2S, which eventually ends up in off-gas streams. Amine absorbers remove the H_2S, leaving only 10–20 wppm in the treated gas streams. H_2S is steam stripped from the amines, which are returned to the absorbers. The H_2S goes to the refinery sulfur plant.

Claus process. At the sulfur plant, H_2S is combined with sour-water stripper off-gas and sent to a Claus unit. Almost every refinery in the world uses some version of this process to convert H_2S into elemental sulfur. H_2S and a carefully controlled amount of air are mixed and sent to a burner, where about 33 % of the H_2S is converted to SO_2 and water. From the burner, the hot gases go to a reaction chamber, where the reactants and products reach equilibrium. In several units, the air is enriched with oxygen to increase plant capacity. Elemental sulfur is produced by the reversible reaction between SO_2 and H_2S. Ammonia comes in with the sour-water stripper off-gas. In the Claus process, it is thermally decomposed into nitrogen and water.

In the Claus burner, combustion temperatures reach 2,200 °F (1,200 °C). Much of the heat is recovered in a waste-heat boiler, which generates steam as it drops the temperature to 700 °F (370 °C). Next, the process gas goes to a condenser, where it is cooled to about 450 °F (232 °C). At this temperature, sulfur vapors condense, and the resulting molten sulfur flows through a drain to a heated sulfur-collection pit. At the bottom of the drain, a seal leg maintains system pressure and keeps unconverted gases out of the pit. Uncondensed sulfur and other gases flow to a series of catalyst beds, which recover additional sulfur by promoting the reaction between leftover H_2S and SO_2. With fresh catalyst and a stoichiometric gas composition, the cumulative recovery of sulfur across four condensers is about 50 %, 80 %, 95 %, and 96–98 %, respectively.

Tail-gas treating units (TGTU) bring the total sulfur recovery up to >99.9 %. Most tail-gas treating processes send the tail gas to a hydrotreater, which converts all sulfur-containing compounds (SO_2, SO_3, COS, CS_2, and various forms of S_x) into H_2S. In the SCOT process, offered by Shell Global Solutions, the H_2S is absorbed by an amine and returned to the front of the Claus furnace. In the LO-CAT® process, offered by Merichem, H_2S is air-oxidized to sulfur in an aqueous solution containing a chelated iron catalyst. For all tail-gas treatment processes, the last traces of unrecovered sulfur go to an incinerator, where they are converted into SO_2 and dispersed into the atmosphere.

Wastewater treatment. Wastewater treatment is used to purify process water, runoff, and sewage. As much as possible, purified wastewater steams are reused in the refinery. Wastewater streams may contain suspended solids, dissolved salts, phenols, ammonia, sulfides, and other compounds. The streams come from just about every process unit, especially those that use wash water, condensate, stripping water, caustic, or neutralization acids.

Primary treatment uses a settling pond to allow most hydrocarbons and suspended solids to separate from the wastewater. The solids drift to the bottom of the pond, hydrocarbons are skimmed off the top, and oily sludge is removed. Difficult oil-in-water emulsions are heated to expedite separation. Acidic wastewater is neutralized with ammonia, lime, or sodium carbonate. Alkaline wastewater is treated with sulfuric acid, hydrochloric acid, carbon dioxide-rich flue gas, or sulfur.

Some suspended solids remain in the water after primary treatment. These are removed by filtration, sedimentation, or air flotation. Flocculation agents may be added to consolidate the solids, making them easier to remove by sedimentation or filtration. Activated sludge is used to digest water-soluble organic compounds, either in aerated or anaerobic lagoons. Steam stripping is used to remove sulfides and/or ammonia, and solvent extraction is used to remove phenols.

Tertiary treatment processes remove specific pollutants, including traces of benzene and other partially soluble hydrocarbons. Tertiary water treatment can include ion exchange, chlorination, ozonation, reverse osmosis, or adsorption onto activated carbon. Compressed oxygen may be used to enhance oxidation. Spraying the water into the air or bubbling air through the water removes remaining traces of volatile chemicals such as phenol and ammonia.

Solid waste handling. Refinery solid wastes may include spent catalyst and catalyst fines, acid sludge from alkylation units, and miscellaneous oil-contaminated solids. All oil-contaminated solids are treated as hazardous and sent to sanitary landfills. Recently, supercritical extraction with carbon dioxide has been used with great success to remove oil from contaminated dirt.

The Future

Petroleum is one of mankind's primary sources of energy, so anything that affects global energy affects petroleum demand.

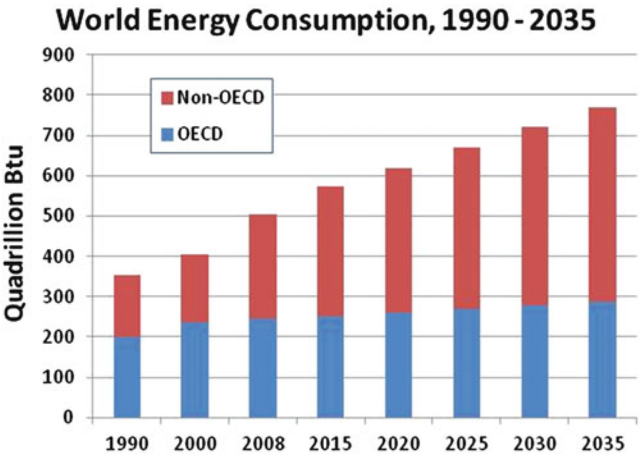


Fig. 18.42 Historical and projected world energy consumption, 1990–2035 [2]

Demand growth in China and India. According to Fig. 18.42, world consumption of energy grew from 350 to 500 quadrillion Btu between 1990 and 2008 and is expected to grow to 780 quadrillion Btu. Non-OECD countries [44], especially China and India, will account for most of the growth. Most likely, on a per capita basis, consumption in OECD countries will still exceed consumption in China and India by wide margin (Fig. 18.43).

Alternatives to petroleum. On average, conventional crude oil is getting harder to find. Over time, unconventional alternatives will supply more and more of our energy and petrochemical monomers. Liquid hydrocarbons will come from natural gas, biomass, bitumen, kerogen, and coal. Some of these contain daunting amounts of contaminants. Even biomass, which includes ethanol and vegetable oils, presents challenges, due the presence of high concentrations of oxygen and olefins. Compared to conventional oil, converting these materials will be costly and more difficult.

EIA's Annual Energy Outlook 2012 [2] projects supplies of unconventional liquids through 2035. The projections were based on three scenarios. In the high-oil-price scenario, future prices will reach $200 per barrel in 2012. In the reference scenario, which is based on present-day prices, the future price in 2035 reaches $150. In the low-price scenario, oil reaches $60 per barrel in 2035. As shown in Fig. 18.45, the largest expected increases come from bitumen and biofuel; biofuel is biomass that has been converted into fuels. In 2009, unconventional liquids supplied about 5 % of world demand. In

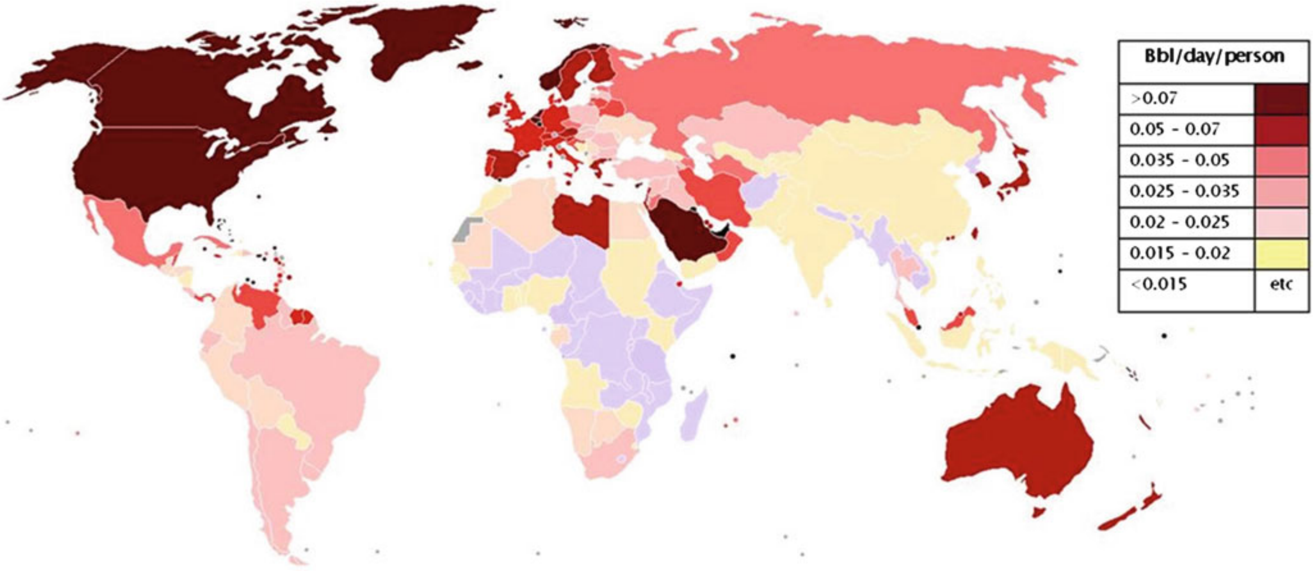


Fig. 18.43 Per capita energy consumption by country, 2009 [45]

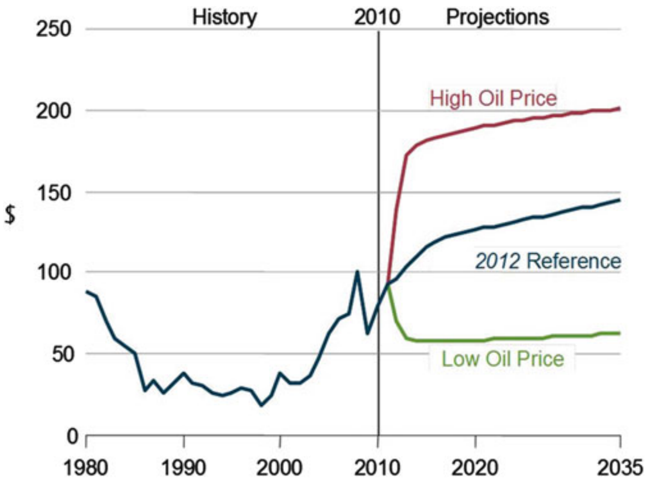


Fig. 18.44 Basis for the EIA forecast on the future use of unconventional liquids as replacements for conventional petroleum

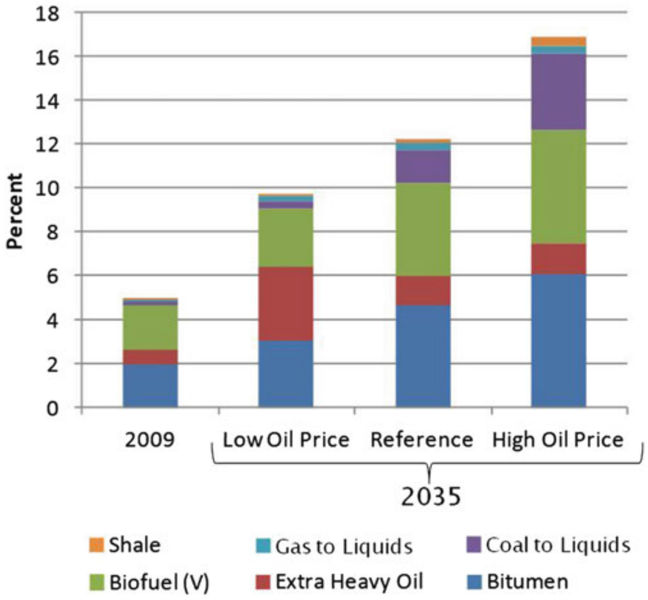


Fig. 18.45 Projected sources of unconventional hydrocarbon liquids in 2035, as a function of the scenarios shown in Fig. 18.44

EIA's high-oil-price scenario, by 2035 they will account for nearly 17 %. In the low-price scenario, by 2035 they will supply less than 10 %. The percentage of unconventional liquids will be higher at high oil prices, because the relative production cost (compared to conventional crude) will be lower.

Ethanol and other biofuels. Many countries now require the blending of ethanol or other oxygen-containing chemicals into gasoline. The oxygen content of fossil hydrocarbons is low. So is the oxygen content of most petrochemical monomers. In contrast, ethanol, cellulose, and vegetable oil contain a lot of oxygen, which must be removed before they can replace petroleum or natural gas in existing petrochemical plants. Removing oxygen is expensive and inefficient. Given enough time, we could develop a new bio-based chemical industry [46]. But in meantime, if we would decrease the amount of petroleum we burn, we would extend the time we could use it for polymers and plastics.

Adding 10 % ethanol to gasoline decreases the demand for petroleum, not by 10 %, but to some extent. This is because the production of ethanol requires energy. As mentioned, the estimated energy gain for ethanol from corn ranges from 1.1 to 1.7, while the energy gain for ethanol from sugar is 8:1 [36–38].

Automation. The ever-increasing development and use of better instruments and analyzers, engineering models, and real-time online optimization continue to improve the efficiency and safety of petroleum processing plants. In one case [47], implementing advanced process control increased production by 2.5 % at the Chemopetrol ethylene plant in Litvinov, Czech Republic. Other examples are provided on the web sites of companies such as Applied Manufacturing Technologies, Aspen Technology, Emerson Process Management, Honeywell Process Solutions, Invensys Operations Management, Yokogawa, and others.

Hampton and Robinson [48] provide an example of the use of advanced regulatory control to improve the safety of hydrocracking units.

Protection of the environmental. The push to control anthropogenic CO_2 emissions is gaining momentum (see the inset for additional comments). Government agencies around the world continue to tighten regulations on sulfur in fuels, air quality, wastewater quality, and the disposal of hazardous solids. Such regulations are now being adopted in less developed countries. They improve the quality of air, water, and land, but they also require investment.

Conservation. Conservation also decreases the demand for petroleum. As mentioned above, with modest investment in proven, widely applied technology, the United States could increase the efficiency of power production by 30 % [4]. Prices have a tremendous effect on energy conservation, which can be accomplished by turning off lights and driving less, but also due to innovation.

Encouraging energy conservation is an obvious, low-cost, win–win way to reduce CO_2 generation on a very large scale [4]. Other solutions, such as replacing coal-bed methane with CO_2, are technically sound, but their potential impact is minor. However, the sum of several minor solutions can be quite significant.

At present, it looks like the world is moving toward punitive measures, such as carbon taxes and/or cap-and-trade. The impact of these could be limited, because businesses with high emissions can avoid investing in energy efficiency by purchasing carbon credits [49].

Never underestimate the ingenuity of a hungry human.

Acknowledgement The author wishes to recognize the authors of previous editions of this chapter, S. Romanow and H.L. Hoffman, and to gratefully acknowledge that much of the content of this edition derives from their work.

References

1. International Energy Statistics (2011) United States Energy Information Administration. Washington, DC. http://www.eia.gov/cfapps/ipdbproject/IEDIndex3.cfm. Accessed Nov 2011
2. Annual Energy Outlook 2011 and Annual Energy Outlook 2012. United States Energy Information Administration. Washington, DC. www.eia.gov/forecasts/aeo/
3. BP Statistical Review of World Energy (2011) www.bp.com/statisticalreview.
4. Mims N, Bell M, Doig S (2009) Assessing the electric productivity gap and the U.S. efficiency opportunity. Rocky Mountain Institute, Snowmass, Colorado
5. To quantify the word “huge,” please refer to data presented by the United States Environmental Protection Agency, “Human-Related Sources and Sinks of Carbon Dioxide,” http://epa.gov/climatechange/emissions/co2_human.html
6. Web elements: the periodic table on the web. http://www.webelements.com/helium/ Accessed Nov 2011
7. Crude property data come from various sources. The values differ somewhat from source to source, in part because different laboratories get different results, and in part because the properties of commercial crudes change with time. Up-to-date data can be obtained from major oil companies, including government-owned companies such as the Kuwait Petroleum Company, PDVSA (Venezuela), Petronas (Malaysia) and Pertamina (Indonesia)
8. Robinson PR (2006) Petroleum processing overview (chapter 1). In: Hsu CS, Robinson PR (eds) Practical advances in petroleum processing. Springer, Hedielberg, Germany
9. Government of Alberta: energy. http://www.energy.alberta.ca/OurBusiness/oilsands.asp
10. United States Energy Information Administration (2011) http://www.eia.doe.gov/international/oilreserves.html. Accessed Nov 2011
11. Energy Density (2011) http://en.wikipedia.org/wiki/Energy_density. Accessed Nov 2011
12. Fenton DM, Hennig H, Richardson RL (1980) The chemistry of shale oil and its refined products In: Symposium on oil shale, tar sands and related materials. American Chemical Society Annual Meeting, San Francisco, Aug 1980
13. Speight JG (ed) (2008) Synthetic fuels handbook: properties, process, and performance. McGraw-Hill, New York
14. Altgelt KH, Boduszynski MM (1994) Composition and analysis of heavy petroleum fractions. Dekker, New York, Adapted from Figure 3. Used with kind permission from the publisher. This version of the figure also appeared as Figure 10.8 in Reference 15
15. Robinson PR (2011) Hydroconversion processes and technology for clean fuel and chemical production (chapter 10). In: Khan MR (ed) Advances in clean hydrocarbon fuel processing. Woodhead, Philadelphia
16. Source: JW Bunger & Associates, Salt Lake City, Utah. Used with kind permission from Dr. James W Bunger
17. Walters CC (2006) The origin of petroleum (chapter 2). In: Hsu CS, Robinson PR (eds) Practical advances in petroleum processing. Springer, New York
18. Conaway CR (1999) The petroleum industry: a nontechnical guide. Tulsa, PennWell
19. Finkl CW (ed) (1999) Encyclopedia of earth sciences series. p 454–456, Figure 6. Used with kind permission from Springer Images
20. Tarr BA, Flak L (2011) Underground blowouts. http://www.jwco.com/technical-litterature/p06.htm. Accessed Nov 2011
21. Duir JH, Randle AC, Reeg CP (1990) Unocal’s parachute creek oil shale project. Fuel Process Technol 25(2):101–117
22. U.S. heating oil crack spread widest in 24 months (2011) Reuters. http://in.reuters.com/resources/archive/in/20110126.html. Accessed 26 Jan 2011
23. Robinson PR (2006) Deactivation of hydroprocessing catalysts: thermochemistry of coke formation. International symposium on hydrocracking/hydrotreating technologies, 232nd American Chemical Society Annual Meeting, San Francisco, California, 10 Sept 2006
24. Sullivan RF, Boduszynski MM, Fetzer JC (1989) Molecular transformations in hydrotreating and hydrocracking. Energ Fuel 3:603
25. Scholl R, Seer C (1912) Abspaltung aromatisch gebundenen Wasserstoffs und Verknüpfung aromatischer Kerne durch Aluminiumchlorid. Justus Liebig’s Annalen der Chemie 394 (2):111–177
26. Van Speybroek V, Reyniers M-F, Marinm GB, Waroquier M (2002) Chem Phys Chem 3:863–870 and references cited therein
27. Satterfield CN, Smith CN (1986) Ind Eng Chem Process Des Dev 25:942
28. Greensfelder BS, Voge HH, Good GM (1949) Catalytic and thermal cracking of pure hydrocarbons. Ind Eng Chem 41(11):2573–2584
29. Mota CJA, Rowet R (1995) Mechanism of aromatic hydrocarbon formation in FCC naphtha. Ind Eng Chem Res 34(12):4326–4332
30. Lapinas AT, Klein MT, Gates BC, Macris A, Lyons JE (1991) Catalytic hydrogenation and hydrocracking of fluorene: reaction pathways, kinetics, and mechanisms. Ind Eng Chem Res 30(1):42–50
31. Krantz K (2003) Alkylation chemistry: mechanisms, operating variables, and olefin interactions. Stratco, Leawood, Kansas
32. Robinson PR, Dolbear GE (2006) Hydrotreating and hydrocracking fundamentals (chapter 7). In: Hsu CS, Robinson PR (eds) Practical advances in petroleum processing. Springer, New York
33. Motaghi M, Ulrich B, Subramanian A (2011) Slurry-phase hydrocracking: possible solution to refining margins. Hydrocarb Process 37–43
34. Schroder M (2010) Functional metal organic frameworks: gas storage, separation and catalysis. Top Curr Chem 293:175–205 (Used with kind permission from Springer Images)
35. Olah GA (1994) My search for carbocations and their role in chemistry. Nobel Lecture. 8 Dec 1994
36. World Fuel Ethanol Production (2011) Renewable Fuels Association. http://ethanolrfa.org/pages/World-Fuel-Ethanol-Production. Accessed 11 Oct 2011
37. Shapouri H, Duffield JA, Wang M (2005) The energy balance of corn ethanol: an update. U.S. Department of Agriculture, Office of the Chief Economist, Office of Energy Policy and New Uses. Agricultural Economic Report No. 813
38. Ethanol Fuel in Brazil (2011) http://en.wikipedia.org/wiki/Ethanol_fuel_in_Brazil#cite_note-Wilson-2. Accessed 11 Oct 2011
39. Zhenhong Y (2006) Utilization of ethanol and biodiesel in today and future. International conference on biofuel—China. http://martinot.info/Yuan_WBS2006.pdf

40. Schor E (2011) Senators reach deal on ethanol subsidy repeal, Urge Swift Congressional action for maximum benefit. The New York Times 7 July 2011
41. Cellulosic ethanol. http://en.wikipedia.org/wiki/Cellulosic_ethanol#Cellulolytic_processes
42. Executive Agency Publications. Deep water: the Gulf Oil disaster and the future of offshore drilling, Report to the President. http://www.oilspillcommission.gov, also available through Amazon and Barnes and Noble
43. Robinson PR, Shaheen EE, Shaheen EI (2006) Environmental pollution control (chapter 14). In: Hsu CS, Robinson PR (eds) Practical advances in petroleum processing. Springer, New York
44. OECD is the Organisation for Economic Co-operation and Development. http://www.oecd.org/about/membersandpartners/ It includes 34 of the world's wealthiest countries, including Australia, Canada, Chile, members of the EU, Israel, Japan, Mexico, New Zeeland, South Korea, Turkey and the United States
45. http://en.wikipedia.org/wiki/File: OilConsumptionpercapita.png. Accessed 6 Oct 2012
46. Designing the next generation of bio-refineries: the EuroBioRef project gets underway. CHEManager. 1 Mar 2010
47. Chemopetrol Increases Ethylene Production Capacity with AspenTech Manufacturing Solution. http://www.aspentech.com/uploadedFiles/News/Press_Releases/Pre-2004/pr4-17-02.htm
48. Hampton GW, Robinson PR (2011) Controlling hydrocracker temperature excursions. NPRA Q&A and Technology Forum, PD-11-01, San Antonio, Texas, 9–12 Oct 2011
49. Emissions Trading. http://en.wikipedia.org/wiki/Emissions_trading

Coal Technology for Power, Liquid Fuels, and Chemicals

19

Burtron H. Davis and James Hower

Introduction

The United States contains about 23% of the world's coal reserves, with Russia, China, Australia, and India each having more than 5% of the reserves. Coal represents over 90% of US proven reserves of fossil fuels. Recoverable reserves of US coal are estimated to be 227 Gt (billion tons; all numbers quoted are in metric tons). Bituminous coals (with a heating value of 23.26–34.89 MJ/kg) comprise nearly one-half of total US coal reserves. Eastern US coals are generally bituminous. Western and southwestern US coals are mainly subbituminous (with a heating value of 20.93–27.91 MJ/kg) and lignite (with a heating value of 18.61–23.26 MJ/kg). Coal is a major source of energy for electric power production and process heat and can serve as a source of synthetic fuels and feedstock for the petrochemical industry.

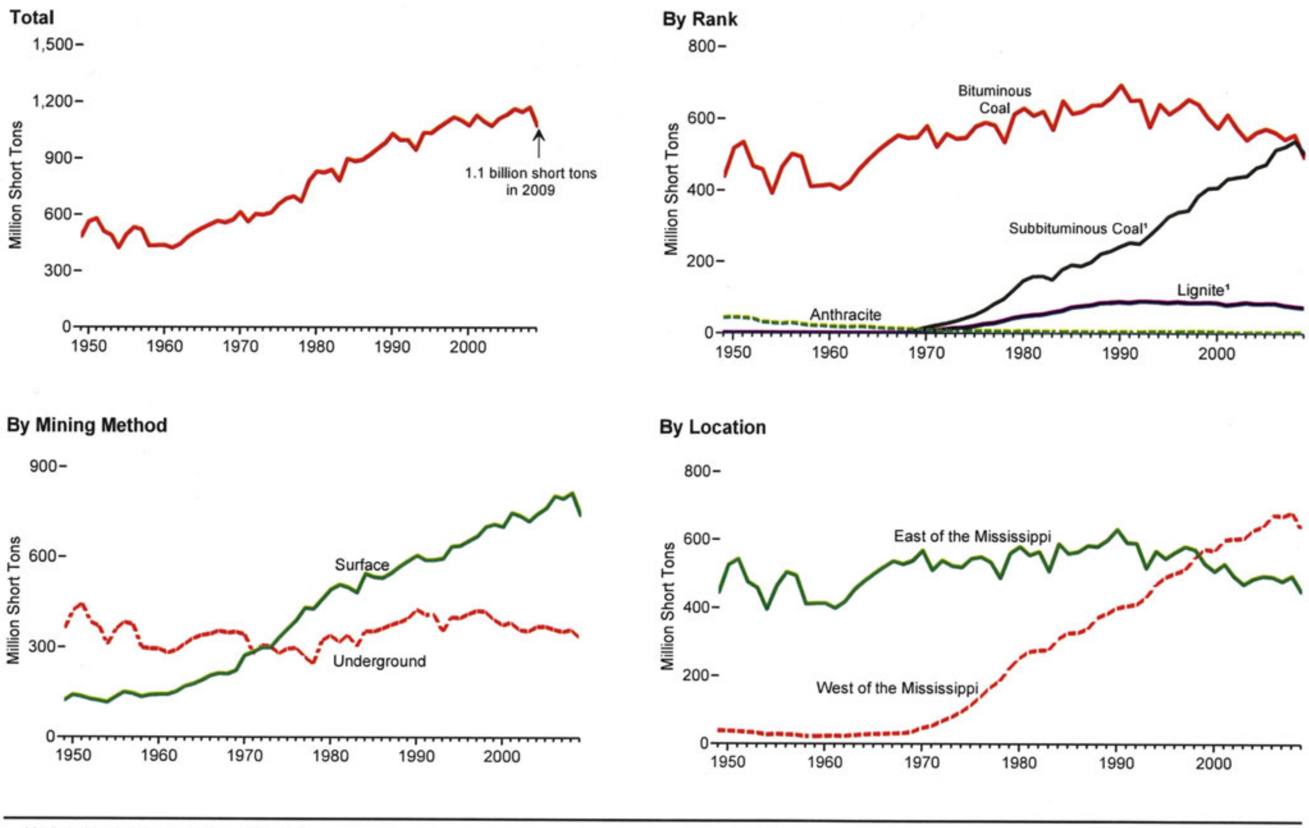
Historically, interest in coal technology has been cyclic, depending primarily on the price of petroleum. Coal dominated the US energy picture until the discovery of petroleum. Each subsequent threat to oil supplies spurred is new interest in coal, which rapidly faded as major new oil discoveries increased the supply of oil and lowered its price. The threat to oil supplies in 1973 produced the historic response, generating increased interest in converting coal to more convenient, cleaner fuels. Coal use for electric power generation has increased, but the economic conversion of coal to clean transportation fuels still stands as one of the key technical challenges of our time.

Bituminous and subbituminous coals account for the major share of all coal produced in the United States. In 2000, production of all types of coal totaled almost 1 Gt, of which about 95% was bituminous or subbituminous (Fig. 19.1). Despite its clean burning characteristics, anthracite, most of which is mined in northeastern Pennsylvania, accounts for a diminishing share of total coal production. Figure 19.2 presents data on coal consumption in the United States. Environmental concerns have led to an increased demand for low-sulfur coal, which is concentrated in the West. Also, surface mining, with its higher than average productivity, is much more prevalent in the West than in the East. Although more coal is mined east of the Mississippi than west, the share of western production has increased significantly since 1965, when production of western coal was only 24.5 Mt (5% of the total). By 1999, western production had increased to 518 Mt (52% of the total). In the period from 1990 to 1998, medium- and high-sulfur coal production decreased from 635 to 577 Mt.

Electric utilities are the dominant consumers of coal with consumption growing from 76 Mt in 1949 (a 17% share) to 892 Mt in 2000 (a 90% share). The percentage of total consumption in 2000 for all other sectors was lower than it was in 1949. The most dramatic declines occurred in the transportation sector (as railroads switched to diesel engines) and the residential and commercial sectors (transitions from coal to oil and gas for heating). In 1949, residential and commercial use accounted for 170 Mt (39% of the total), but in 2000, consumption in these sectors totaled less than 5.5 Mt, less than 1% of US coal usage. Consumption by the industrial sector trended downward after the mid-1960s, falling to about 59 Mt in 2000 from 186 Mt in 1966. As Table 19.1 shows, the United States is the second largest producer of coal in the world, trailing China. Average productivity of all types of mines in the United States has increased each year since 1949, reaching 3.36 t per miner per hour in 1989. In 1998, average productivity reached 5.5 t per miner per hour, largely due to increased production from large western surface mines and increased use of longwall mining in the eastern United States. Continued improvements in mine productivity (averaging 6.7% a year since 1978) are projected to cause falling real mine mouth prices. Higher electricity demand and lower prices, in turn, yield increasing coal demand. However, this demand is subject to a fixed sulfur emissions cap set by the

B.H. Davis • J. Hower (✉)
Center for Applied Energy Research, University of Kentucky,
2450 Research Park Drive, Lexington, KY 40511-8479, USA
e-mail: burtron.davis@uky.edu; james.hower@uky.edu

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_19,



¹ Included in bituminous coal prior to 1969.

Fig. 19.1 Coal production, 1949–2009

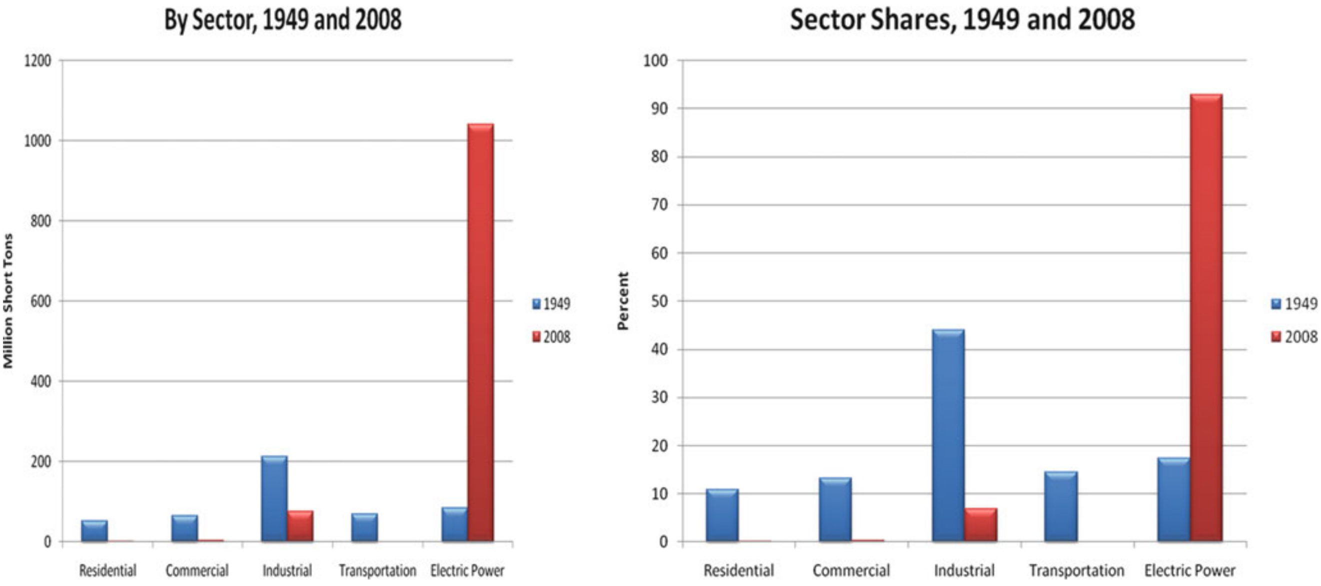


Fig. 19.2 US coal consumption by sector

Table 19.1 Top coal-producing countries, 2008

Country	Short tons	Metric tons	Quad BTU	PJ
China	2847.98	2583.63	57.37	54.38
United States	1171.48	1062.75	23.86	22.61
India	568.32	515.57	9.36	8.87
Australia	4358.51	397.80	8.99	8.52
Russia	356.19	323.12	6.60	6.25
Indonesia	313.23	284.16	7.15	6.78
South Africa	259.60	235.50	5.50	5.21
Germany	214.35	194.45	2.01	1.91
Poland	157.88	143.23	2.40	2.27
Kazakhstan	119.81	108.69	1.99	1.89
Colombia	86.66	78.61	2.13	2.02
Turkey	83.51	75.75	0.70	0.66
Canada	75.07	68.11	1.58	1.50
Greece	72.44	65.72	0.33	0.31
Czech Republic	66.35	60.19	0.71	0.67
Ukraine	65.74	59.64	1.29	1.23
Vietnam	44.03	39.94	0.89	0.84
North Korea	41.35	37.51	0.77	0.73
World	7271.42	6596.49	137.07	129.91

US EIA, http://tonto.eia.doe.gov/cfapps/ipdbproject/IEDIndex3.cfm?tid=1&pid=7&aid=1

Clean Air Act Amendments of 1990 (CAAA 90), which has led to progressively greater reliance on low-sulfur western coal. Anticipated US Environmental Protection Agency guidelines for mercury control will further complicate the mix of coal utilization.[1] The use of western coals can result in up to 85% reduction in sulfur emissions compared with many eastern coals. As coal demand grows, however, new coal-fired generating capacity is required to use the best available control technology: scrubbers and other advanced techniques that can reduce sulfur emissions by 90% or more. Further, the need to reduce Hg emissions has led to increased retrofitting of scrubbers. Thus, there will still be a market for low-cost, higher-sulfur coal.

The literature on the subject of coal composition, structure, and use has grown significantly during the last two decades, as interest in coal research has continued. The future of coal science looks bright as researchers continue to make significant contributions to the elucidation of the structure, composition, and physicochemical behavior of coal. New analytical techniques have made an important contribution to these advances. The objective of this chapter is to provide a brief review of the state of the art of coal science and technology.

Origin and Classification of Coal

Coal is a brown to black carbonaceous sedimentary rock composed of macerals (organic material) and minerals (inorganic material), both subject to metamorphism. Coal cannot be understood outside of the context of all three end members of its composition.

Coal originated from the remnants of various forms of plant life (mosses, ferns, shrubs, trees, etc.) that flourished in swamps and bogs millions of years ago during prolonged periods of warm temperatures and abundant rainfall. This accumulated organic matter, known as peat, was the precursor of coal. Although the peat may have been altered by bacterial and chemical action, overlying water excluded oxygen and retarded the rate of decay. This exclusion of oxygen from the peat was critical for preserving the organic matter. Over time, inorganic sediments accumulated on top of the deposits of plant debris. The gradual rising and lowering of the sea level over time led to deposits of peat alternating with deposits of sand and mud, resulting in the current situation of coal seams interlayered with rock layers.

As overlying sediments became thicker, the peat was compressed into a solid rock by the process of lithification. (Typically, 6 m of peat will form 1 m of coal.) This was followed by coalification, the complex progressive chemical changes that convert peat into coal. A major reaction occurring during coalification is the loss of oxygen (in the form of water) from the organic material. Thus, as coalification proceeds, the oxygen and hydrogen contents of the coal decrease, with a concomitant increase in carbon content. The length of time that these reactions proceed, the temperatures, and, to a lesser degree, the pressures to which the organic molecules are exposed determine the degree of coalification (maturity) of the coal. Coal rank is a measure of the degree of coalification; as coalification proceeds, coal rank increases from lignite through subbituminous and bituminous to anthracite.

Coal is an inhomogeneous rock, because different types of plant material behave differently under coalification conditions, and mineral matter is not uniformly distributed, either by type or by amount. Microscopic examination reveals this inhomogeneity by showing that different areas of the coal structure have a different appearance. The organic components of coal are classified into three maceral groups: vitrinite, liptinite, and inertinite. Each group includes a number of macerals; the vitrinites further divided

---

[1] The US Environmental Protection Agency's Clean Air Interstate Rule (CAIR) (http://www.epa.gov/CAIR/) and Clean Air Mercury Rule (CAMR) (http://www.epa.gov/camr/) would have dictated limits the amount of Hg emissions from coal-fired power plants. In 2008, CAIR and CAMR were vacated by the United States Court of Appeals District of Columbia Circuit (http://pacer.cadc.uscourts.gov/common/opinions/200807/05-1244-1127017.pdf; http://pacer.cadc.uscourts.gov/docs/common/opinions/200802/05-1097a.pdf), the latter upheld by the United States Supreme Court. In place of CAIR, the US EPA proposed the cross-state air Pollution Rule (CSAPR) on 6 July 2011, on 21 August 2012, the US Supreme court of appeal for the D.C. circuit nocated the ruling.

Table 19.2 Classification of coals by rank[a]

		Fixed carbon limits (%) (dry, mineral-matter-free basis)		Volatile matter limits (%) (dry, mineral-matter-free basis)		Calorific value limits (BTU/lb) (moist[b], mineral-matter-free basis)		
Class	Group	Equal or greater than	Less than	Greater than	Equal or less than	Equal or greater than	Less than	Agglomerating character
I. Anthracite	1. Meta-anthracite	98	–	–	2	–	–	–
	2. Anthracite	92	98	2	8	–	–	–
	3. Semianthracite	86	92	8	14	–	–	Nonagglomerating[c]
II. Bituminous	1. Low volatile bituminous coal	78	86	14	22	–	–	–
	2. Medium volatile bituminous coal	69	78	22	31	–	–	Commonly
	3. High volatile A bituminous coal	–	69	31	–	14,000[d]	–	Agglomerating
	4. High volatile B bituminous coal	–	–	–	–	13,000[d]	14,000	Agglomerating
	5. High volatile C bituminous coal	–	–	–	–	11,500	13,000	Agglomerating[e]
III. Subbituminous	1. Subbituminous A coal	–	–	–	–	10,500	11,500	Nonagglomerating
	2. Subbituminous B coal	–	–	–	–	9,500	10,500	–
	3. Subbituminous C coal	–	–	–	–	8,300	9,500	–
IV. Lignitic	1. Lignite A	–	–	–	–	6,300	8,300	–
	2. Lignite B	–	–	–	–	–	6,300	–

[a] This classification does not include a few coals, principally nonbanded varieties, that have unusual physical and chemical properties and which come within the limits of fixed carbon or calorific value of the high-volatile bituminous and subbituminous ranks. All these coals either contain less than 48% dry, mineral-matter-free fixed carbon or have a heating value of more than 15,500 moist, mineral-matter-free British thermal units per pound

[b] Moist refers to coal containing its natural inherent moisture but not including visible water on the surface of the coal

[c] If agglomerating, classify in low volatile group of the bituminous class

[d] Coals having 69% or more fixed carbon on the dry mineral-matter-free basis shall be classified according to fixed carbon, regardless of calorific value

[e] It is recognized that there may be nonagglomerating varieties in these groups of the bituminous class, and there are notable exceptions in high volatile C bituminous group

into the huminite group for lignite and subbituminous coal [1] and the vitrinite group for higher-rank vitrinite forms [2]. The first two of the maceral groups are reactive (relatively easily hydrogenated or oxidized). Inertinites [3], with the exception of the lower-reflectance semifusinite, are relatively inert. Most coals are banded, that is, show a banded structure. Besides banded coals, there are two types of dull, nonbanded coals: (1) cannels, which are rich in spores, and (2) bogheads, which contain abundant remains of algae.

Coals can be classified by rank and/or type. Table 19.2 presents a classification of coals in the United States according to rank. The parameters used to quantify rank (degree of coalification/metamorphism) are moisture, volatile matter (material that is volatilized when coal is heated at a certain temperature for a certain length of time), fixed carbon (the ash-free residue after the loss of moisture and volatile matter), heating value, caking tendency, and weathering properties. Vitrinite reflectance (a commonly used rank indicator for high-rank coals) is measured by comparing the intensity of a beam of light incident on a polished vitrinite surface with the light reflected directly back from that surface.

An international coal classification (Table 19.3) classifies high-rank coals according to their volatile-matter content, calculated on a dry, ash-free (daf) basis. Because volatile matter is not a suitable parameter for classifying coals containing more than 33% volatile matter, the calorific value on a moist, ash-free basis is included as a parameter for such coals. The resulting classes of coal, based on volatile-matter content and calorific value, are grouped according to their plastic properties by employing either the free-swelling or the Roga test. They are then placed in subgroups according to plastic properties, by using either the Audibert–Arnu or the Gray–King test. A three-digit code number is used to identify a coal; the first digit indicates the class of the coal, the second digit indicates the group, and the third digit indicates the subgroup.

In assessing the suitability of a coal for a particular purpose, the above classification of coal types must be used with caution. The heterogeneous nature of the wide range of plant materials and minerals that were the genesis of coal guarantees an almost unlimited range of chemical constituents within the various types of coal, and these constituents determine the coal's reactivity under specific conditions.

Table 19.3 International classification of coals by type (Courtesy of Bureau of Mines, US Department of the Interior)

Groups (determined by caking properties)			Code numbers												Subgroups (determined by coking properties)		
	Alternative group parameters															Alternative subgroup parameters	
Group number	Free-swelling index (crucible-swelling number)	Roga Index	The first figure of the code number indicates the class of the coal, determined by volatile-matter content up to 33% V.M. and by calorific parameter above 33% V.M. The second figure indicates the group of coal, determined by caking properties. The third figure indicates the subgroup, determined by coking properties												Subgroup number	Dilatometer	Gray–King
3	>4	>45							435	535	635				5	>140	>G_8
							334		434	534	634				4	>50–140	G_5–G_8
							333		433	533	633	733			3	>0–50	G_1–G_4
							332a	332b	432	532	632	732	832		2	≤0	E–G
2	2 ½–4	>20–45					323		423	523	623	723	823		3	>0–50	G_1–G_4
							322		422	522	622	722	822		2	≤0	E–G
							321		421	521	621	721	821		1	Contraction only	B–D
1	1–2	>5–20				212	312		412	512	612	712	812		3	≤0	E–G
						211	311		411	511	611	711	811		1	Contraction only	B–D
0	0–1/2	0–5		100		200	300		400	500	600	700	800	900	0	Nonsoftening	A
				A	B												
Class number→			0	1		2	3		4	5	6	7	8	9	As an indication, the following classes have an approximate volatile matter content of *Class 6*: 33–41% volatile matter; 7: 33–44% volatile matter; 8: 35–50% volatile matter; 9: 42–50% volatile matter		
Class parameters	Volatile matter (dry, ash-free) →		0–3	>3–10		>10–14	>14–20		>20–28	>28–33	>33	>33	>33	>33			
				>3–6.5	>6.5–10												
	Calorific parameter[a] →		–	–		–	–		–	–	>13.950	>12.960–13.950	>10.980–12.960	10.260–10.980			
Classes (determined by volatile matter up to 33% V.M. and by calorific parameter above 33% V.M.)																	

[a] Gross calorific value on moist, ash-free basis (30°C, 96% relative humidity) BTU/lb. Note: (1) Where the ash content of coal is too high to allow classification according to the present systems, it must be reduced by laboratory flow-and-sink method (or any other appropriate means). The specific gravity selected for flotation should allow a maximum yield of coal with 5–10% of ash. (2) 332a. . . >14–16% V.M., 332b. . . >16–20% V.M.

Coal Structure

From the above discussion, it should be clear that there is no such thing as pure coal. Because of variations in the organic and inorganic material that went into the formation of the coal, each seam will be different. Furthermore, there will be significant differences from point to point in the same seam. Composition and structure can vary dramatically between coals of different ranks or even between coals of the same rank. However, some structural features are common to all coals [4].

In general, at a molecular level, coals are composed of aromatic and hydroaromatic building blocks with a variety of substituent groups and some cross-linking between adjacent units [5]. Some of these units are not strictly planar because of the presence of heteroatoms (oxygen, nitrogen, or sulfur) and hydroaromatic portions. Graphite consists of parallel, mutually oriented layers of carbon atoms separated by a distance of 0.335 μm. The parallel stacking of layers in coal is similar but lacks mutual orientation between the layers, and the average spacing between layers is somewhat larger in coal than in graphite.

Low-rank coals are composed of small layers that are more or less randomly oriented and connected by cross-links. Thus, the structure is highly porous. Medium-rank coking coals show a greater degree of orientation and a greater tendency toward parallel stacking. These coals have fewer cross-links and fewer pores than lignites. High-rank coals (primarily anthracites) show a growth in the individual layers, a marked increase in the degree of orientation, and the development of a new type of pore that is elongated and parallel to the layers. These coals are the result of the highest degree of coalification.

The identification of the constituents of a complex mixture, such as coal, by molecular type may proceed in a variety of ways but generally consists of three types of analyses: chemical, spectroscopic, and physical. Mathematical formulas have been developed that use the properties measured by these techniques to derive structural parameters. Several good reviews on this subject are available [6, 7].

Coal Composition and Analyses

Coal is composed of carbon, hydrogen, oxygen, nitrogen, sulfur, chlorine, other elements in trace amounts, and mineral matter. The percentage of these elements varies considerably, depending on rank and even within coals of the same rank. For coals of different ranks, the largest variation in composition occurs in the carbon and hydrogen contents, while for coals of the same rank, sulfur and mineral matter exhibit the greatest variation.

The two most common coal analyses are the proximate analysis and the ultimate analysis. In addition, there are a number of miscellaneous analyses. Many of these tests are empirical, requiring strict adherence to specified conditions. In the United States, the American Society for Testing and Materials (ASTM) standardizes these methods. International standards are developed by the International Organization for Standardization (ISO).

A proximate analysis includes the determination of moisture, ash, volatile matter, and fixed carbon (by difference). The proximate analysis, together with total sulfur and heating value, is usually sufficient to characterize coal burned for process heating and steam generation. An ultimate analysis includes the determination of carbon, hydrogen, nitrogen, total sulfur, chlorine (sometimes), ash, and oxygen (by difference). Other analyses that may be conducted include not only standard tests, such as heating value, free-swelling index, and ash composition, but also special tests to determine trace elements, mineral phases, maceral types, surface properties, conversion potential, and plastic properties.

Standard analyses are described in ASTM and Bureau of Mines publications, but nonstandard methods must be acquired from publications dealing with the specific area of interest. For example, the plastic properties of coal, which are important considerations in many process designs, can be characterized by various techniques, each having its own advantages and disadvantages (Table 19.4).

Figure 19.3 shows typical heats of combustion, and Fig. 19.4 shows typical proximate analyses (reported on an ash-free basis) of various coal ranks. Analyses listed by coal bed, county, and state can be found in a number of publications, such as the data bank developed by the US Department of Energy and the Pennsylvania State University Coal Bank, Bureau of Mines bulletins, Illinois State Geological Survey circulars, and the Argonne National Laboratory Coal Data Base (for a limited number of coal samples).

The heating value (usually reported in MJ/kg or, in the US, in BTU/lb; MJ/kg = (BTU/lb)/429.92) given in Fig. 19.3 is the higher-heating value (HHV), also referred to as the gross heating value. This quantity is defined as the heat produced by combustion of a unit quantity of coal at a constant volume, under specified conditions, with all water produced being condensed. The lower heating value (LHV), also referred to as the net heating value, is calculated from the HHV by deducting 2.40 MJ for each kilogram of water originally present or formed during combustion. The heating value can be estimated from a number of formulas based on the ultimate analysis for carbon (C), hydrogen (H), oxygen (O), and sulfur (S). Some correlations include additional parameters.

Nitrogen and sulfur are the two primary heteroatoms in coal and tend to form environmental pollutants when coal is

Table 19.4 Summary of various techniques to characterize the plastic properties of coal

	Description	Advantages	Disadvantages
Free-swelling index/Gray–King	Visual comparison of cokes carbonized under well-defined conditions, with standard conditions	Inexpensive, rapid, reproducible	Lack of flexibility; transition phenomena from coals to coke not known
Hot-stage microscopy	Morphological changes observed by optical microscopy	Appropriate to supply the changes in macerals when heated	Provides qualitative information "subjective"
Foxwell/gas flow method	Measures the resistance of the gas flow through a bed of coal during continuous heating	Indirectly measures the fluidity of the plastic mass by the pressure differences between the inlet and outlet ports	Requires highly purified (oxygen-free) gas to avoid oxidation
Plastometer			
(a) Constant torque	(a) Records the angular velocity of a rotating shaft through a bed of coal as a function of temperature	Measures the changes in fluidity as a function of temperature; quite sensitive for determining extent of oxidation or weathering	Does not provide data on swelling properties, limited heating rates (−3°C/min)
(b) Variable torque	(b) Records the changes in torque required to rotate a cylindrical retort (when coal is carbonized) at constant speed as a function of temperature		
Dilatometer	Expansion, contraction, and volume change on resolidification measured as a function of temperature when heated	Provides data on dilation parameters and transition temperatures; wide range of heating rates (<100°C/min)	Does not provide quantitative information on changes in fluidity of coal

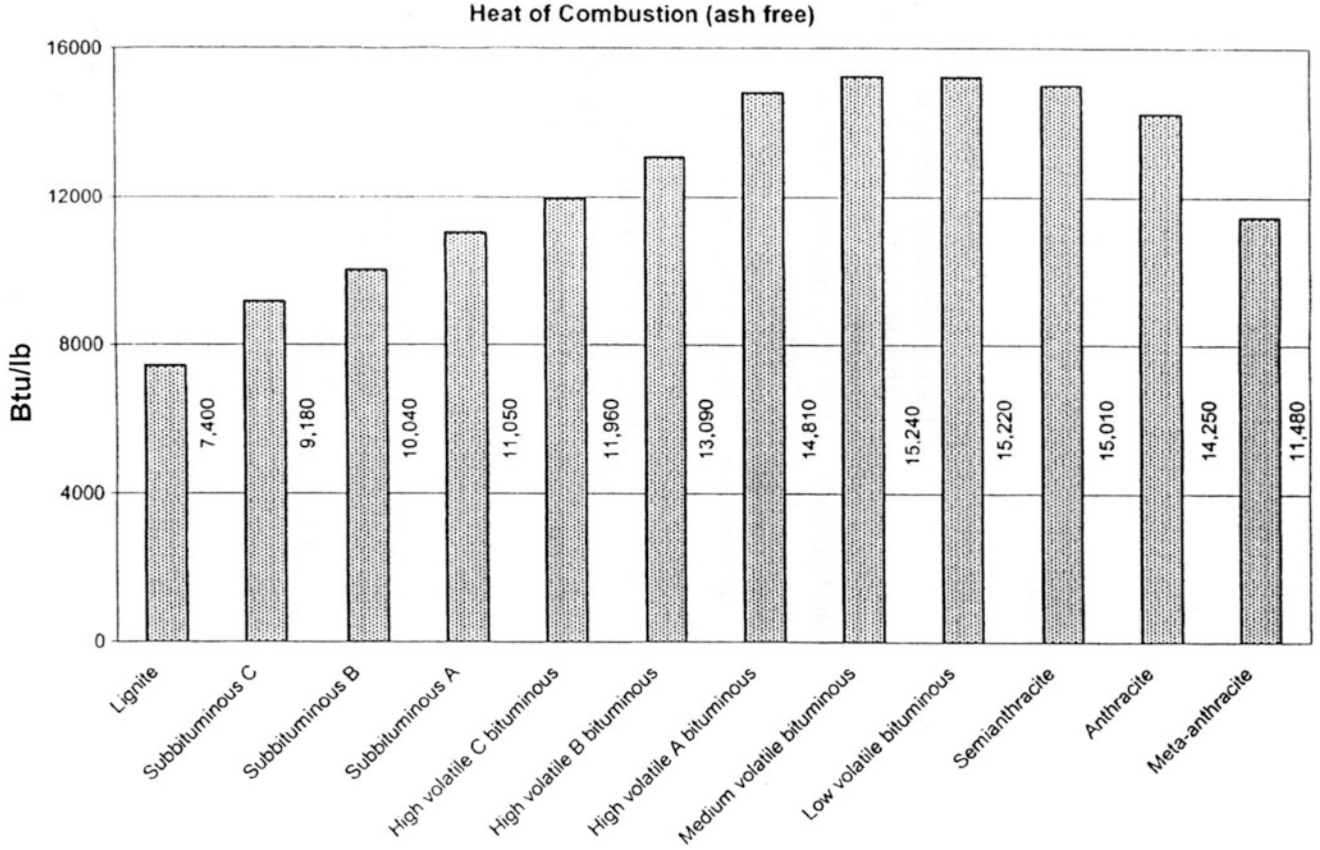


Fig. 19.3 Heat of combustion of US coal selected to represent the various ranks (Courtesy of Bureau of Mines)

processed. Therefore, the structure of sulfur and nitrogen compounds in coals, their evolution during thermal decomposition, and the distribution of these species in products have been extensively investigated. These studies suggest that nitrogen in coal is almost entirely contained in tightly bound ring structures. The nitrogen-containing structures

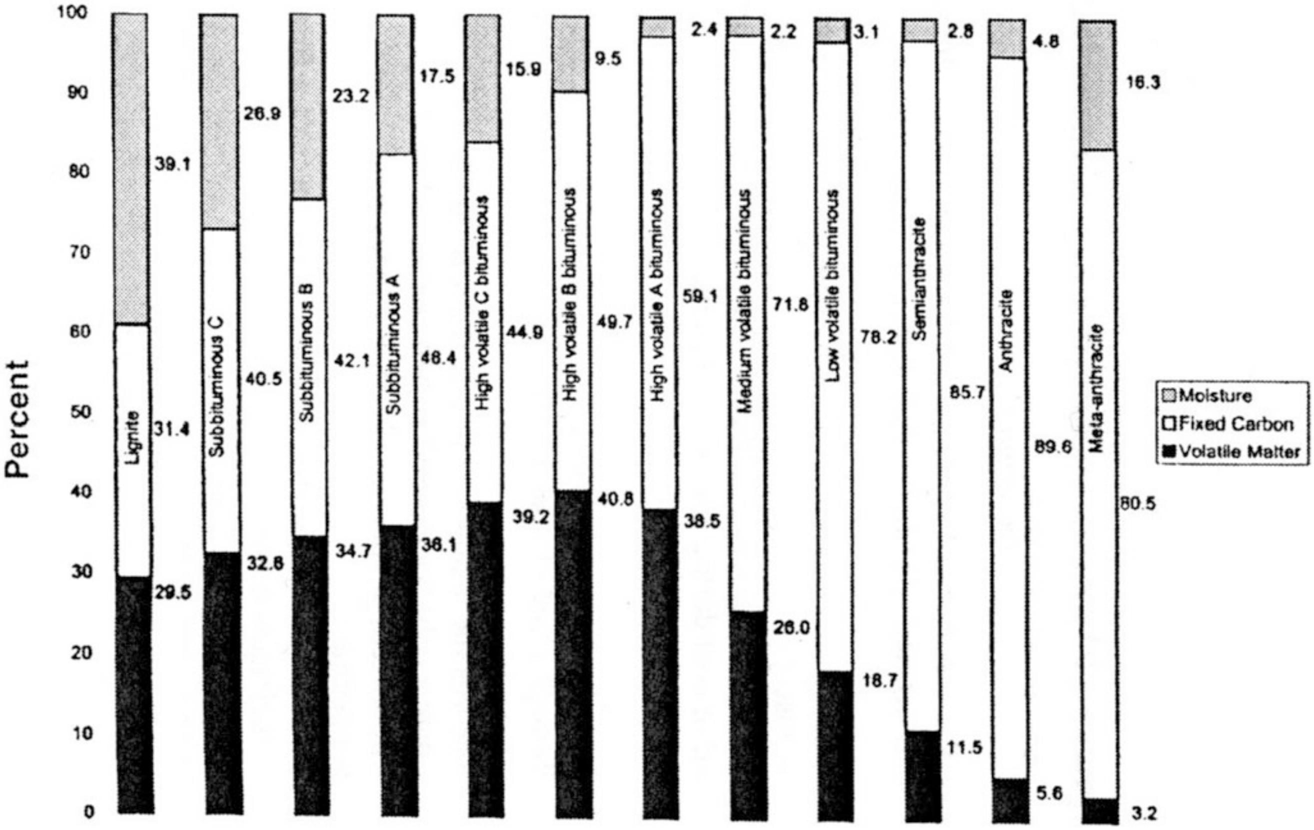


Fig. 19.4 Analysis of US coals selected to represent the various ranks (Courtesy of Bureau of Mines)

evolve from coal without cleavage during pyrolysis. Nitrogen compounds in coal can interfere with synthetic fuels processing owing to the fact that most of these compounds are basic and are strongly adsorbed on the acidic catalysts used, thus poisoning them. Nitrogen present in coal is converted to NO_x during combustion, the extent of the conversion depending on a variety of factors, including oxygen concentration, flame temperature, residence time at temperature, and the degree of fuel–air mixing.

Unlike nitrogen that is almost entirely associated with the organic structure of coal, sulfur is present in both the organic and inorganic constituents. Inorganic sulfur is present in coal in several forms, such as pyrite (FeS_2) and sulfates (small amounts). Organic sulfur is mainly present in ring compounds. Pyrite decomposes at a relatively low temperature (approximately 400°C), producing FeS, H_2S, and organic sulfur compounds. Organic sulfur compounds can be classified into three groups, those with loosely bound sulfur, those with tightly bound sulfur, and an intermediate category. The tightly bound sulfur compounds decompose above 500°C, while the loosely bound sulfur compounds decompose at low temperatures (about 100°C). Some of the tightly bound organic sulfur compounds are formed by incorporating sulfur from pyrite decomposition into ring structures.

Coal Mining and Preparation

Coal mines fall into two general classifications—surface and underground. Production of coal by type of mining for the United States is shown in Fig. 19.1. Most developed countries would use similar techniques.

Surface Mining

Surface mining techniques are used when the coal is present near the surface, and the overlying strata (overburden) are thin enough to permit economic removal. Surface mining techniques include contour mining, strip mining, and auger mining.

Contour mining is used in hilly areas, where the slope of the surface will permit only a narrow bench cut around the side of a hill; the excavation is backfilled immediately after the removal of coal. It is the only method that can be used on slopes of 15° or higher.

Strip mining is used in flat or gently rolling land, predominantly in the Midwest and West, where large, efficient equipment can be used. In this technique, the coal is exposed by removing the overburden. Blast holes are drilled, and

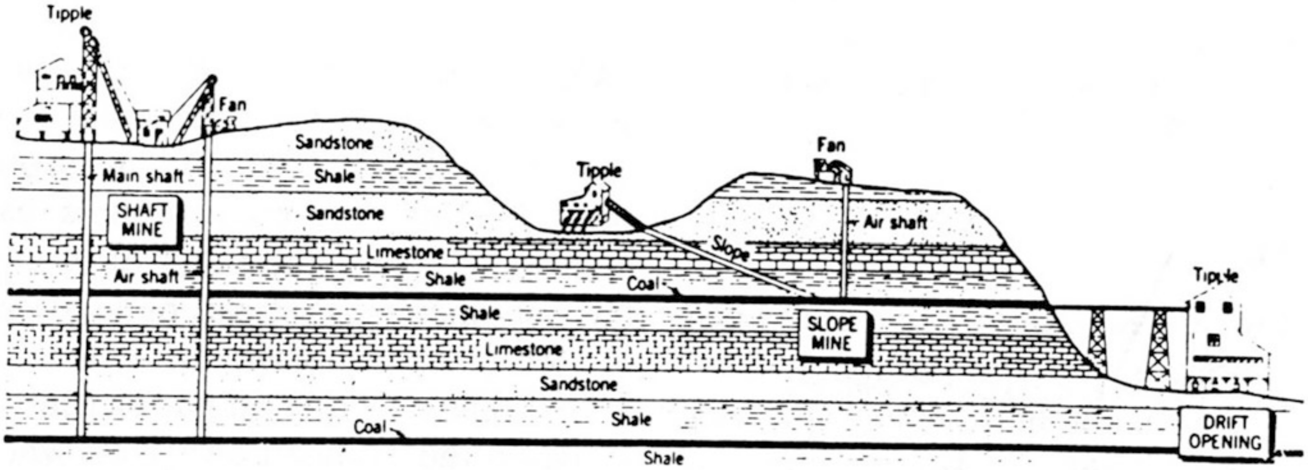


Fig. 19.5 Three types of entrances to underground mines—shaft, slope and drift (Courtesy of Bureau of Mines)

explosives are loaded into these holes to shatter the rock cover; earth-moving equipment is used to remove the soil and the shattered rock. The coal is then collected with power shovels or other coal-digging machines and loaded directly into trucks. When the seam is near the top of a hill, the entire hilltop may be removed to expose the coal. Strip mining can be very efficient, especially in the West, where coal seams can be 30 m thick or more. For example, the Wyodak seam averages nearly 20 m thick.

Auger mining is a supplementary method used to reach coal in stripped areas where the overburden has become too thick to be removed economically. Large augers are operated from the floor of the surface mine and bore horizontally into the coal face to produce some reserves not otherwise minable. This technique is frequently used to supplement coal recovery from contour mining.

Underground Mining

Underground mining is somewhat more labor-intensive than surface mining and is used to remove coal located too deep for surface mining, but here too, machines are used in most instances to dig, load, and haul the coal. Access to the coal seam is through a drift (horizontal passage), a slope, or a shaft (Fig. 19.5), depending on the location of the coal seam.

A drift mine is one that enters a coal seam exposed at the surface on the side of a hill or mountain. The mine follows the coal horizontally. A slope mine is one where an inclined tunnel is driven through the overlying rock to the coal, with the mined coal removed by conveyors or trucks.

A shaft mine is one where a vertical shaft is dug from the surface to reach the coal, which may be at great depth. The coal is then mined by horizontal entry into the seam, with the recovered coal hoisted to the surface through the vertical shaft.

Two general systems are used in underground mines: room-and-pillar mining and longwall mining.

In room-and-pillar mining, coal removal progresses in a nearly horizontal or low-angle direction by opening multiple stopes or rooms, leaving coal pillars to support the vertical load. This system recovers about 50% of the coal and leaves the mined area looking much like a checkerboard. It is used in areas where the overlying roof rock has geologic characteristics that provide good support. This system was used in old mines, where the coal was hand-dug. Two current methods for extracting the coal from the seam are the conventional method, where the coal is undercut and blasted free, and the continuous method, where a machine with a rotating cutter head moves along the coal face to extract the coal instead of blasting it loose. Roof control is the major problem for room-and-pillar mining. Ninety-eight percent of room-and-pillar mining is by the continuous method.

Longwall mining uses a machine that moves back and forth across the face of the coal seam in large rooms to remove up to 100% of the coal along faces of up to 305 m or more across. As coal is removed, the mining machine moves forward, and the roof is allowed to collapse behind. Coal recovery using this method is greater than in room-and-pillar mining, and the method can be used where roof conditions are fair to poor. Strong roof rock, however, can be a problem, since roof collapse is desired in this type of mining. The seam should be over 1.07 m thick to accommodate the large coal cutter or plow that is used, and a large reserve is necessary. The use of longwall mining has increased significantly in the past several years to the point where it accounts for 30–40% of underground production. Despite the increase in the use of longwall mining, subsidence remains a significant problem. Surface structures can suffer significant damage, and wells and streams can be disrupted.

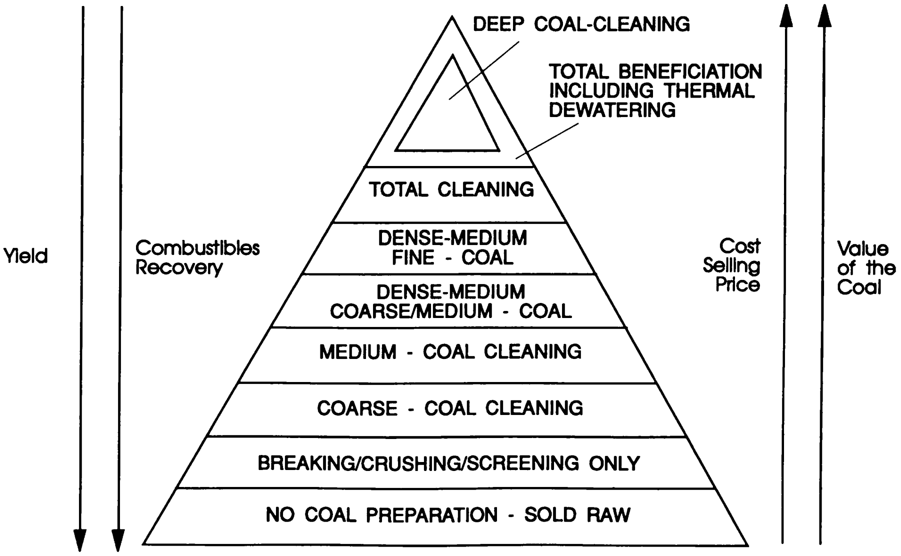


Fig. 19.6 Different levels of coal cleaning and the effect on coal recovery and economics

A modification of this method, using a continuous mining machine on faces up to 46 m long, is known as shortwall mining. It uses the roof support system of self-advancing chocks developed for longwall operations.

Coal Preparation

Coal preparation is a term applied to upgrading coal to make it suitable for a particular use; this includes blending and homogenization, size reduction, and beneficiation or cleaning. It is this last aspect, as well as the degree to which it is required, that most significantly governs the cost of coal preparation. Figure 19.6 shows levels of cleaning in terms of broad categories. Only about one-third of the 3.3 Gt of coal produced worldwide every year is at present cleaned by breaking, crushing, screening, and wet and dry concentrating processes. The purpose of coal cleaning [8, 9] is to remove mineral matter, consisting of rock, slate, pyrite, and other impurities, and to increase the organic content. Coal cleaning typically involves size reduction, particle sizing, cleaning, and dewatering. In general, cleaning processes utilize differences in density or surface properties between the organic and mineral matter in the coal. Depending upon the size of the coal particles being cleaned, cleaning techniques can be divided into three areas: coarse, medium, and fine. The vast majority of the coal cleaned commercially is separated by processes that rely on the density differential of the various components. Technologies that make separations on this basis are particularly effective for treating the coarse to intermediate sizes of coal (i.e., particle sizes ranging from greater than 10 cm to about 0.1–0.5 mm), with efficiency decreasing as particle size decreases. Centrifugal force is often applied to improve separation, particularly for finer particles. Density-based processes are particularly effective in removing pyrite from coal. With very few exceptions, coal cleaning is done using aqueous media. Figure 19.7 shows various coal-cleaning methods and the corresponding size ranges of the coal treated.

The first step in most coal-cleaning operations is size reduction, the main objective being to liberate mineral matter from organic matter. Typically, coal is crushed to 5–10-cm top size, and no grinding is performed. Size reduction equipment ranges from heavy-duty crushers and breakers, capable of crushing lumps up to a meter in size, to coal pulverization equipment, capable of milling coal to a fine powder. Grinding the coal finer liberates more mineral matter, but fine coal is more difficult and expensive to process. The ease with which coal can be ground (coal hardness) is measured in terms of the Hardgrove Grindability Index (HGI) [10]. A high HGI (80) indicates soft, easy to grind coal, and a low HGI (35) indicates difficult-to-grind coal. Details on size reduction can be found in various books on coal preparation [9, 11].

After the coal is crushed, it is generally screened to separate the raw coal into various sizes for cleaning operations. Screening of coal particles above 1.27-cm size usually is carried out dry. Double-deck vibrating screens commonly are used for this purpose. For sizing below 1.27 cm, wet screening is used, employing either high-frequency vibrating screens or the Sieve Bend. Classifying cyclones are used to size coals finer than 100 mesh.

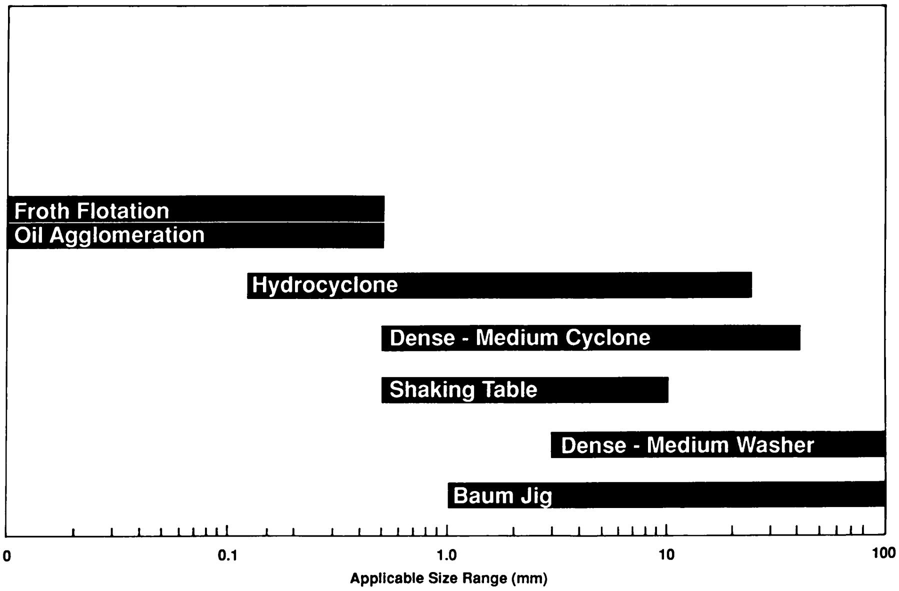


Fig. 19.7 Coal-cleaning equipment in common use in the coal industry with respect to the coal size processed

Coarse-Coal Cleaning

Coarse coal, which is usually the largest weight fraction in the coal feed to the cleaning plant, typically has a size greater than 1.9 cm. Jigging and dense-medium separation are the most frequently used coarse-coal cleaning techniques. In jigging, the separation of coal from mineral matter is accomplished via a fluidized bed created by a pulsating column of water, which produces a stratifying effect on the raw coal. The lighter coal particles rise to the top, overflow at the end of the jig, and are removed as clean product. The denser mineral matter settles and is removed as refuse. Dense-medium separations include processes that clean raw coal by immersing it in a fluid with a density intermediate between the density of clean coal and that of reject mineral matter. Most dense-medium washers use a suspension of fine magnetite in water to achieve a fluid with the desired density. Dense-medium systems are more expensive to install and operate than water-only systems, but their superior performance makes them particularly well suited for cleaning difficult-to-clean coals.

Medium-Coal Cleaning

Medium-size coal, which is usually the second largest weight fraction, includes coal ranging in size from 1.9 cm to 28 mesh (0.5 mm). The principal techniques used for medium-coal cleaning are dense-medium cyclones, hydrocyclones, spirals, and wet concentrating tables. All of these equipment types are widely used in the coal industry, with the dense-medium cyclone being the most efficient and prevalent. In a dense-medium cyclone, a mixture of raw coal and a dense medium (magnetite suspension) enters the cyclone tangentially near the top, producing free-vortex flow. The refuse is discharged through the underflow orifice, and the clean coal is discharged from an axial outlet located at the top of the cyclone. When only water is used in a cyclone for cleaning coal, it is called a hydrocyclone. The spiral separator is usually 2.44–3.05 m in height and consists of a trough going downward in a spiral. The coal slurry is fed in at the top, and, as the slurry follows the spiral down, centrifugal force separates the coal from the denser mineral matter. Wet concentrating tables, also known as shaking tables, work much like the classical miner's pan.

Fine-Coal Cleaning

Coal below 0.5 mm in size, classified as fine coal, typically makes up 5–20% of the feed to the cleaning plant. It is generally processed by methods such as froth flotation, which depend on surface properties. In froth flotation, the fine-coal slurry, to which a small amount of flotation agent (usually fuel oil and a short-chain alcohol, such as methyl isobutylcarbinol) is added, is processed through a flotation cell. In the cell, fine bubbles are generated by using either forced air or suction. The coal, being hydrophobic, attaches to the air bubbles and rises to the top, where it is removed as froth. The refuse, being hydrophilic, remains in the water and is removed from the bottom. The process is effective in recovering high-grade coal at moderate cost.

Conventional mechanical flotation cells have been widely used in the industry for many years and remain the primary means of cleaning coal particles smaller than 100 mesh. These units consist of a number of cells (usually 4–6) connected in series, each cell containing a mechanical agitator and aerator. The feed slurry, pretreated with frother and, usually, collector reagents, passes from one cell to the next with the most floatable coal being recovered in the first cell and the least floatable coal in the last cell.

A newly developed alternative to mechanical flotation cells is column flotation. Column flotation differs from conventional flotation in that columns have a much greater height-to-width ratio and do not require mechanical agitation to induce particle–bubble attachment. Columns are claimed to be more effective than conventional cells for beneficiating finer particles due to better control of bubble size (smaller is better), improved particle–bubble interaction, and froth washing capabilities. There are a number of similar column processes, with each possessing one or more unique technological advances. Among the most significant of these are Microcel™, Kenflote™, Flotaire, and the packed column. The Jameson Cell differs from other columns in that particle–bubble contact is achieved in a downcomer tube, where air and feed mixing occurs in a venturi-type system, rather than in the slurry column itself. The Microcel™ has been applied commercially for the recovery of fines from both freshly mined coal and refuse ponds and has also realized significant applications in the mineral industry.

In addition, improved separation of fine particles has resulted from recent advances in density-based separation techniques. This includes the development of several new water-only devices, such as various concentrating spirals, the Kelsey Jig, the Multi-Gravity Separator, the Falcon Concentrator, and the Knelson Concentrator. The development of advanced ultrafine dense-medium systems, such as Carefree Coal™ and MicroMag, also has the potential to improve cleaning of coal fines. However, none of these processes have yet achieved any significant degree of commercialization.

Oil agglomeration, a process that utilizes oil, or a similar hydrocarbon, to agglomerate coal and leave mineral matter in aqueous suspension, has been tested on pilot and commercial scales. Other processes (high-gradient magnetic and electrostatic cleaning) utilize differences in magnetic and electric charge properties of the mineral matter present in coal. None of these processes have achieved significant commercial success.

Conventional mechanical dewatering devices, such as vibrating screens and vibrating basket centrifuges, are effective in dewatering coarse and medium coal to moisture levels of less than 5%. Vibrating and high-speed vibrating screens are typically used to dewater coal larger than about 0.98 cm. For coal in the 0.98-cm × 30-mesh size range, vibrating basket centrifuges have proven to be highly effective and are widely used throughout industry.

One of the biggest problems in processing fine coal is dewatering and drying. Fine-coal dewatering is one of the most costly operations in the coal-cleaning plant, and, therefore, a significant number of operators bypass this operation by discarding coal fines directly to their waste slurry impoundments. Although these impoundments create a tremendous environmental liability and represent a loss of valuable coal resources, they still are widely used with an estimated 30–50 Mt of coal being added to ponds each year. Abandoned and active ponds in the United States are estimated to contain 2–3 Gt of coal.

Vacuum filtration and screen-bowl centrifuges are the principal unit operations used for mechanical dewatering of fine coal. Pressure filtration (plate and frame filter presses and continuous belt filter presses) is used to a much lesser extent to dewater fine coal and refuse. Sedimentation devices (thickeners, including conventional, high-rate, and multiplate units) are widely used to provide clarified water for recycle back to the plant and to preconcentrate dilute slurries (usually refuse fines) prior to dewatering or disposal. Thickener operation almost always requires the use of flocculants (primarily synthetic polymers or, to a lesser extent, modified starches) to promote particle settling.

The use of screen-bowl centrifuges has increased significantly in popularity to the point where now they are the predominant method for dewatering fine coal. Compared with vacuum filters, these units provide lower operating costs and increased moisture reductions; however, typically 10–15% of the coal entering the unit is lost to the effluent. These units typically produce filter cakes in the range of 15–20% moisture. Solid-bowl centrifuges occasionally are used on very fine coal, where coal losses using screen-bowl centrifuges would be excessive.

The primary form of vacuum filtration is the rotary-disk vacuum filter, which consists of multiple vertical disks, connected by a horizontal hollow shaft, that are partially submerged in open tanks containing the slurry to be dewatered. A vacuum is applied to the portion of the disk that is submerged to draw slurry toward the filter medium supported by the disk to form a filter cake on the surface of the filter as water passes through. As the disk rotates out of the slurry, the filter cake is dried by the continuing vacuum action, and a final product is produced by pressurized air blowback and mechanical scrapers. Flocculating agents may be added to the slurry prior to filtration to improve dewatering performance. These systems typically produce final product moistures in the range of 20–30% depending on the percentage of −200-mesh material in the feed.

Filtration devices that utilize high-pressure forces are capable of reducing moisture in the filter cake to about 20%, but capital and operating costs are high for these devices. Other newly developed techniques that have been tested on a pilot scale include an ultrahigh-g centrifuge, which generates forces up to 4,000 × *g*, and an electro-acoustic technique that utilizes the synergistic effect of electric, ultrasonic, mechanical, and surface chemical forces to remove moisture from the cake.

Chemical Coal Cleaning

Chemical cleaning processes utilize alkali or acid to leach out the impurities present in coal. Chemical methods are effective in removing both organic and pyritic sulfur. For example, the Molten Caustic Leaching Process is quite effective in its ability to remove both organic and pyritic sulfur, mineral matter, and trace elements from coal [12]. The process consists of treating coal with molten caustic (usually a eutectic mixture of sodium and potassium hydroxides) at 350–400°C for up to 4 h. The treated coal is washed with water and dilute acid to give a low-ash, low-sulfur product. The major drawbacks of chemical cleaning processes are economics and the need to use corrosive materials. The optimum coal desulfurization scheme may include both physical and chemical treatment.

Coal Utilization

The vast majority of the coal produced is burned to generate heat and power. The next most important use is for the production of metallurgical coke. Other uses are for the production of synthetic fuels for transportation and energy production, and as a raw material for the production of chemicals [13]. A unique feature of coal utilization is that, with a few exceptions such as Japan and Australia, the use of coal occurs in the country where it is produced (Fig. 19.8) [14]. This section discusses the technologies involved in these uses of coal.

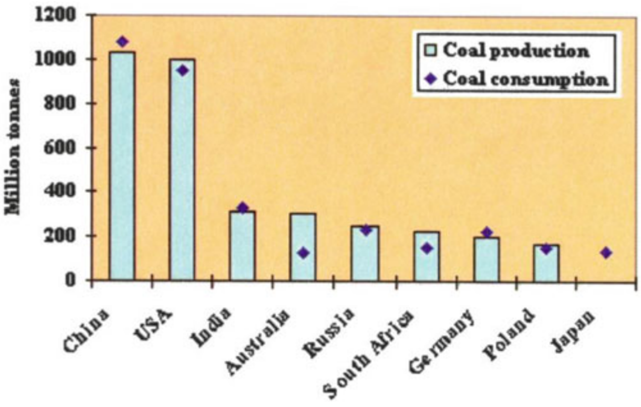


Fig. 19.8 Coal production and coal consumption (*filled diamonds*) in selected countries

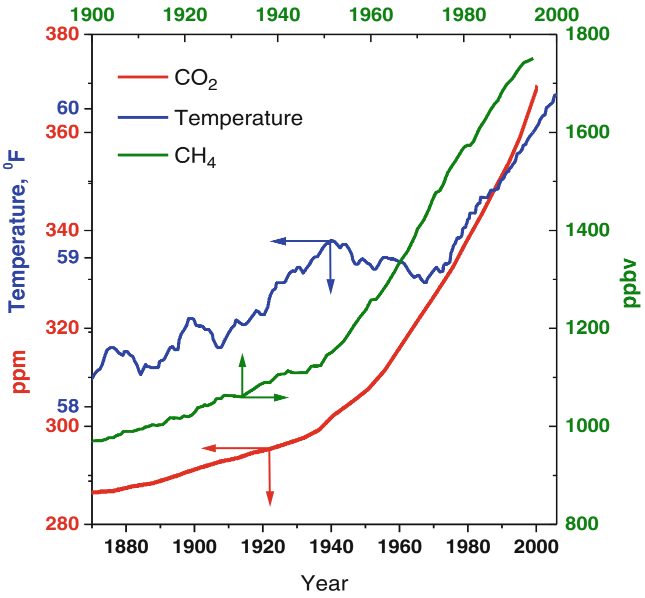


Fig. 19.9 Change in surface temperature, atmosphere carbon dioxide, and atmosphere methane during 1870–2000

Environmental Concerns Related to Coal Use

As discussed earlier, coal contains many of the elements in the periodic table, at least in trace amounts, and, during processing and/or combustion, these elements or their compounds can be released into the environment. Some of these harmful pollutants that can result from coal use are sulfur and nitrogen oxides, particulates, hydrogen chloride, mercury vapor, and a wide range of trace metals. In addition, most uses of coal result in the production of CO_2 which is considered the most important greenhouse gas from the point of view of the quantity in the atmosphere. At the same time, CO_2 is the feedstock for plant growth—a feature crucial to animal life. While CO_2 receives much attention because of the potential for global warming, there does not seem to be a direct relationship between the content of CO_2 in the atmosphere and the average surface temperature of the earth during the last 60 years (Fig. 19.9). As seen in Fig. 19.9, between 1940 and 1980, the temperature remained essentially constant, while the CO_2 concentration was increasing; then following 1980, the temperature rapidly increased, while the CO_2 concentration followed its steady increase. Thus, during this period, there is no direct relationship between the temperature of the lower atmosphere and the concentration of CO_2. In fact, a projection of the 1910–1950 period temperature increase leads to a projected temperature that is slightly higher than was actually reached.

Table 19.5 Summary of direct combustion coal process (Smoot)

Process type	Description	Extent of coal use in the United States (% of total used)	Commercial use	Scale size (TPD)[a]	Coal types	Coal size
Power station	Commercial electricity production	78–80				
Pulverized	Rapid burning of finely grained goal		Common	1,000–10,000	All	0.01–0.025 mm
Fluidized bed[b]	Well-stirred combustion		Pilot plant	2,000–8,000	All	0.15–0.6 cm
Stoker	Mechanically fed fixed bed		Small	100	Noncaking	1–5 cm
MHD	Combustion energy capture by magnetic fields		Laboratory	800–4,000		
Coal/oil mixture (COM)	Burning coal/oil mixtures in oil furnaces		Demonstration			
Industrial heat/steam	Industrial plant power	8–11				
Pulverized			Small	1–100		
Fluidized bed			Pilot	1–100	Same as above	
Stoker			Common	1–100		
COM			Demonstration	1–100		
Domestic/commercial	Hand-stoked space heating	1		0.005–0.05	Noncaking	3–10 cm
Transportation	Fuel for railroads	0.01–0.02		0.1–1	Noncaking	

Source: Excerpted from Smoot LD (1991) In: Bartok W, Sarafin AF (eds) Fossil fuel combustion. Wiley, New York. Copyright © 1991 by John Wiley & Sons, Inc. and used by permission of the copyright owner

[a] TPD = t/day

[b] Very adaptable, exhibits high heat transfer in bed as well as low-level pollutant products

Combustion

In the United States, approximately 90% of the coal consumed annually is burned in boilers to generate steam [15–17] to run turbines for electric power production. The burning of coal in small-scale combustion devices for direct heating is more significant outside the United States, so that, on a worldwide basis, about 55% of the coal produced is used for electricity generation.

Combustion Equipment

A summary of the most important direct combustion processes for coal is given in Table 19.5. Options for firing industrial and electric utility boilers are stokers, pulverized-coal-fired burners, and cyclone burners, the choice depending on the kind of coal and the amount of steam needed. Table 19.6 lists the types of coal that can be burned with the various combustion systems. Although good results usually will be obtained with the indicated coal and firing equipment, this table should be used only as a rough guide, and final equipment selection should be based on a sound engineering analysis.

Table 19.6 Suitability of various combustion systems for various coals

Fuel	Under feed	Stokers traveling or chain gate	Spreader	Pulverized-coal burner	Cyclone burner
Anthracite		X		X	
Bituminous					
17–25% volatile	X		X	X	X
23–35% volatile				X	X
Strongly coking	X	X	X	X	X
Weakly coking		X	X	X	X
Lignite		X	X	X	X

Electric utility and large industrial plants favor pulverized-coal-fired and cyclone-fired furnaces because of their inherent capability to effectively burn a range of coal types with varying qualities, their comparatively high availability, their quick response to load changes, and their extremely large steam-generating capacity. Large units can generate as much as 9.5 million pounds of steam per hour at 3,500 psig and 1,000°F. Smaller industrial installations typically favor stoker firing.

Stoker Firing. Stokers have been in operation for more than 150 years and are still used to fire small-capacity boilers, generally for steam production. There are several types of stokers, which are classified by the method of coal feeding and the relative flow directions of the fuel and air. The three main types are underfeed, crossfeed, and overfeed. Both fuel and air flow in the same direction in retort stokers; this arrangement is called underfeed burning. The fuel moves across the air direction in chain- or traveling-grate stokers; this design is called crossfeed burning. The spreader stoker approximates overfeed burning, the incoming fuel moving toward the air. Except for certain types of coal gasifiers, in which lump coal moves downward toward a grate against air (or oxygen and steam) coming through the grate, no conventional combustion system operates purely in the overfeed mode. Table 19.7 gives the approximate capacity range for each type of stoker.

An example of a crossfeed chain-grate stoker being used to heat a fire-tube boiler is given in Fig. 19.10. This stoker consists of an endless track of grate bars that pass over rotating sprockets at each end, similar to a conveyor belt. Coal is fed onto one end of the grate, using a hopper, to a depth of about 5 in. The coal burns as it moves along the grate, and, at the far end of the grate, only ash remains. The ignition of the coal is assisted by radiant heat from a refractory lining.

Table 19.7 Approximate capacity range of stokers

Type	Steam production (1,000 lb/h)	Grate heat release (1,000 BTU/ht/ft^2) (max)
Single retort	5–50	200
Multiple retort	40–300	300
Traveling or chain grate	10–300	300
Spreader	10–300	1,000

Pulverized-Coal Firing. The burner and furnace configurations for the main types of pulverized-coal firing (often called suspension firing) are shown in Fig. 19.11. There are some design variations among various vertical, impact, and horizontal suspension firing systems, but these schematic drawings serve to illustrate the general principles.

Coal-fired boilers are further differentiated by the operating model, either dry-bottom or wet-bottom (slagging), depending on whether the ash leaves the boiler as a solid or as a liquid. Most modern pulverized-coal-fired boilers are of the dry-bottom type, in which the ash is cooled before contacting any surface [18]. Wet-bottom firing provides a more convenient method of ash removal and disposal but has several disadvantages, including less flexibility in fuel selection, higher incidence of fouling and corrosion, and higher levels of NO_2 formation, that have led to its virtual elimination in the United States [19].

The first suspension-fired furnace in the United States was designed like the one shown for vertical firing in Fig. 19.11. Pulverized coal (about 70% through a 200-mesh screen) is transported to the burner with primary air, amounting to about 20% of the total air needed for complete combustion. The balance of the air, known as secondary air, is admitted through openings in the furnace wall. Because a large percentage of the total combustion air is withheld from the fuel stream until it projects well down into the furnace, ignition stability is good. This type of firing is well suited for coals that are difficult to ignite, such as those with less than 15% volatile matter. Although no longer used in central-station power plants, this design, with delayed admission of secondary air, may find favor again if low-volatile chars from various coal-conversion processes become available for burning for heat and power.

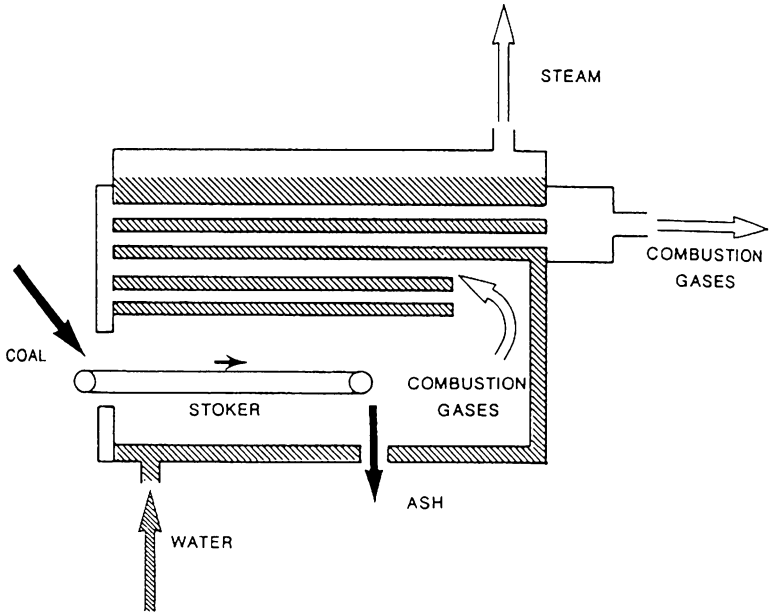


Fig. 19.10 Example of a chain-grate stoker equipped with a fire-tube boiler (reprinted with permission from Schobert [35], Copyright 1987, American Chemical Society)

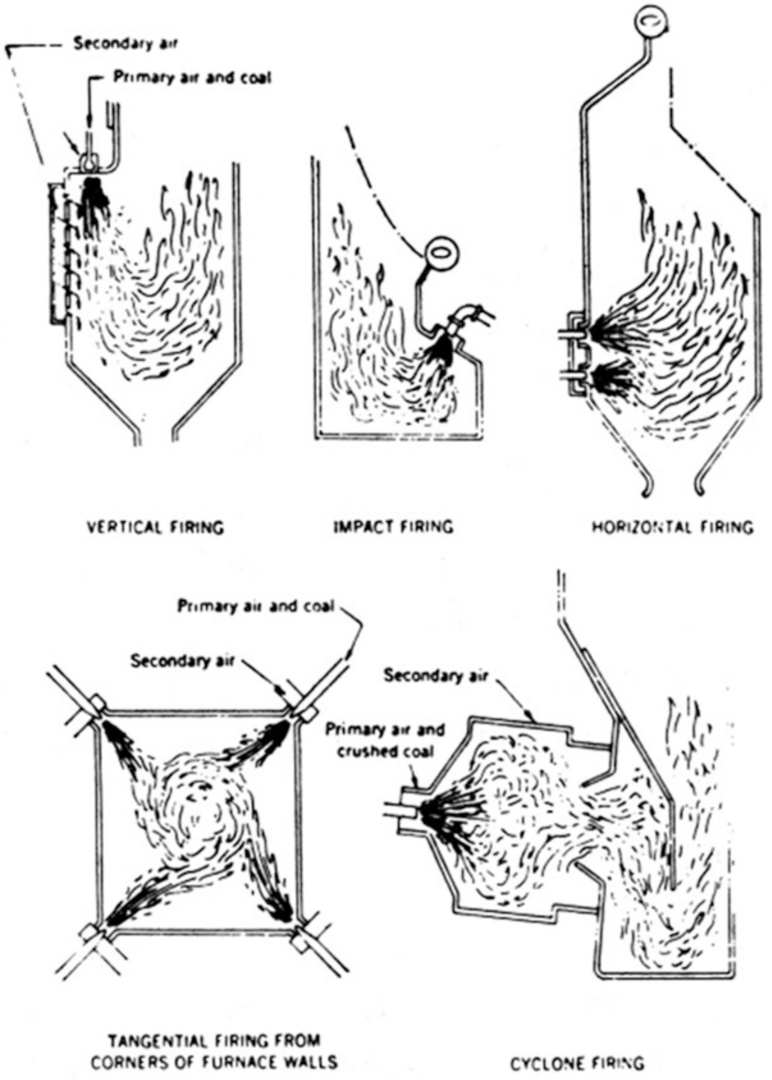


Fig. 19.11 Methods of firing pulverized and crushed coal (Courtesy of Bureau of Mines)

The other types of suspension firing use burners in which the primary air, coal, and secondary air are mixed just before or immediately after entering the furnace. With tangential firing, the burners are arranged in vertical banks at each corner of a square, or nearly square, furnace and directed tangentially toward an imaginary circle in the center of the furnace. This produces a vortex with its axis along the vertical centerline. The burners consist of an arrangement of slots, one above the other, admitting through alternate slots the primary air–fuel mixture and secondary air. The burners can be tilted upward or downward 30° from the horizontal plane, enabling the operator to control superheat and to permit selective utilization furnace heat-adsorbing surfaces. In tangential firing, the turbulence needed for mixing the fuel and air is generated in the furnace instead of in the burners.

The other major type of pulverized-coal-fired boiler is the wall-fired boiler. In wall-fired units, instead of being mounted in the corners, the burners are mounted on the walls of the furnace. They may be mounted on only one wall, or they may be mounted on opposing walls. The burners typically are mounted in a grid pattern.

Cyclone Firing. Cyclone firing is a form of two-stage wet-bottom firing in which the coal is not pulverized, as for suspension firing, but is crushed to 4-mesh size and admitted tangentially with primary air to a water-cooled cylindrical chamber called a cyclone furnace, as shown in Fig. 19.11. The finer particles burn in suspension, while the coarser ones are thrown by centrifugal force to the furnace wall. The wall, having a sticky coating of molten slag, retains the coal particles until they complete combustion. Secondary air, which is admitted tangentially along the top of the furnace, completes the combustion of the coarse particles. Slag drains continuously into the main furnace and through a tap hole into a water-filled slag tank, where it solidifies. The main advantage of cyclone firing is the ability to use larger coal particles. However, being a wet-bottom technique, it suffers from the disadvantages discussed above. In particular, the high temperature required to melt the slag results in high NO_x production.

Fluidized-Bed Combustion (*FBC*). In the past 30 years, FBC has advanced from the laboratory and pilot-plant scale to commercial status. A fluidized-bed combustor consists of a vessel containing a granular material, such as sand, coal ash, or limestone. This bed is fluidized by blowing air up through it. Fuel particles are injected into the bed and burn as they contact the oxygen in the air. Well-balanced air and fuel flow at the burner is needed to get the highest performance from existing and new low-NO_x firing systems. Developments to measure the air and fuel distribution continuously at the burners are the backbone for automated systems which constantly monitor and adjust the stoichiometry at the burners [20]. As ash particles agglomerate, they sink to the bottom of the bed and are removed. Fluidized-bed combustors are very flexible with respect to the fuel that can be burned, and almost any carbonaceous material that can be injected into the bed can be used. Steam tubes are generally immersed in the bed, where the heat-transfer coefficient is high.

Although this technology is not yet widely used in the electric power industry, the potential advantages are clear. (1) The bed temperature is low, about 840°C, which means less formation of NO_x and the retention of some of the sulfur in the ash of certain coals. (Adding dolomite or limestone to the bed greatly improves sulfur retention.) Of equal importance is less volatilization of sodium and potassium in the coal; consequently, there are fewer deposits on, and corrosion of, the furnace, superheater, and reheater tubes. (2) Heat-transfer rates from the fluidized-bed to the immersed heat-transfer surfaces are relatively high, as much as 100 BTU (hr $ft^{2\circ}F$). (3) Temperature control is easier, because rapid mixing in the bed results in a relatively uniform bed temperature. (4) FBC has the capability to burn many kinds of coal and is less sensitive to changes in coal heating value or ash characteristics than other combustion methods. (5) FBC is adaptable to several kinds of power generation cycles (e.g., cogeneration or combined cycle), can operate at atmospheric or elevated pressure, and can use bubbling or circulating beds with in-bed or external heat exchangers. This versatility allows engineers to develop FBC systems that are tailored to specific applications. This can result in significant cost savings when compared with other technologies.

The application of fluidized-bed boilers to electric utility stations should reduce operating costs and increase efficiency compared with conventional pulverized-coal-fired plants of the same capacity equipped with emission controls. The principle of the fluidized-bed combustor is illustrated schematically in Fig. 19.12.

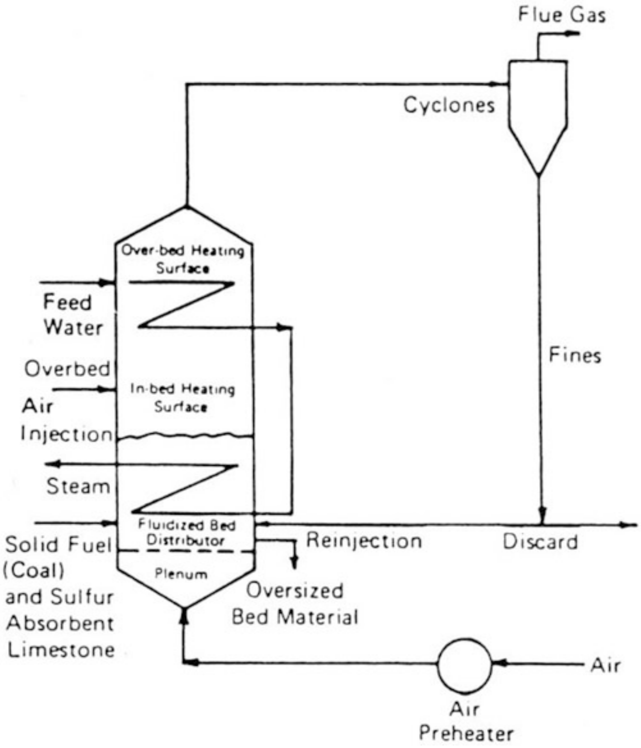


Fig. 19.12 Basic components of a fluidized-bed combustor (Shang YJ, Department of Energy, Morgantown, WV)

Interest is increasing in running units at elevated pressures (6–15 atm) in order to reduce equipment size, lower emissions, reduce sorbent requirements, and increase the thermal efficiency by using a gas turbine in addition to a steam turbine in a combined cycle. There is also a trend away from bubbling beds and toward circulating beds, which use finer solids and higher velocities and are able to use sorbents more efficiently.

Several large demonstrations of FBC technology for electric power generation have proven the viability of this technology. These include the 110-MW(e) atmospheric pressure circulating fluidized-bed boiler at the Colorado-Ute Electric Association's Nucla Station, which began operation in August 1988. In 2002, the JEA utility declared their two atmospheric fluidized-bed systems fully operational to generate 350 MW(e). A second project is the American Electric Power Tidd 70-MW(e) pressurized fluidized-bed combustor/combined-cycle demonstration, which began generating power in December 1990. These projects were partly funded by DOE's Clean Coal Technology program [21].

The Tidd plant utilized a nearly stationary bubbling bed technology (a first-generation technology) with a heat exchanger immersed in the bed to generate steam. The flue gas is utilized in a rugged gas turbine to generate additional electricity. A second-generation technology utilizes a circulating fluid bed that introduces a number of efficiency

Table 19.8 Approximate range of capacities of various types of industrial and utility boilers

Type	Capacity (lb steam/h)	Maximum design pressure (psig)
Fire tube (HRT)	1,000–15,000	250
Water tube		
Straight	15,000–150,000	2,000
Bent, three drum, low head	1,000–35,000	400
Two-drum vertical	1,000–350,000	1,000
Electric utility	Up to 9,000,000	3,500

enhancements. A major efficiency enhancing measure is the integration of a coal gasifier (carbonizer) to produce fuel gas. This fuel gas is combusted and adds to the combustor's flue gas energy fed to the gas turbine. While this adds complexity, it offers opportunity to improve efficiency.

Boiler Types

Industrial and utility boilers are broadly classified as fire tube or water tube. In fire-tube boilers, the hot combustion gases pass through tubes, and heat is transferred to water outside the tubes. Most steam locomotives had this type of boiler. The most common and least expensive boiler of this type is the horizontal return tubular (HRT) boiler. However, because of the design and construction of fire-tube boilers, there is a definite limitation to their size and the pressure that they can tolerate.

Water-tube boilers may be broadly classified as straight-tube and bent-tube types; the latter have several variations in design and are preferred for applications where higher capacities and steam pressures are required. In both types, heat is transferred by radiation or convection to the outside of the tubes, and water flows inside the tubes as a result of thermal circulation, or in the case of certain bent-tube boilers, as the result of forced circulation. A comparatively new version of the forced-circulation, bent-tube boiler for central-station power plants is the "once-through" type. The feed-water passes progressively through the heating, evaporation, and superheater sections; no drum is used for separating steam from unvaporized water as in other boilers, so the ratio of water circulated in the boiler to steam generated is unity. Table 19.8 presents approximate ranges of steam capacity and pressure for the principal types of boilers.

Combustion Going Forward

Unless some means of reducing CO_2 emissions is developed, generating electricity by coal combustion will be declining [22]. Existing US coal plants (over 600) have been listed, and some projections of their future have been made [23, 24]. Geisbrecht and Dipietro [25] have evaluated options for US coal-fired power plants and concluded that while retrofitting or repowering offers hope, the impact will depend on the extent to which efficiency and other cost factors can be collectively upgraded. The average efficiency of all coal power stations in the world currently is around 31%, and this leaves much potential to reduce coal consumption and CO_2 emissions [26]. Roberts and Goudarzi [27] concluded that an 8% improvement in the heat rate could be achieved if all coal generating units could reach the average thermal conversion performance of the 274 most efficient generating units.

Pollution Controls

Coal combustion is drawing increasing scrutiny because of concerns over coal's contribution to several environmental problems: acid rain (from oxides of sulfur and nitrogen), global warming (from CO_2), and particulates and trace elements (from ash) [28]. Despite increasingly stringent environmental regulations, coal is likely to remain in plants in the United States and many other countries [29]. The amount of coal burned is likely to increase steadily in developing countries like China and India. The combustion of coal is also important in a number of eastern European economies, such as Poland. Consequently, there is a strong impetus for developing highly efficient clean coal technologies. Achieving significant reduction of SO_2 and NO_x emissions from coal combustion is a major goal of the 1990 Clean Air Act Amendments. The achievement of these goals has been assisted by a US DOE effort known as the Clean Coal Technology program, which has provided government funding for a range of advanced pre- and post-combustion technologies to mitigate environmental problems.

Emissions of sulfur oxides, nitrogen oxides, and particulates from coal combustion are problems of increasing concern and regulation. Coal combustion contributes about 25% of the particulate matter, 25% of the sulfur oxides, and 5% of the nitrogen oxides emitted to the atmosphere. Much of the particulates are derived from the mineral content of the coal, but some particulates also result from sulfur and nitrogen oxides that react to form various sulfate and nitrate salts. A major concern about particulate matter is that the smallest particles are respirable and may pose a health hazard. Particulate matter is recovered in most power plants by the use of electrostatic precipitators, which have been developed to very high efficiencies (>99%). Other methods of particulate removal include baghouses and cyclone collectors.

An important technology for removal of SO_2 is flue gas desulfurization (FGD), carried out in units known as scrubbers. Most scrubbers contact the flue gas with a slurry of lime or limestone to capture the sulfur oxides and produce a sludge containing calcium sulfite and calcium sulfate. However, disposal of sludge is another environmental problem, and some scrubbers include oxidation to convert all the calcium sulfite to sulfate (gypsum), which can be used for wallboard manufacture. FBC units add a sulfur capture agent, such as limestone, to the bed, and integrated gasification combined cycle (IGCC) systems convert the sulfur in the coal to hydrogen sulfide, which can be removed easily from the fuel gas stream before it is burned.

Most of the NO_x emitted from coal combustion is in the form of nitric oxide (NO). Some of the NO_x originates from oxidation of nitrogen in the coal, and the rest comes from the reaction of nitrogen and oxygen in the combustion air. Consequently, partial control of NO_x can be achieved by changing the conditions of the combustion process. This usually involves the use of staged combustion, in which coal is first burned at fuel-rich conditions. Staged combustion decreases NO_x by lowering both oxygen concentration and temperature during combustion. So-called low-NO_x burners have been developed that can be retrofitted to most utility boilers to reduce NO_x emissions.

Other techniques that have been developed to reduce NO_x emissions involve post-combustion operations, such as reburning, selective noncatalytic reduction (SNCR), and selective catalytic reduction (SCR). In reburning, additional fuel (natural gas, oil, or pulverized coal) is added to the combustion gases downstream of the main combustion to result in fuel-rich, reducing conditions, which convert NO_x into nitrogen and water. Overfire air is then added to complete combustion, but because of the lower temperature, additional NO_x is not formed. In SNCR, ammonia or urea is injected into the furnace to react with NO_x and form N_2. For SNCR to be effective, the flue gas, after urea or ammonia injection, must spend a minimum residence time within a specific temperature range. In SCR, ammonia is injected into the flue gas, which then passes over a catalyst where reduction of NO_x to N_2 occurs. SCR is used extensively in Japan and Germany, and, because of decreasing costs, its use in the United States is steadily increasing. SCR is attractive because it can remove upward of 90% of the NO_x. Fluidized-bed combustors, in which combustion occurs at a lower temperature than in PC-fired furnaces, generally produce less NO_x.

Another pollutant which is coming under increasing scrutiny is mercury [30, 31]. Coal contains trace amounts of mercury (an average of about 0.1 ppm) which, because of its high volatility, is released into the atmosphere when the coal is burned. Approximately 50 t/year of mercury is released by power plants. Although coal cleaning can reduce the mercury content of coal, it will prove difficult to achieve complete removal. Therefore, post-combustion technology is the most probable approach to controlling mercury emissions. Possible technologies include adsorbents and scrubbers. Scrubbers are already used at many power plants to remove SO_2 from the flue gas, and many of these scrubbers have been found to remove a significant fraction of the mercury. One adsorbent that has been used effectively is activated carbon. The EPA is under mandate to develop regulations that may require removal of up to 90% of mercury emissions, and much progress has been made to accomplish this.

Pollution

Finally, the impact of CO_2 on the environment must be taken into account. This involves two issues for coal combustion: capture and sequestration [32]. The cost of carbon capture is not a simple issue, but it appears that carbon capture is less expensive for IGCC than for supercritical power coal (SCPC) plants (Fig. 19.13). The work to define the cost and evaluate the effectiveness of several approaches for sequestration is in its infancy; however, the data generated during decades of using CO_2 for enhanced oil recovery provides an excellent start in this direction [33].

Advances in Combustion Technology

Because of improvements in steam-generating equipment and turbines, remarkable advances have been made in the energy efficiency of electric utility plants. One method of expressing the efficiency of the conversion of coal to electricity is the heat rate, which is the amount of heat energy required to generate 1 kWh of electricity [34]. The heat rate of a modern, well-maintained plant in today is less than 9,000 BTU/kWh, compared with an average value of 17,500 BTU/kWh in 1937 and 10,000 BTU/kWh in 1968. However, the age of US plants causes US plants to have a rating of about 10,000 BTU/kWh. Some new designs project heat rates close to 7,500 BTU/kWh, corresponding to an overall thermal efficiency of 45% [35].

Further improvements in thermal efficiency may be achieved with new power generation systems now being developed. One promising technology is IGCC, in which both a gas turbine and a steam turbine are used to generate electricity. In this concept, the coal is first gasified to produce fuel gas (consisting mainly of CO and H_2, with small amounts of methane and other gases), which is cleaned of H_2S, NH_3, HCl, and particulates and then burned in a gas turbine. The hot exhaust gases from the gas turbine are used to generate steam that drives a steam turbine [36]. An important advantage is that the thermal efficiency of a combined-cycle plant exceeds 40%, whereas that of a conventional pulverized-coal-fired boiler equipped with FGD is

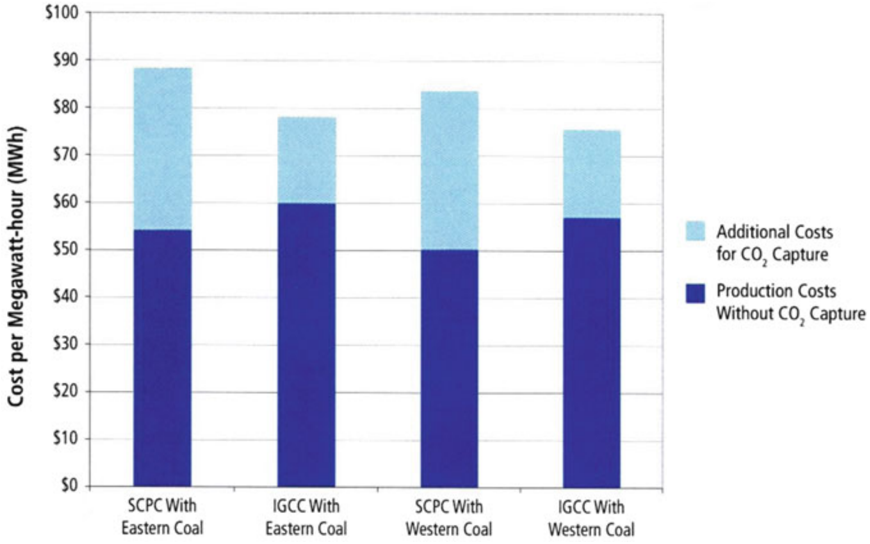


Fig. 19.13 Costs of electricity production from IGCC and SCPC with and without carbon capture technology [32]

about 35%. The development of more advanced gas turbines, capable of operating at higher gas inlet temperatures, will significantly increase the efficiency advantage of combined-cycle plants. IGCC is discussed in more detail later in this chapter in the section "Gasification."

Combined-cycle plants require clean fuel gas for the gas turbine. To maximize efficiency, gas cleanup should be carried out at high temperature. Removing trace amounts of impurities from the gasifier effluent without cooling the gas stream presents a technical challenge. An advantage of combined-cycle plants is that, because of fuel gas volume, pollutant removal equipment can be much smaller. Also, the sulfur and nitrogen are in forms that are more easily removed. Another advantage is the modular nature of combined-cycle plants, which offers flexibility in construction and the ability to add capacity in increments. The feasibility of combined-cycle plants is being demonstrated in two of DOE's Clean Coal Technology programs, the Tampa Electric Company IGCC Project and the Wabash River Coal Gasification Repowering Project. These plants are recognized as being two of the cleanest coal-fired power plants in the world.

The use of coal slurries, either coal–oil mixtures (COM) or coal–water mixtures (CWM), has been the subject of considerable research because of the potential to replace fuel oil in boilers. Recent work has shifted in favor of CWM over COM, as the cost savings for the latter were not enough to offset conversion costs. CWM is typically composed of 60–75% coal, 24–39% water, and 1% chemical additives. The advantages of CWM compared with dry pulverized fuel are the case and safety of handling, transporting (e.g., via pipeline), and injecting the fuel. It is possible that the manufacture of CWM can be integrated into coal beneficiation processes, such as froth flotation. Potential problems from replacing oil with CWM are (1) limited flame stability and turndown, (2) incomplete combustion, and (3) slagging and fouling caused by inorganic constituents. This technology is still in the development stage. Several demonstration projects have been completed or are under way, with the primary focus being on utilization of waste coal fines from coal-cleaning operations.

Chemical Looping

Chemical-looping combustion (CLC), first proposed as a method of improving power plant efficiency, is presently being studied for its potential for carbon capture. Still at an early stage of development [37], it may be used for either combustion [38] or gasification [39]. The concept is simple, but there are many impediments to be overcome before it becomes a commercial process. Two reactors are utilized with the oxidized and reduced solid, the oxygen carrier, being transported between the reactors as illustrated in Fig. 19.14. The fuel is mixed with the oxidized solid in the fuel reactor, where the fuel is oxidized to water and CO or CO_2, depending upon whether the process is combustion or gasification. In the air reactor, the reduced form of the solid that is generated in and transported from the fuel reactor is oxidized to produce CO_2 and water, together with the oxidized form of the catalyst. This process therefore allows CO_2 to be captured by condensing the water, and not from a dilute concentration in nitrogen, as is the case with normal combustion.

There are many problems to overcome. The circulation of the solid as well as the many oxidation–reduction cycles it undergoes can present severe attrition problems as well as significant losses of surface area. The final evaluation of this process is years into the future.

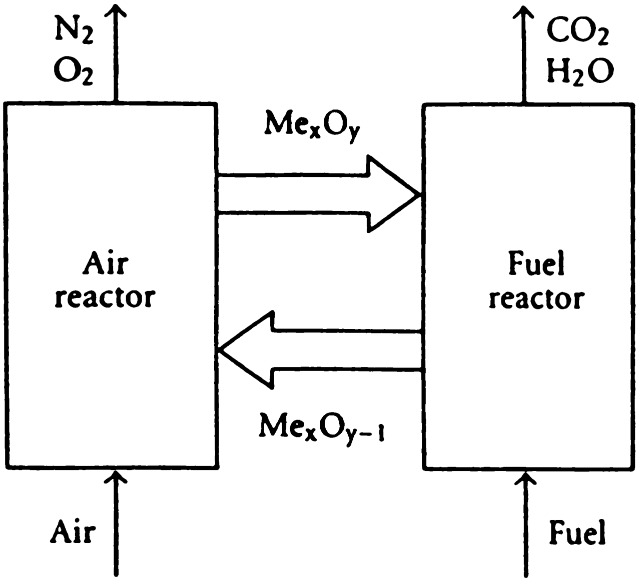


Fig. 19.14 Chemical-looping combustion (CLC)

Coke Production

The production of metals frequently requires the reduction of oxide-containing ores, the most important being production of iron from various iron oxide ores. Carbon in the form of coke is often used as the reducing agent in a blast furnace, which is basically a vertical tubular vessel to the top of which is fed alternate layers of iron ore, coke, and limestone. Coke is the porous residual solid remaining after a coal's volatile matter is released at high temperature. Coal cannot be fed directly at the top of a blast furnace, because it does not have the structural strength to support the column of iron ore and limestone in the furnace while maintaining sufficient porosity for the air blast to pass upward through the furnace.

Not all coals can produce coke that is suitable for use in a blast furnace. The property that distinguishes coking coals is their caking ability. Various tests such as maceral analysis, free-swelling index, and chemical analysis are performed to identify suitable coals. Frequently, to achieve the necessary coal properties to produce a suitable coke, several coals are blended to generate the feed to the coke plant. Commercial cokemaking processes can be divided into two categories: nonrecovery cokemaking and by-product cokemaking.

Combustion Future

A baseline performance and cost estimate for today's fossil energy plants was established [40]. This baseline can be used to benchmark the progress of energy progress. The study provided an accurate and independent assessment of the cost and performance of pulverized-coal (PC) combustion, IGCC and natural gas combined cycles with and without carbon capture and sequestration assuming the plants use the technology available in 2010. They also reviewed the current and future technologies for gasification-based power generation [41].

Nonrecovery Cokemaking

In nonrecovery coke plants, the volatile components released during cokemaking are not recovered, but rather are burned to produce heat for the coke oven and for auxiliary power production. One of the earliest nonrecovery units was the beehive oven, which for many years produced most of the coke used by the iron and steel industry. With these ovens, none of the by-products produced during coking were recovered.

A beehive oven consists of an old-fashioned beehive-shaped brick chamber. It has a charging hole in the top of the dome and a discharging hole in the oven's wall near the bottom. Coal is charged through the opening in the top and spreads over the oven floor. Retained heat from the previous charge is sufficient to start devolatilization. Evolved gases mix with air entering through the discharge door and burn, with the combustion products exiting through the charging port in the top. The heat generated by the burning gases is sufficient to complete the coking process. The finished coke is removed through the discharge door. Because of their low efficiency and pollution problems, beehive ovens are no longer used in the United States.

The general principle of operation of modern nonrecovery coke plants is the same as for the beehive oven, except that the units are designed to prevent the escape of volatile organic compounds (VOCs) and other pollutants to the atmosphere. Coking takes place in large rectangular chambers that are heated from the top by radiant heat transfer and from the bottom by conduction through the floor. Primary air for the combustion of evolved volatiles is introduced through several ports located above the charge level. Combustion gases exit the chamber through downcomers in the oven walls and enter the floor flue, thereby heating the floor of the oven. Combustion gases from all the chambers collect in a common tunnel and exit via a stack that creates a natural draft for the oven. To improve efficiency, a waste heat boiler can be added before the stack to recover waste heat and generate steam for power production.

At the completion of the coking process, the doors of the chamber are opened, and a ram pushes the hot coke (approximately 2,000°F) into a quench car, where it is typically cooled by spraying it with water. The coke is then screened and transported to the blast furnace.

By-Product Coke Production

The majority of coke produced in the United States comes from wet-charge, by-product coke-oven batteries. By-product cokemaking consists of the following operations: (1) selected coals are blended, pulverized, and oiled for bulk density control; (2) the blended coal is charged to a number of slot-type ovens, each oven sharing a common heating flue with the adjacent oven; (3) the coal is carbonized in a reducing atmosphere with the off-gas being collected and sent to the by-product plant for by-product recovery; and (4) the hot coke is discharged, quenched, and shipped to the blast furnace.

After the coke oven is charged with coal, heat is transferred from the heated brick walls to the coal charge. In the temperature range from 375°C to 475°C, the coal decomposes to form a plastic layer near the walls. From 475°C to 600°C, there is marked evolution of aromatic hydrocarbons and tar, followed by resolidification into semicoke. At 600–1,100°C, coke stabilization occurs, characterized by contraction of the coke mass, structural development of coke, and final hydrogen evolution. As time progresses, the plastic phase moves from the walls to the center of the oven. Some gas is trapped in the plastic mass, giving the coke its porous character. When coking is complete, the incandescent coke mass is pushed from the oven and wet or dry quenched prior to being sent to the blast furnace. Modern coke ovens trap the emissions released during coke pushing and quenching, so that air pollution is at a minimum.

By-Product Recovery. Coke-oven gas is released during the coking process. A typical composition for raw coke-oven gas is given in Table 19.9. In addition to the components shown in the table, raw coke-oven gas also contains various chemicals, such as tar vapors, light oil vapors (mainly benzene, toluene, and xylene), naphthalene, ammonia, hydrogen sulfide, and hydrogen cyanide. The by-product plant removes these chemicals so that the gas can be used as fuel. The volatiles emitted during the coking process are recovered as four major by-products: clean coke-oven gas, coal tar, ammonium sulfate, and light oil. The following description represents a typical coke-oven gas cleanup scheme, but considerable variations are possible from plant to plant.

Tar is the first by-product recovered. Hot coke-oven gas at a temperature of about 1,100°C is quenched to about 80°C by contact with an aqueous spray (flushing liquor) in the primary cooler. The cooled gas, along with excess flushing liquor, is collected in the coke-oven battery gas-collecting main. The condensed tars and water are transported to a tar decanter, where the heavier tars are separated out, and the water is recycled to a primary cooler. The gas passes through a compressor, called the exhauster, which sucks the gas from the ovens and provides sufficient pressure to allow the gas to flow through the rest of the by-product train. During cooling of the coke-oven gas, some of the tar forms an aerosol, which is carried along with the gas. Because these tar droplets would contaminate and foul downstream equipment, the tar mist is removed in an electrostatic precipitator. The recovered tar is pumped to the tar decanter.

Table 19.9 Composition of coke-oven gas

Component	Composition (vol. % (dry))
Hydrogen	54
Methane	25
Nitrogen	10
Carbon monoxide	6
Carbon dioxide	3
Higher hydrocarbons (ethane, propane, etc.)	2

Ammonia removal is the next step in the gas cleanup process. A typical approach is to contact the coke-oven gas with a solution of sulfuric acid to form ammonium sulfate, which is then recovered by crystallization, dried, and sold as fertilizer. More modern processes for ammonia removal include the water wash process, in which the coke-oven gas is scrubbed with water that dissolves the ammonia along with the hot blast near the bottom of the blast furnace. The coal reacts with oxygen to form H_2 and CO, which act as reducing agents as they rise through the furnace. As long as sufficient coke is used to preserve bed porosity, a portion of the coke can be displaced by injected coal on a roughly pound per pound basis. Reducing coke requirement not only reduces costs, because coal is cheaper than coke, but also reduces the environmental problems associated with coke production. Furthermore, furnace capacity is increased, because the smaller volume of coke charged means a larger volume of iron ore can be charged.

A wide variety of noncoking coals have been successfully tested for direct injection, ranging from lignite to anthracites. The choice depends on price and availability rather than on attaining the highest injection rates. Among coals of similar cost and availability, important factors are ash, volatile matter, moisture content, and grindability. Also, sulfur and alkali contents should be low. Because ash lowers temperature and affects grindability, ash should be less than 10%. When injecting pulverized coal, volatile matter should be as high as possible, since low-volatile coals tend not to burn completely and contribute particulate matter to the off-gas. Moisture should be minimized, as water adsorbs heat in the blast furnace, causes difficulties during grinding, and can cause plugging problems in storage vessels and injection equipment. In addition to coal, natural gas, oil, and tar are also injected into blast furnaces.

Another approach receiving increased interest is direct reduction of iron ore. In this process, coal is added directly to the smelting vessel to become the source of the reducing gases and thermal energy. Thus, easily crushed and handled, inexpensive noncoking coals of high calorific value can be used. Particle size requirements are variable, but less than 1 mm is most often used, and, in some cases, there is an effort to restrict moisture content to below 6%.

Mild Gasification

Mild gasification is actually a pyrolysis-based process, closely related to coking. It is designed to produce the most profitable slate of alternative fuels and chemicals by decomposing coal at relatively mild conditions of temperature and pressure [42–45]. Mild gasification differs from coking in that in mild gasification, the goal is the production of liquid and gaseous fuels with the accompanying char being the by-product, whereas in coking, the coke is the main product, and the liquids and gases are the by-products. A major problem with this technology is the low quality of the liquid products, which are hydrogen deficient and require expensive upgrading to convert them into gasoline, diesel fuel, and other products. Pyrolysis-based processing for the production of liquids is discussed in more detail later in this chapter under "Pyrolysis-Based Processes" in the section "Coal Liquefaction."

Gasification

Coal was first gasified in England by William Murdock in 1792, and the world's first coal gas company was chartered in England in 1812. Coal gas was first produced in America in 1816 in Baltimore, and by 1850, over 55 commercial coal gasification plants in the United States were generating gas for lighting and heating. During the late 1800s and the early 1900s, a large number of coal gasifiers operated commercially in the United States and Europe to produce industrial and residential fuel gas. In the early 1930s, there were over 11,000 coal gasifiers operating in the United States, consuming about 13 million tons of coal per year and producing 500 million BTU/day of coal gas.

Most of the early gasifiers were moving-bed units, charged with sized coal and blown with steam and air to generate "producer gas" (150 BTU/scf). Operation of the moving-bed gasifiers in a cyclic mode (blowing first with air to heat the coal, followed by contact with steam to produce "water gas") increased the heating value of the product gas to 300 BTU/scf. The heating value of the fuel gas was further increased to about 500 BTU/scf by co-feeding oil with steam in the moving-bed process to produce "carbureted water gas," which contained hydrocarbons in addition to H_2 and CO. Coke-oven gas was also used as "town gas." An early gasification process, still in use today, was that developed by Lurgi. This gasifier operates at elevated pressure and uses oxygen, instead of air, thus increasing the heating value of the product gas.

By the mid-1950s, the availability of inexpensive natural gas and petroleum had led to the abandonment of most coal gasification units. However, even as oil and natural gas use increased, coal gasification plants continued to be built on a limited scale in certain areas of the world, particularly in countries with a limited supply of petroleum.

Extensive process development was carried out in the United States in the late 1940s to mid-1950s, prompted by a concern that natural gas reserves were limited. This work included (1) development of a 100-t/day Texaco entrained-flow gasifier at Morgantown, West Virginia, by Texaco and Olin Matheson; (2) demonstration of a 360-t/day Babcock and Wilcox entrained-flow gasifier at Belle, West Virginia; and (3) development of a 24-t/day Koppers-Totzek entrained-flow gasifier at Louisiana, Missouri. Further work has resulted in the development of several additional gasifiers, including the Shell, Kellogg-Rust Westinghouse (KRW), E-Gas™, British Gas/Lurgi (BGL), Kilngas, and Dow gasifiers. Recent interest in coal gasification has been driven by the potential of IGCC facilities to increase the efficiency of power production and reduce emissions. Development of improved coal gasification processes increases the potential for implementation of coal gasification technology for a variety of applications [46, 47].

In modern gasification systems [47, 48], a bed of coal or other carbonaceous material is blown with a mixture of steam and air or oxygen under pressure to produce a low to medium heating value gas whose main constituents are CO and H_2 with smaller amounts of CO_2, CH_4, and other components. (A mixture of H_2 and CO is called synthesis gas or syngas. Chapter 10 discusses some of the chemicals that can be made from syngas.) Most of the mineral matter (ash) in the feed leaves the bottom of the gasifier either as a dry solid or as a liquid, which solidifies to form an inert glassy slag, suitable for sale or disposal. Only a small fraction of the ash is entrained in the syngas and requires removal downstream. During gasification, sulfur in the feed coal is reduced to H_2S, and nitrogen is reduced to NH_3. These materials, along with HCl and entrained particulates, are removed in the downstream gas cleanup system. A typical gasifier raw gas composition is shown in Table 19.10.

Chemistry of Coal Gasification

The initial step in coal gasification involves grinding and/or pretreatment of the coal to put it into a form suitable for injection into the gasifier. In the gasifier, the coal is heated in the presence of a reactive gas whose composition depends on

Table 19.10 Typical coal gasifier raw gas composition

Component	Concentration (vol. %)
Hydrogen	25–30
Carbon monoxide	30–60
Carbon dioxide	5–15
Water	2–30
Methane	0–5
Hydrogen sulfide	0.2–1
Carbonyl sulfide	0–0.1
Nitrogen	0.5–4
Argon	0.2–1
Ammonia/hydrogen cyanide	0–0.3

Table 19.11 Reactions occurring during coal gasification

	ΔH_r (kcal/gmol)	Name	Number
Pyrolysis reactions			
$CH_x = C + CH_4$		Pyrolysis	(17.1)
$CH_x + H_2 = C + CH_4$		Hydropyrolysis	(17.2)
Gas–solid reactions			
$C + 0.5O_2 = CO$	−26.45	Combustion	(17.3)
$C + O_2 = CO_2$	−94.1	Combustion	(17.4)
$C + H_2O = CO + H_2$	41.9	Steam/carbon	(17.5)
$C + CO_2 = 2CO$	41.25	Reserve boudouard	(17.6)
$C + 2H_2 = CH_4$	−17.9	Hydrogasification	(17.7)
Gas–gas reactions			
$CO + H_2O = H_2 + CO_2$	−0.7	Water–gas shift	(17.8)
$CO + 3H_2 = CH_4 + H_2O$	−59.8	Methanation	(17.9)

the product desired. Walker [49] has presented an extensive review of gasification reaction fundamentals. Breault [50] and Shinner [51] have reviewed both the older and newer major processes for gasification.

Reactions occurring during gasification of coal can be divided into three groups: (1) pyrolysis reactions (thermal decomposition of the coal), (2) gasification reactions (gas–solid reactions), and (3) gas–gas reactions. The major reactions taking place are listed in Table 19.11.

The first reaction that occurs is pyrolysis or devolatilization (Reaction (17.1)). Reaction (17.2) represents the conversion of gaseous pyrolysis products into predominantly methane and coke, the extent of this reaction depending upon gasifier design. Gasification is related to pyrolysis/coking in that, when coal is gasified by reacting it at a high temperature with steam and an oxidant (air or oxygen), pyrolysis is the first reaction as the coal is heated to reaction temperature. Depending on the type of gasifier, condensable hydrocarbons may be collected as a by-product or may be completely destroyed.

Combustion reactions, Reactions (17.3) and (17.4), are the fastest gasification reactions and are highly exothermic. The oxygen can come from air, enriched air, or pure oxygen. In practice, the oxidation reactions occur rapidly with essentially complete oxygen consumption, so that most of the reactor volume is utilized for the other reactions listed in Table 19.11. The primary combustion products are CO and CO_2. The equilibrium constants for both Reactions (17.3) and (17.4) are greater than 1 up to 3,500°C, and, as a result, both reactions are essentially irreversible at typical gasifier temperatures. The CO/CO_2 ratio should increase with lower pressure and higher temperature, because pressure does not influence the equilibrium of Reaction (17.4), and Reaction (17.3) is not favored under pressure.

Reactions (17.5)–(17.7) illustrate the gasification of char by reaction with various gases. The carbon–steam Reaction (17.5) is an endothermic reversible reaction. Steam undergoes a side reaction, Reaction (17.8), called the water–gas shift reaction. This reaction, which is very rapid, is catalyzed by various impurities and surfaces. The carbon–CO_2 reaction, Reaction (17.6), is favored at high temperatures and low pressures, whereas the carbon–H_2 reaction, Reaction (17.7), is favored at low temperatures and high pressure. Since only three of Reactions (17.5)–(17.9) are independent, if the equilibrium constants for Reactions (17.6)–(17.8) are known, the equilibrium constants for the other reactions can be calculated.

The reactions in Table 19.11 are useful for examining the effect of changes in gasifier operating conditions on product gas composition. As temperature is increased, endothermic reactions are favored over exothermic reactions, which means that methane production will decrease and CO formation will increase. An increase in pressure favors reactions in which there are fewer moles of products than moles of reactants. Thus, higher pressure favors methane and CO_2 formation. Coal char gasification is generally equilibrium controlled. For design purposes, it is common to consider char as thermodynamically equivalent to graphite, although this is not strictly true.

Even when gasifier operating conditions are known, it is not possible to calculate the composition of the product gas exactly. First, the pyrolysis reactions and those of the pyrolysis products are highly dependent on coal properties and gasifier operating mode and vary significantly from case to case. Second, the thermodynamic equations treat coal char as if it were pure carbon, which is not the case. Finally, the system may not reach equilibrium. Thus, although useful predictions of gas composition can be made, exact calculations are not possible. Nevertheless, knowledge of stoichiometry, equilibrium constants, and reaction rates provides a reasonable basis for the modeling and design of gasification systems.

Pore structure can markedly affect char reactivity. Coals in general are highly porous with a polymodal pore size distribution. Pores normally are classified into macropores (>500 Å in diameter), mesopores (20–500 Å in diameter), and micropores (<20 Å in diameter). Upon pyrolysis, the pores in the coal open up, but the char still contains microporosity. Coal chars in general, and lignitic chars in particular, retain coal's polymodal pore distribution. Surface area of coal chars are in the range 100–800 m^2/g [52]. Most of this surface area and, therefore, most of the active surface reside inside the pores, so accessibility of reactive gases to active sites is very important.

The fundamental reactions occurring during gasification can be described by the Langmuir–Hinshelwood and Rideal–Eley mechanisms. The Langmuir–Hinshelwood mechanism involves three steps: (1) adsorption of the gas onto the solid surface, (2) surface migration/reaction, and (3) desorption of the products from the solid surface. In the Rideal–Eley mechanism, the basic steps are (1) reaction between gas molecules and surface atoms by direct collision and (2) desorption of products.

Because coal chars are highly microporous, most of the gasification reactions take place inside the char particles. Therefore, diffusion of gas into, and products out of, porous particles is required. The overall diffusion process can be described by the following steps: (1) diffusion of the reactant from the bulk gas to the solid surface (film diffusion), (2) diffusion of the reactant from the particle's surface to its interior (internal diffusion), (3) diffusion of the product from the interior to the particle's surface (internal diffusion), and (4) diffusion of the product from the surface to the bulk gas (film diffusion).

If the pore diameter and process conditions are well defined, the rates of internal and film diffusion can be calculated. The temperature dependency of the rate can be presented in the form of an Arrhenius plot, that is, log rate vs. reciprocal of the absolute temperature. Gasification rates can be divided into three zones, depending on whether reaction rate is controlling, diffusion is controlling, or both reaction rate and diffusion are important (Fig. 19.15). To determine the overall gasification rate, it is necessary to model the actual gasifier.

Because of the reducing nature of the product gas, heteroatoms (sulfur, nitrogen, and chlorine) appear in reduced form; that is, sulfur appears as hydrogen sulfide, nitrogen as ammonia, and chlorine as hydrogen chloride. In most cases, these materials are scrubbed from the product gas before it is burned. Ammonia and HCl are very water soluble and are easily removed by a water wash. A number of processes have been developed for H_2S removal; many of these processes are based on absorption in solutions of amines, such as monoethanolamine (MEA).

Types of Coal Gasifiers

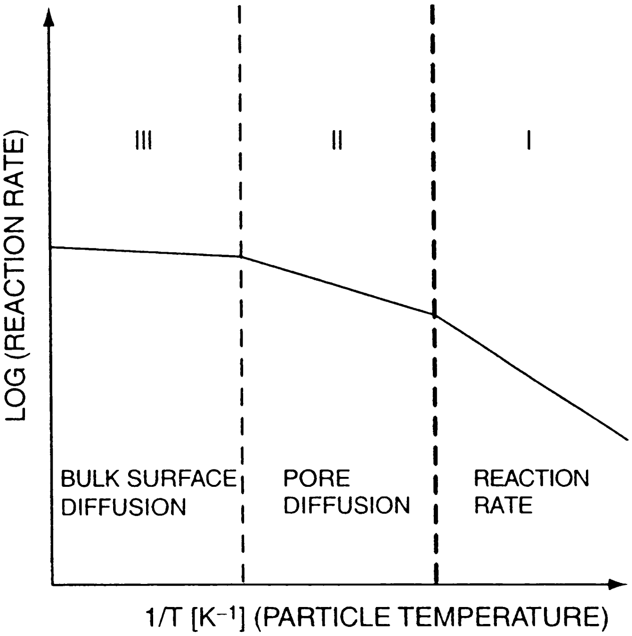


Fig. 19.15 Effective reaction rates in temperature zones

Gasification processes can be separated into three major types: (1) moving-bed (countercurrent flow) reactors, (2) fluidized-bed (back-mixed) reactors, and (3) entrained-flow (not back-mixed) reactors. Figure 19.16 shows the types of gasification reactors together with temperature profiles and locations of feed and product streams. Table 19.12 summarizes the important characteristics of each type of gasifier, and Table 19.13 presents the performance characteristics of selected gasifiers.

Moving-Bed Gasifiers. Moving-bed gasifiers consist of a downward-moving bed of coal contacted by a countercurrent flow of gases moving upward through the bed. In moving down the bed, coal sequentially undergoes dying, devolatilization, gasification, combustion, and ash cooling. Moving-bed gasifiers can be operated at atmospheric pressure or can be pressurized with either air or oxygen as the oxidant, with either dry ash or oxygen as the oxidant, with either dry ash or molten ash removal, and with or without stirrers to prevent agglomeration.

In the drying zone at the top of the bed, the hot upward-flowing gases remove moisture from the coal. As the coal moves down the bed, its temperature increases, and at about 315–480°C, pyrolysis occurs, liberating gases, oils, and tars. As the devolatilized coal (char) further descends, it is gasified by reaction with steam, carbon dioxide, and hydrogen to produce a mixture of carbon monoxide, hydrogen, and methane, containing unreacted steam and carbon dioxide. Below this, char is combusted by reaction with oxygen. The final composition of the gas produced and the rates of reaction are strongly influenced by gasifier temperature and pressure.

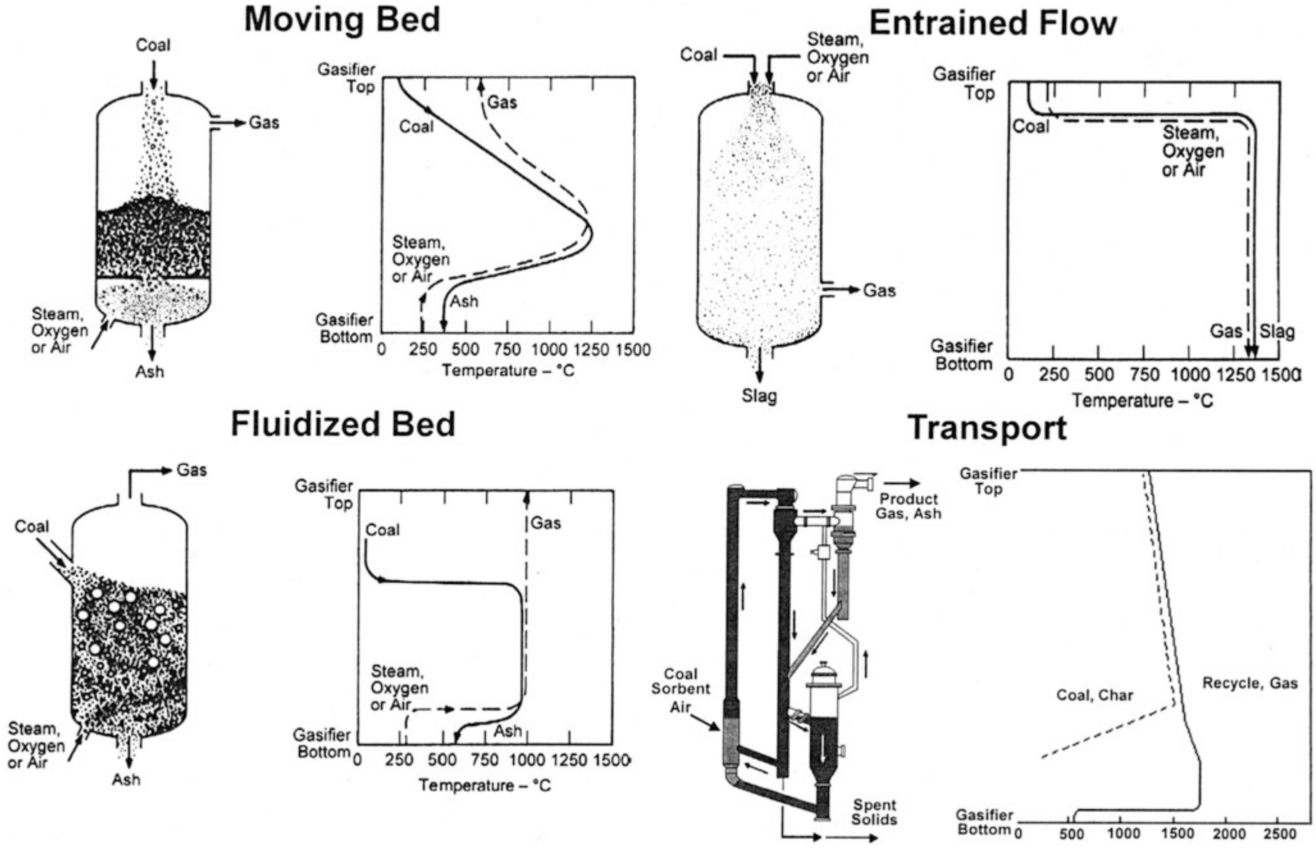


Fig. 19.16 Gasifier configurations

Table 19.12 Features of fixed-bed, fluidized-bed, and entrained-bed gasifiers

Moving bed	Fluidized bed	Entrained bed
Parameters		
Bed of coal descending under the influence of gravity	One or more fluidized-bed zones	Upflow or downflow suspension gasification
Mechanical grates/distributors	Uniform temperature and composition throughout each fluidized zone	High-temperature–high-rate process
	Moderate process response	Fast process response
Discrete zones		
Preheating-drying devolatilization Gasification Combustion		
Advantages		
High carbon conversion efficiency	High degree of process uniformity	Handles all types of coal without pretreatment
Low-ash carryover	Excellent solids/gas contact	Low steam consumption
Lower-temperature operation	Lower residence time than fixed-bed gasifier	Excellent solids–gas contact
Lowest air/oxygen requirement	Higher coal throughput per unit volume of reactor	No tar formation No phenol formation Ability to slag ash High capacity per unit volume of reactor Produces inert slagged ash
Disadvantages		
Sized coal required	Sized coal required	Requires finely crushed coal
Coal fines must be briquetted	Dry coal required for feeding	Small surge capacity requiring close control
Low off-gas temperature	Requires complicated gas distributor	

Table 19.13 Performance characteristics of selected gasifiers

	Illinois No. 6 coal		Pittsburgh No. 8 coal				West Kent coal
Parameters	Lurgi dry ash	Lurgi slagger	Texaco	Shell	KRW PDU	KRW design estimate	U-Gas
T (°F)	1,600	2,700	2,700	3,000	1,793	1,850	1,831
P (psia)	315	300	600	365	230	600	15
H_2O–coal (mole–mole)	2.42	0.293	0.027	0	13	0.29	1.34
O_2–coal (mole–mole)	0.286	0.259	0.463	0.435	0.44	0.30	0.45
C conversion (%)	99.3	99.3	99	99.3	76.0	89.6	93.3
Gas composition (dry)							
CO	15.36	58.05	51.69	61.46	43.38	51.5	26.1
CO_2	31.1	1.94	10.6	1.65	35.30	9.3	23.4
H_2	42.9	30.41	35.1	30.6	18.04	25.9	37.4
CH_4	3.78	7.76	0.09	0.04	1.33	10.4	2.8

Source: Shinner [51]

The ratio of steam to oxygen (air) in the gasifier controls the peak temperature in the combustion zone. If a nonslagging operation is desired, sufficient steam is added to offset exothermic oxidation reactions with endothermic steam–carbon reactions to stay safely below the ash-fusion temperature. Slagging gasifiers operate at a higher temperature, and ash is removed in a molten state and then quenched in a water bath.

Moving-bed gasifiers require sized coal for proper operation; typically, coal between 0.6 and 5 cm in diameter is used. The countercurrent flow of gases and solids leads to high thermal efficiency, with exit gas temperature usually in the range of 260–540°C. The raw product gas contains fines and tar/oil produced during gasification. These materials are removed prior to downstream processing. Figure 19.17 illustrates the mechanical configuration for a Lurgi dry-bottom gasifier.

Fluidized-Bed Gasifiers. In a fluidized-bed gasifier, reactant gases are introduced through a distributor at the bottom of the bed at a velocity sufficient to suspend the feed particles. The result is a bed of highly mixed solids in intimate contact with the gas phase. The agitation leads to a uniform temperature throughout the bed. The high degree of mixing results in reaction rates that are typically higher than those experienced in moving-bed gasifiers, although fluidized-bed gasifiers generally do not achieve as high a thermal efficiency as moving-bed gasifiers.

The exit gas temperature for a fluidized-bed gasifier is higher than that for a moving-bed gasifier. The higher temperature results in further reaction of the pyrolysis products, and, therefore, the product gas contains a much lower concentration of tar/oil. However, particle attrition generally leads to higher levels of fines/dust carryover. Unconverted char and ash are removed as dry solids. Pretreatment is required if strongly caking coals are used.

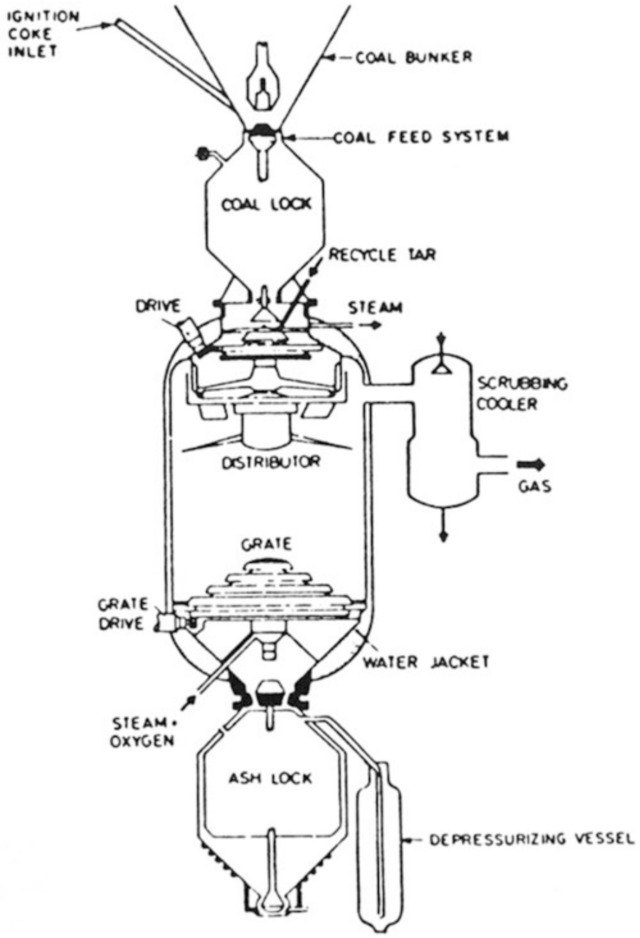


Fig. 19.17 Lurgi dry-bottom gasifier ("Comparative study of coal gasification processes—Koppers-Totzek, Lurgi and Winkler," Koppers Company, Pittsburgh, PA)

Fluidized-bed gasifiers can be operated at atmospheric or higher pressure. Fluidizing gases can be mixtures of steam with either air or oxygen and, in some cases, mixtures of hydrogen with other gases. The KRW gasifier is illustrated in Fig. 19.18.

Entrained-Flow Gasifiers. In an entrained-flow gasifier, a mixture of finely ground coal entrained in a reactant gas

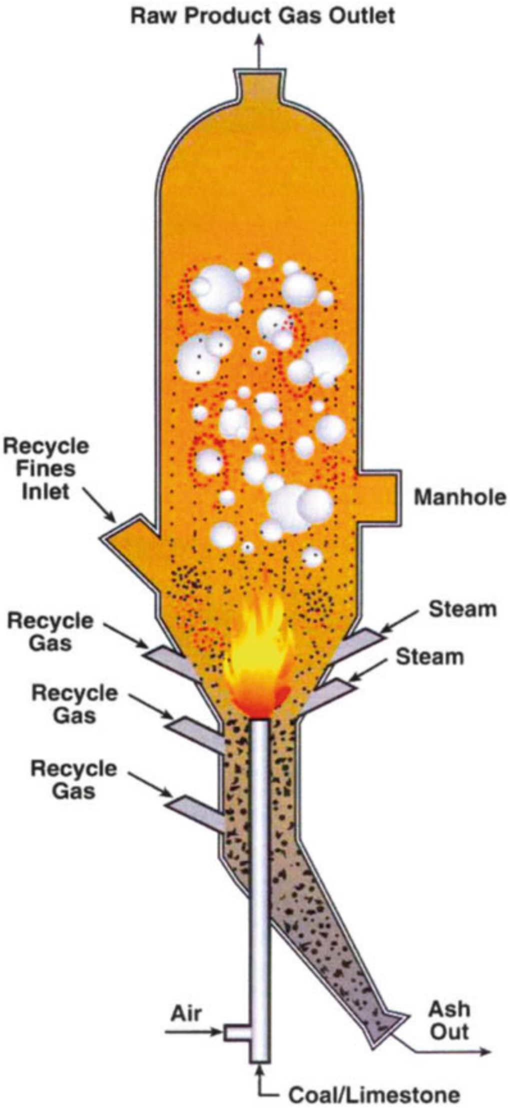


Fig. 19.18 KRW gasifier

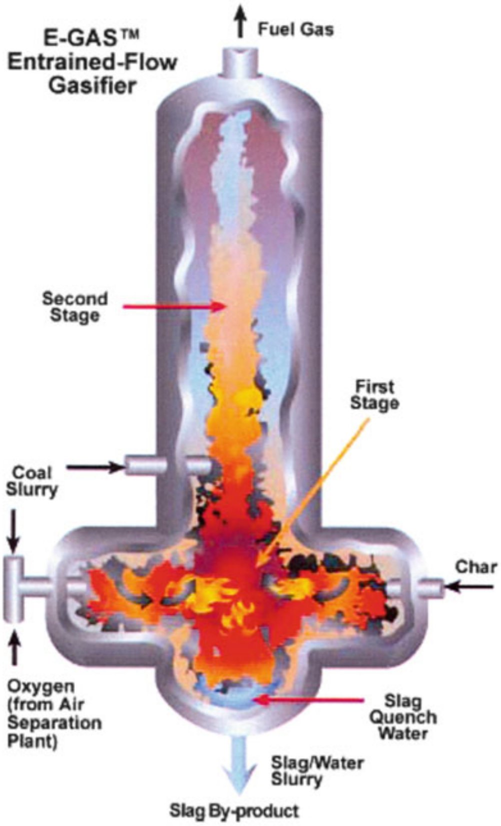


Fig. 19.19 Schematic of E-Gas™ entrained-flow gasifier

flows co-currently through the reactor, with little or no backmixing. This type of gasifier may be either single-stage or two-stage.

In general, high temperatures (1,200–1,760°C) are used to achieve complete gasification of the coal in a mixture with steam and oxygen or air. Because of the high temperatures employed, gasification rates are considerably higher than those achieved with either fixed-bed or fluidized-bed gasifiers. Entrained-flow gasifiers can handle all coals, including strongly caking coals, without pretreatment. The high temperature of operation produces a gas devoid of both methane and tars/oils. In two-stage gasifiers, the incoming coal is first entrained with reactant gases to produce gas; the resultant char is gasified further in a second stage, which may or may not be entrained. As is the case with fluidized-bed gasifiers, a staged operation achieves better overall thermal efficiency without sacrificing higher throughput, because the more reactive incoming coal can be gasified at a lower temperature than the less reactive char. Entrained-flow gasifiers can be operated at atmospheric pressure or pressurized, and ash may be removed either dry or molten. Figure 19.19 illustrates the two-stage entrained-flow E-Gas™ gasifier.

Gasification for Power Generation

High reliability, acceptable capital and operating costs, and minimal environmental impact are requirements for gasification systems proposed for utility applications. Operating costs can be minimized by using a gasifier that is capable of operating on run-of-mine coal, including caking coals and fines. The potential also exists to feed other materials such as petroleum coke, lignite, sewage sludge, refuse-derived fuel (RDF), and biomass to the gasifier. Use of such opportunity fuels can lower electricity costs.

Numerous studies confirm that gasifiers coupled with gas turbine–steam turbine combined cycles represent one of the most promising technologies for future coal-based power

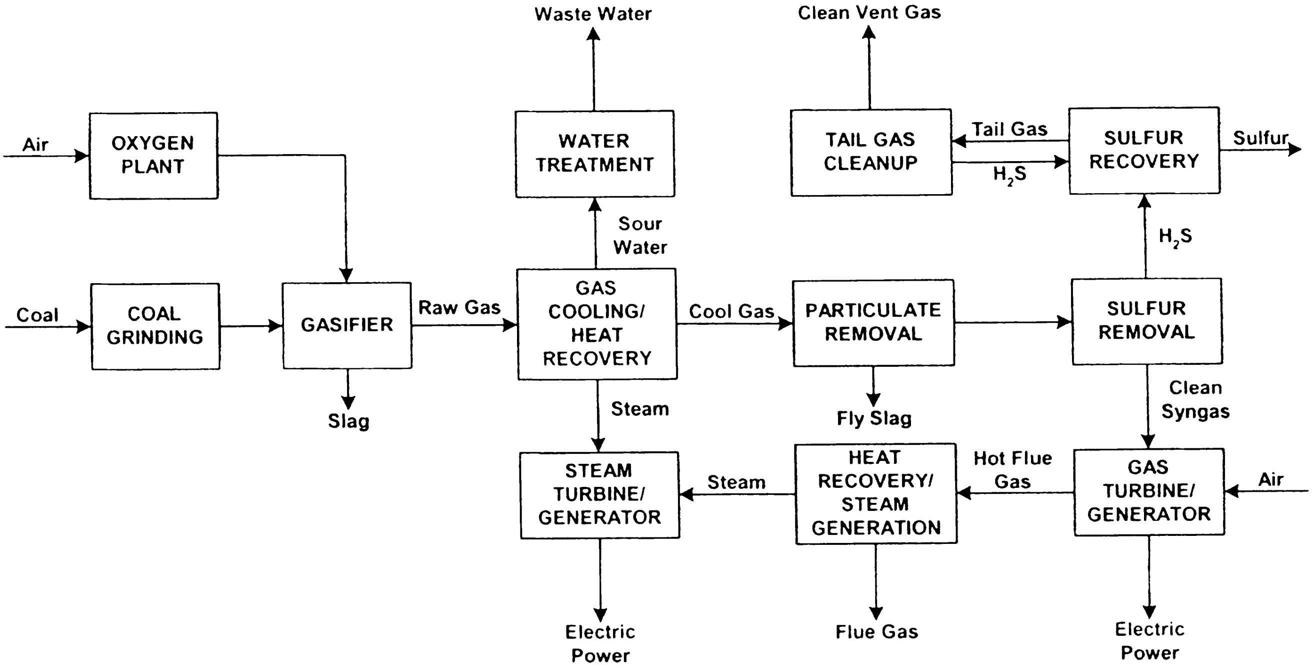


Fig. 19.20 Block flow diagram of a typical integrated gasification combined-cycle power plant

generation systems. IGCC technologies offer the potential for high efficiencies with low pollutant emissions. High efficiencies are achieved in IGCC operation by combining efficient combustion turbines with steam turbine bottoming cycles. The gas turbine converts high-temperature (1,820°C) heat efficiently, while the steam turbine utilizes lower-temperature heat (565°C) in the form of steam generated by recovering heat from the gas turbine exhaust. A typical IGCC facility is illustrated in Fig. 19.20.

A variation on IGCC operation is to combine power generation with chemicals or fuels production, the so-called coproduction option. In this operating mode, part of the syngas produced by the gasifier is used for power production, and part is sent to a Fischer–Tropsch [53], methanol, or similar facility for fuels/chemicals production. The coproduction gasifier facility provides the flexibility to maximize power generation during peak demand while maximizing fuels/chemicals production during off-peak periods. This makes gasification more attractive in the power generation market and allows flexibility in the geographical location of plants based on local grid peak shaving needs. Although not yet commercial, coproduction is being investigated by several energy companies.

A major goal of power production is minimal environmental impact. Because the product gas from IGCC systems is purified before combustion [54], burning this clean fuel results in low pollutant emission levels. Another potential environmental problem is eliminated by designing and operating the gasifier so that tars are not formed. Ash leaving the system is usually in the form of molten slag, which is water quenched to form benign vitreous material suitable for use or disposal. On balance, coal gasification systems are environmentally superior to other alternative coal utilization technologies and can meet rigorous environmental standards for SO_2, NO_x, and particulates. Furthermore, because of their increased efficiency, IGCC plants emit less CO, the major greenhouse gas, per unit of electricity generated.

Description of Selected Gasification Processes

Many different gasifiers have been developed, at least through the demonstration stage, by a variety of organizations, both public and private. However, not all of these gasifiers have achieved commercial success, and improved processes now supercede some technologies that were widely used in the past. In 2007, the GE, Shell, and Lurgi (dry ash) processes accounted for over 90% of the installed coal gasification capacity. It is beyond the scope of this chapter to describe all existing gasifiers. The following processes cover a wide variety of gasifier types and designs.

GE Energy Gasifier (*GEEG*). The GEEG (formerly Texaco Gasification) offers the following design features: a down-fed, pressurized, entrained-flow, slagging reactor; air or oxygen blown; high operating temperature; flexible feedstocks and products; and system coupling for cogeneration. The GEEG is designed to operate at pressures between

20 and 30 atm and temperatures between 1,200 and 1,500°C. Figure 19.21 shows a direct quench gas cooling configuration for the GEEG. The coal is wet ground and mixed with water to make a slurry containing 60–70 wt.% solids. The slurry is mixed with oxygen or air in the gasifier burner. By properly adjusting the oxygen/slurry ratio, gasifier temperature is maintained above the ash-fusion temperature.

The raw product gas is cooled either by direct contact with quench water, as in Fig. 19.21, or by passage through a radiative cooler followed by a convective cooler (Fig. 19.22). The cooled gases go to a scrubber to remove entrained char and ash. The product gas leaving the scrubber contains H_2, CO, CO_2, H_2O, and traces of Ar, N_2, CH_2, H_2S, and COS. Spencer and others have reported on the performance of the GEEG [55–57].

Various products have been prepared in commercial plants employing GEEGs, including H_2, NH_3, fuel gas for electricity generation, and syngas for production of methanol and other oxo products. The direct water quench cooling method is preferred when the desired final product is NH_3 or H_2, as the required shift reactor is easily integrated into the quench train.

Any carbonaceous material that can be processed into a concentrated, pumpable slurry can be gasified in a GEEG. For example, petroleum coke is converted to syngas in a 30-t/day plant in Ube City, Japan, which has been operating since 1982. In 1991, a mixture of coal and sewage sludge was successfully gasified. The ability of the GEEG to handle a wide range of feedstocks without the need for additional equipment or changes in operating procedures ensures flexibility in selecting the least expensive feed alternative while

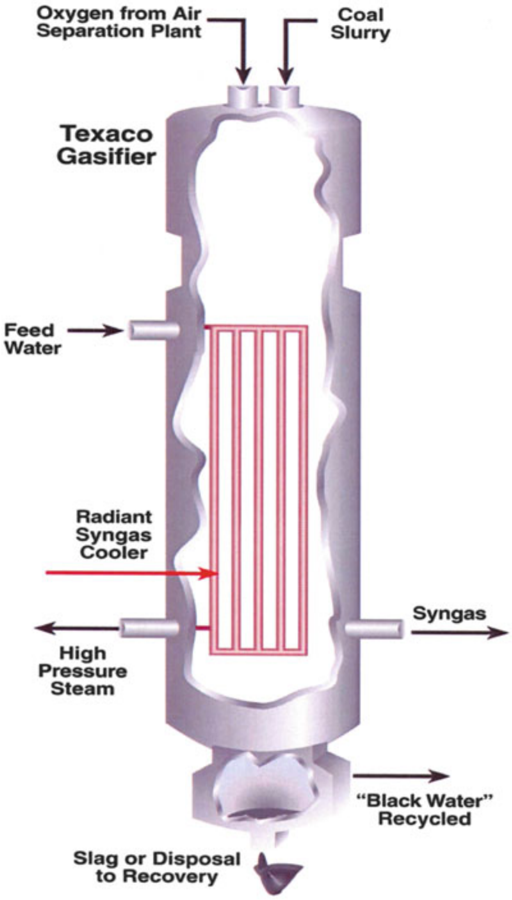


Fig. 19.21 Flow diagram of GE (Texaco) gasifier (Adapted from "Clean Coal Technology," DOE/NETL-2004/1207, August, 2004)

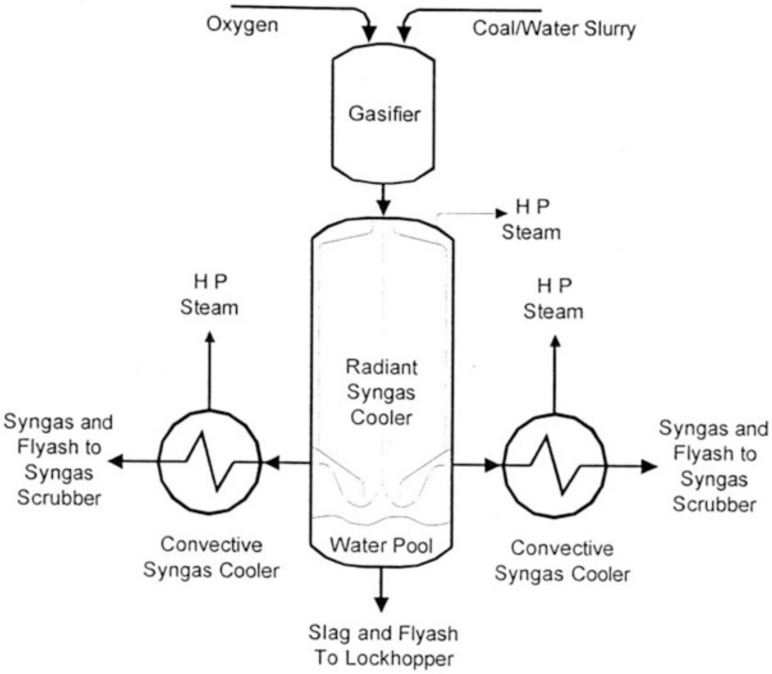


Fig. 19.22 Flow diagram of syngas cooling system (Adapted from "Clean Coal Technology," DOE/NETL-2004/1207, August, 2004)

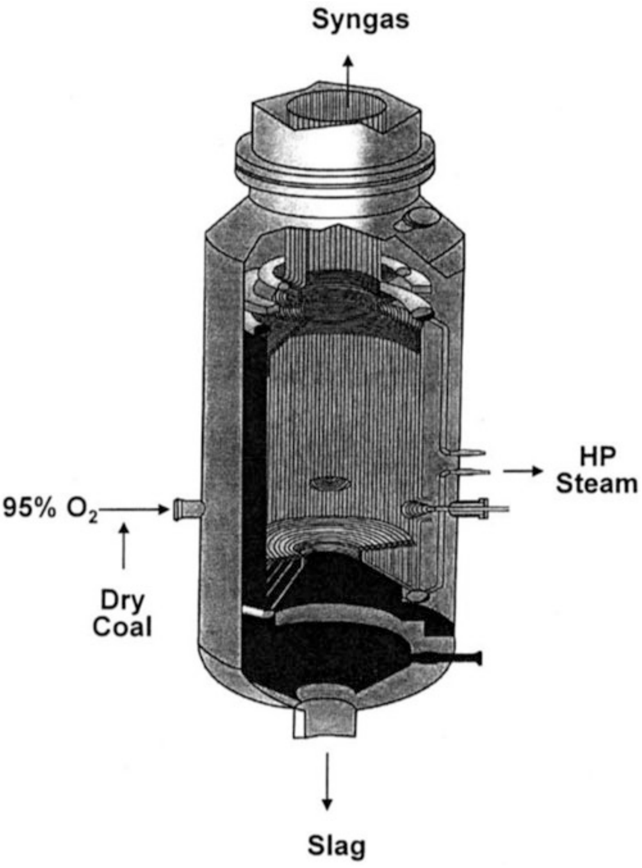


Fig. 19.23 Schematic for a shell gasifier

maintaining high availability and reliability. About 70 GEEGs are operating in 2010.

Shell Gasification Process. Shell's experience with coal gasification started with operation of a 6-t/day pilot plant in Amsterdam. This was followed by a 150-t/day unit in Hamburg, Germany, and a 250–400-t/day unit in Houston, TX. This led to the successful operation of a 2,000-t/day commercial unit.

The Shell gasifier (Fig. 19.23) consists of a pressure vessel containing a gasification chamber whose inner wall consists of closely spaced tubes. The wall temperature is controlled by circulating water through these tubes to generate saturated steam. Coal is ground (90% less than 100 μm) and dried in an inert atmosphere and pneumatically transported to the coal pressurization and feeding system. Coal, oxygen, and steam are injected into the gasifier through pairs of opposed burners. Oxygen purity is typically 95%, but other purities can be used, depending on the application. Gasifier pressure is 20–40 bar, and temperature is above 1,500°C.

Most of the mineral content of the feed coal leaves the gasification zone in the form of molten slag. The high operating temperature ensures that the molten slag flows freely down the tube walls and into a water bath at the bottom of the gasifier, where it solidifies into dense glassy granules that are washed, depressured, and sent to storage or disposal. Greater than 99% carbon conversion is obtained. Because of the high temperature, no hydrocarbons heavier than methane exist in the raw gas. The insulation provided by the molten slag layer minimizes heat loss.

The raw gas leaving the gasifier is quenched with recycled cooled product gas to solidify any entrained slag droplets. The gas then enters a syngas cooler that generates high-pressure steam. The cooled gas flows through filters or cyclones for fly slag removal. The recovered fly slag can be recycled via the coal feed system to enhance gasification efficiency. The syngas goes to a scrubber or removal of remaining solids and water-soluble contaminates and then to an acid gas removal system. Recovered H_2S is converted to elemental sulfur. Overall efficiency (including high-pressure steam production) is very high.

The Shell process can handle a wide variety of coals, varying from lignite to bituminous, in an environmentally acceptable way and produce a high-purity, medium heating value gas for use as a fuel for power generation, as a chemical feedstock, or as a source of hydrogen. As of 2010, 45 plants were operating.

Lurgi Dry Ash Gasifier. The Lurgi Dry Ash Coal Gasification Process (Fig. 19.17) is among the earliest commercially successful gasifiers. The first commercial plant was built in 1936. Since then, numerous plants have been built. In particular, Lurgi gasifiers provide the syngas for the Sasol indirect coal liquefaction (ICL) plants in South Africa (greater than 90 gasifiers) and also for the Dakota Gasification synthetic natural gas plant (14 gasifiers).

Sized coal enters the top of the bed by way of a lock hopper and moves down through the bed. Ash is withdrawn by a rotating grate and is removed from the gasifier through another lock hopper. Steam and oxygen enter the bottom of the gasifier through the grate and react with the coal as the gas moves up through the bed. As the coal passes down through the gasifier vessel, it undergoes in turn drying, devolatilization, pyrolysis, gasification, and combustion. This countercurrent operation results in a temperature drop across the reactor. Temperature in the combustion zone near the bottom is in the range of 600°C, whereas gas temperature leaving the drying and devolatilization zone near the top is typically 125–280°C. The main factor affecting exit temperature is the moisture content of the feed coal. For feed coals high in moisture, such as lignite, the gasifier exit temperature is at the lower end of the range, and for a low-moisture bituminous coal, it is at the upper end.

The raw gas is quenched directly with recycle water to condense the tar it contains. The gasifier vessel is cooled by water jacketing, which generates a portion of the steam that is used in the gasifier. Sufficient steam is injected to keep the temperature at the bottom of the gasifier below the melting

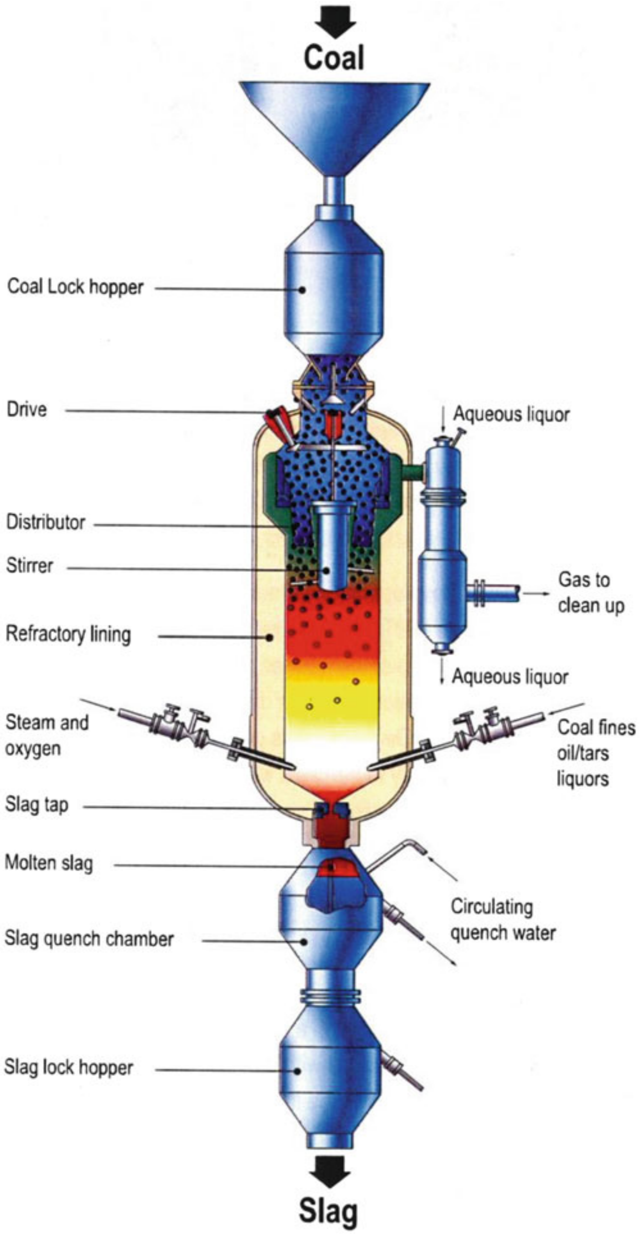


Fig. 19.24 Schematic for a BGL gasifier

temperature of the ash. The quenched gas is sent to cleanup for removal of NH_3, H_2S, and other impurities. After cleanup, the gas has a higher methane and hydrogen content than the gas from other gasifiers. Because of the lower exit temperature, oxygen requirement is reduced. However, recovery and disposal of tar can be a problem, or they can be a source of a variety of chemicals. High-reactivity, low-rank coal with a high ash softening temperature and minimal fines is favored as the feed.

BGL Slagging Gasifier. The BGL slagging gasifier technology (Fig. 19.24) is very similar to the conventional dry ash Lurgi gasifier, the key difference being that the BGL removes ash as molten slag. Advantages of slagging the ash include a significantly lower steam requirement and improved utilization of high-rank coals and coal fines. Slagging operation also eliminates liquid organic by-products by recycling them.

Like the Lurgi dry ash gasifier, the BGL gasifier is a high-pressure, moving-bed, oxygen-blown unit. Coal with a size distribution of 2 × 0 in. (the amount of −¼-in. material allowed depends on the coal's swelling characteristics) and/or briquettes are fed to the top of the bed via a lock hopper system. A rotating internal fuel distributor rakes and levels the top of the bed. The coal is gasified as it passes down through the reactor. Ash is removed from the bottom of the gasifier as molten slag through a slag tap, then quenched in water, and removed by a lock hopper. Oxygen and steam are injected through tuyeres near the bottom of the bed.

As with the Lurgi dry ash gasifier, there is a large temperature difference between the top and bottom, the magnitude of the difference depending mainly of the moisture content of the feed coal. For bituminous coals, this temperature difference is about 540–650°F. The raw gas leaving the gasifier is directly quenched to remove solids and condensable liquids. The solids entrained in the raw product gas, along with hydrocarbon by-products, such as tars and oils, naphtha, and phenols, are recycled to the top of the gasifier and/or reinjected into the gasifier at the tuyeres, where they are gasified. Coal fines can also be fed through the tuyeres, either pneumatically as dried solids or slurried with wastewater.

E-Gas™ Process. The first large-scale experience with the E-Gas™ gasifier (formerly the Destec gasifier) was gained in more than 34,000 h of operation at Louisiana Gasification Technology, Inc., from April 1987 through November 1995. This was followed by construction of an E-Gas™ gasifier at PSI Energy's Wabash River Generating Station in Indiana as part of DOE's Clean Coal Technology program. This plant started operations in 1995 and has the capacity to gasify about 1,000 t/h of bituminous coal.

The E-Gas™ gasifier consists of two stages: an entrained-flow, slagging first stage and a nonslagging second stage, as shown in Fig. 19.19 [58]. The first stage is a horizontal, refractory-lined vessel in which carbonaceous fuel is partially combusted with oxygen at an elevated temperature and pressure (775°C and 27 atm). The gasifier is initially heated with natural gas and then switched to the feed slurry, produced by wet grinding of coal. Oxygen and preheated slurry are fed to two opposed mixing nozzles, one on each end of the horizontal section of the gasifier. The oxygen feed rate is controlled to maintain the gasification temperature above the ash-fusion point. Under gasifier conditions, the feed is almost completely converted to a gas consisting primarily of H_2, CO, CO_2, and H_2O.

Coal ash and any added flux form a molten slag that flows continuously through a tap hole in the floor of the horizontal section into a quench water bath, where the slag solidifies. A continuous removal system transfers the slag to a dewatering bin, where the bulk of the slag settles out. The overflow goes to a settler for removal of the remaining fines. To improve carbon utilization, the fine particles are recycled to the slurry preparation area.

Raw syngas from the gasifier's first stage flows up into the vertical, refractory-lined, second stage, where additional slurry, but no additional oxygen, is injected. The injected fuel undergoes devolatilization and pyrolysis that both cools the syngas (because these reactions are endothermic) and increases its heating value (because of the nature of the products produced). In addition, water reacts with some of the carbon to produce CO and H_2. Unreacted char is carried overhead with the syngas that leaves the gasifier at 565°C.

The hot syngas is cooled to approximately 190°C in the syngas cooler, a vertical fire-tube heat recovery boiler with the hot syngas on the tube side, that generates 110 atm steam. The cooled syngas is sent to the particulate removal system that removes over 99.9% of the particulates. The particulate-free syngas is further cooled and then water scrubbed to remove chlorides and volatile trace metals. The syngas is then cooled to about 5°C before being sent to the acid gas removal system. During cooling, water condenses and dissolves most of the NH_3, along with some CO_2 and H_2S. This water is sent to the sour water treatment unit. The essentially sulfur-free syngas is moisturized and superheated and sent to the combustion turbine.

KRW Fluidized-Bed Gasifier. Development of this process was started by Westinghouse. Later, Kellogg Rust purchased an interest, and the technology was renamed the KRW process (Fig. 19.18). A 15-t/day unit was successfully operated at Waltz Mills, Pennsylvania. Based on this success, a KRW gasifier was included as part of a 100-MW(e) IGCC Clean Coal Technology project near Reno, Nevada. Owing to problems, mainly with the hot gas cleanup system, this unit only operated for short periods of time.

The KRW fluidized-bed gasifier consists of a vertical vessel with a smaller diameter at the bottom than at the top (Fig. 19.18). It is fitted with a central feed tube through which coal, sorbent (limestone), oxidant (air or oxygen), and steam are introduced. Solids are fed through a series of bins, which raise them from atmospheric to gasifier operating pressure. Coal and limestone are then pneumatically transported to the gasifier central feed tube. Oxidant is added to the feed tube, and the streams merge to form a central jet. The coal quickly devolatilizes, and the remaining char and limestone enter the gasifier bed.

Combustion of char and gas occurs within the jet to provide the heat necessary for the endothermic devolatilization, gasification, and desulfurization chemical reactions. Extraction steam from the steam turbine is also fed to the gasifier. CO and H_2 are the major combustible constituents of the product gas. Methane and other hydrocarbons are produced in small quantities during devolatilization of the coal. The gasifier temperature is high enough to crack the tars and oils produced. Because of the reducing atmosphere present during gasification, sulfur in the coal is released as predominantly H_2S. At conditions in the gasifier, limestone is rapidly calcined to form lime that reacts with the H_2S to form CaS. Because chemical equilibrium limits sulfur capture, remaining H_2S in the product gas is captured in a downstream desulfurizer.

As the carbon in the coal is gasified, the particles become enriched in ash. These ash particles tend to agglomerate and, along with dense calcium sulfide/oxide particles, separate from the char bed because of their higher density and different fluidization characteristics. This separation occurs primarily in the region surrounding the central feed tube at the bottom of the gasifier. These solids are cooled in the gasifier annulus by a countercurrent stream of recycle gas. The spent solids (unconverted calcined limestone, sulfide limestone, and ash) leaving the gasifier are transported to the ash collection system, consisting of a series of bins designed to reduce pressure back to atmospheric.

Gas exiting the gasifier flows through a cyclone for removal of entrained solids (char, ash, and sorbent) and then to the product gas cooler, where it is cooled to 540°C while generating 75 atm steam. Solids collected in the cyclone are returned to the gasifier, while the gas goes to sulfur removal before being burned in the gas turbine.

Siemens

The technology was developed originally by Deutsches Brennstoffinstitut in 1975 to gasify brown coal. The technology was acquired in 1991 by the Noell group and operated under other names before Siemens acquired it in 2006. The Siemens gasifier is a dry-feed, pressurized, entrained-flow reactor with the internal gasifier enclosed by a cooling screen (Fig. 19.25 left). The cooling screen consists of a gas-tight membrane that is studded and refractory-lined with a thin layer of silicon carbide. Molten slag cools and solidifies when it contacts the screen, and this layer protects it. Subsequent hot slag flows down the lag coated screen into a water bath where it solidifies to be removed through a lock hopper. For coals with less than 2% ash, the gasification zone is lined with a refractory (Fig. 19.25 right).

This gasifier was used at the Schwarze Pumpe complex in Saxony, Germany, that began commercial operation in 1984. Recently, commercial-scale gasifiers have been installed by Siemens in the US land in China. The company appears to be pursuing a very aggressive campaign in China.

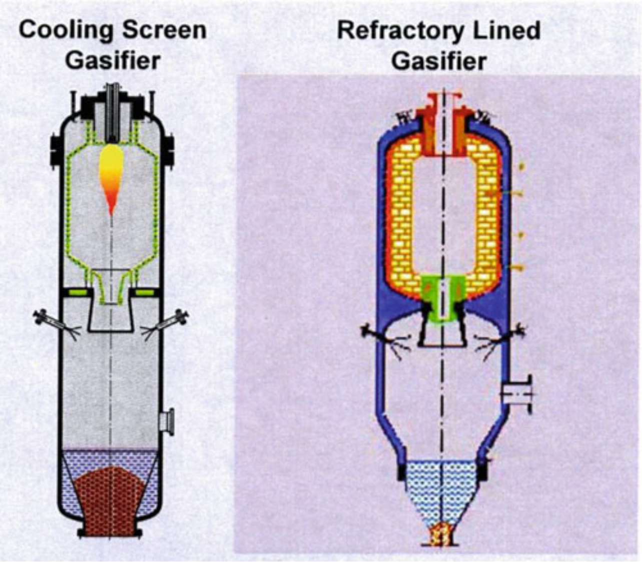


Fig. 19.25 Schematic of the Siemens gasifier for high (>2%) (*left*) and low (<2%) (*right*) ash content coals

Mitsubishi Heavy Industries

The Mitsubishi Heavy Industries (MHI) gasification technology was developed with the goal to maximize the gasifier efficiency (e.g., dry feed using a unique two-stage reactor design) and reduce costs (e.g., air blown to eliminate power consumption and capital cost of an oxygen plant) [59]. A joint venture of several companies and the Japanese government built a 2-t/day PDU and later a 200-t/day pilot plant to test the technology. In 2004, a 1,700-t/day IGCC demonstration plant was built, and the plant started operation in 2007. The MHI gasifier is a pressurized, upflow, entrained-flow slagging reactor with a unique two-state operation, similar to the E-Gas gasifier. Unlike E-Gas, it is a dry-feed operation that uses an internal membrane wall to protect the reactor, similar to the Shell and Siemens designs. MHI's current focus is the utilization of air or enriched air-blown IGCC application.

Compact Gasifier

Pratt & Whitney Rocketdyne has been applying its experience with rocket engines to develop a compact (90% smaller) and cheaper (one-half the cost) gasifier [60]. This gasifier is at an early stage of development.

Other Gasifiers

A number of gasifiers are being introduced for small-scale processes, especially in China. An example of such gasifiers is illustrated by the Synthesis Energy Systems (SES) utilization of the U-Gas technology. Eight 150-t/day U-Gas plants were installed in Shanghai, China, in 1995. Recently, SES has installed two 400-t/day gasifiers in Zaozhuang City, China, and another SES unit is to come on stream in 2011 in China. These SES gasifiers utilize the cheapest coal available, frequently approach 50% ash, and may even exceed this amount. These units are very flexible with a large turndown capability and rapid variation in coal quality.

These smaller gasifiers find wide usage for chemical synthesis, especially for use in a nearby industry. Thus, at this time, China is the scene of widespread activity in gasification and is constructing both small and large gasifiers at a rapid pace.

Underground

The desire to eliminate underground mining and replace it be underground gasification is not new. In 1913, Lenin wrote in Pravda, "Free the miners from going underground by applying underground gasification." During the past 100 years, interest in underground gasification has surged and ebbed [61]. Today, environmental concerns are pushing interest in the topic. The current status of underground coal gasification (UCG) has recently been summarized by Couch [62].

UCG has been practiced in Russia for about 50 years; however, it is possible to operate a process under a communist government under conditions where it would not be practicable under a free market situation. Thus, the Yerostigaz facility in Uzbekistan is the only remaining UCG remaining in the former Soviet Union and Linc Energy now owns more than 90% of this operation. Today, it appears that Linc Energy has demonstrated the first long-term UCG in the Western world. Linc Energy has now opened its fourth syngas generator and is currently developing a Fischer–Tropsch process to utilize the syngas generated by UCG. In South Africa, the electricity supply company Eskom, a state-owned enterprise, is now operating a pilot plant. Sasol and Eskom have joined to develop an environmental impact assessment for UCG in the Mpumalanga coalfields.

Biomass Gasification

One approach to decrease CO_2 emissions is to gasify a coal–biomass mixture [63]. This is not without problems because of their differences in mass and pyrolysis behavior, but its attractiveness to decrease GHG emissions pushes work in this area forward. Thus, the higher reactivity of the biomass leads to pyrolysis at a lower temperature, but this also leads to more tar formation, and it may be necessary to devise methodologies to deal with these tars. The status of US biomass gasification has been reviewed recently [64].

Co-gasification of coal and biomass was demonstrated at the Polk Power Plant in Tampa, Florida, but this was accomplished up to one and a half percent of woody biomass (by weight) being gasified together with a bituminous coal [65]. At the NUON power plant in the Netherlands, they successfully fed a mixture of 30% by weight demolition wood and 79% coal to a Shell dry-feed, high-pressure, entrained gasifier. The wood was reduced to sawdust (1 mm size) and was fed using the existing lock hopper feed system.

Major issues to be defined are the development of high-pressure biomass feed systems, gasifier optimizations for co-gasification at varied biomass feed rates, overcoming hurdles associated with feeding heterogeneous biomass types, design of biomass processing systems, and choice between biomass preprocessing at the plant or at the harvest site [66].

Future of Gasification

It appears that gasification will continue to be utilized at an expanding scale as recent reviews of the US [67] and world [68] programs indicate. In this regard, Collot has recently presented approaches for matching the gasification technology to coal properties. Abadie and Chamorro [69] have provided guidance by providing a market-based approach for the economics of gasification. The Shell, GE, Texaco, High-Temperature Winkler, and E-Gas processes can be classified as high-pressure, oxygen-blown, pulverized-coal, entrained-flow gasifiers having different geometries, operating pressures, and quench mechanisms. A surprising conclusion is that plant availability is the most important factor [70].

Coal Liquefaction

The primary purpose of any coal liquefaction process is to produce a liquid product with an increased hydrogen-to-carbon ratio, a reduced heteroatom (sulfur, nitrogen, and oxygen) content, and a minimal inorganic content. This has been accomplished by a variety of methods, which can generally be grouped into three major categories: pyrolysis, direct liquefaction, and indirect liquefaction. In pyrolysis, coal is heated to generate gases and liquids/tars while leaving a char behind. The liquids/tars are condensed and further treated to produce conventional liquid fuels. A process variation (hydropyrolysis) produces higher-quality products by carrying out the pyrolysis in a hydrogen atmosphere. In DCL, powdered coal is mixed with a solvent and, usually, a catalyst, and heated under pressure in a hydrogen atmosphere to produce gaseous and liquid products. The liquids are separated from ash and unconverted coal and refined to useful fuels. In indirect liquefaction, the coal is first gasified to produce syngas, which is then converted in the presence of a catalyst at an appropriate temperature and pressure to liquid products. Each of these approaches is discussed in more detail below.

Pyrolysis-Based Processes

When coal is heated, it undergoes thermal decomposition, which results in the evolution of gases and liquids/tars from the char that remains. In all coal conversion and utilization processes (including combustion, gasification, and liquefaction), some form of coal pyrolysis occurs. Because coal pyrolysis is an important process for making metallurgical coke, a vast amount of literature exists on the topic. By carefully varying pyrolysis conditions and optionally adding a reactive atmosphere (such as hydrogen), one can optimize production of a desired product (solid, liquid, or gas).

Coal pyrolysis is a very old technique (dating back to the eighteenth century), based on relatively inelegant technology. Most pyrolysis systems in use in the late 1800s and the early 1900s were in Europe and had as their objective the production of smokeless fuel (char) for domestic use. However, within a short time, it was realized that the coal tar fraction contained valuable chemical products. However, as inexpensive petroleum appeared on the scene, interest in coal by-products faded. A detailed review has been published [71].

Interest in pyrolysis was rekindled by the oil embargo of 1973, and emphasis was placed on developing pyrolysis processes that would achieve a maximum yield of liquid products capable of displacing petroleum as a source of distillate fuels. Several processes were developed to the pilot-plant stage in the 1970s, but development was discontinued when oil prices softened in the early 1980s, and none of these processes achieved commercial success.

Pyrolysis conditions determine the extent to which the coal is carbonized and the nature of the liquid and gaseous products. Coal pyrolysis processes are generally classified as low temperature (<700°C), medium temperature (700–900°C), or high temperature (>900°C). A number of reviews on advanced pyrolysis processes are available [72–76]. Two highly developed pyrolysis processes were the COED process, developed by FMC Corporation, which used a series of fluidized beds operating at successively higher temperatures, and the TOSCOAL process, which used a horizontal rotating kiln.

Table 19.14 Summary of effects of process parameters on coal pyrolysis

Heating rate
• A rapid heating rate increases liquid/gas yield and reduced char yield • A rapid heating rate produces poorer quality (lower H/C ratio) tar than that obtained at a slower heating rate [29, 35] • A rapid heating rate (in a reactive gas ‘atmosphere) produces char with increased open structure and reactivity • A rapid heating rate increases the thermoplastic (softening and swelling) behavior of coal • Achieving a rapid heating rate requires a sophisticated (often expensive) system
Temperature
• Low-temperature operation (500–700 °C) improves liquid yield • Temperature affects heteroatom distribution among char, liquid, and gas • At elevated temperatures (>1,300 °C), inorganics are removed as slag • Lower temperatures require a longer residence time for complete reaction
Pressure
Inert gas atmosphere
• Higher pressure operation reduces reactor size (i.e., increases throughput) • Higher pressure reduced tar yield • Coal feeding and product separation are more difficult at high pressure • Higher pressure improves gas–solid heat transfer
H_2 atmosphere
• Improves the yields of liquid and lighter products • Requires a sophisticated pressure control system • May increase undesirable agglomerating properties of coal [81] • H_2 cost must be compared to the increased value of the product
Other atmospheres (H_2O, CO_2, CO, CH_4, CS_2)
• Probably improve liquid–gas yield • Little information available
Vacuum
• Plastic behavior of coal is reduced [48] • Increases liquid/gas yield [48] • Difficult to achieve gas–solid heat transfer (solid–solid heat transfer feasible) • Little information available
Particle size
• Smaller particle size improves gas–liquid yield • Smaller particle size reduced secondary reactions • Grinding cost increases with the reduction in size
Coal rank
• High-volatile A (HVA) bituminous coals produce the largest quantity of tar • Lignites are rich in oxygen functional groups that lead to overall reduction in the calorific value of the product • The type of sulfur (pyritic vs. organic) present influences sulfur distribution among the products

Factors Affecting Coal Pyrolysis

During pyrolysis, the yield of gaseous and liquid products can vary from 25% to 70% by weight, depending on a number of variables, such as coal type, type and composition of the atmosphere present, final pyrolysis temperature, time-temperature path, and pressure. Although certain operating conditions may lead to increased product yield, achieving these conditions may result in increased costs.

Many excellent reviews [7, 77–80] have been published that discuss the factors affecting coal pyrolysis and product composition. In the following sections, major conclusions presented in the literature are discussed (as summarized in Table 19.14).

Temperature and Heating Rate. Coal undergoes many physical and chemical changes when heated gradually from ambient temperatures to approximately 1,000°C. Some changes, such as carbon–carbon bond scission, are observed before the onset of thermal decomposition that occurs above 350°C. When heated to approximately 100°C, physically sorbed moisture is liberated. Heating low-rank coals, such as lignites, that contain appreciable carboxylic functional groups will liberate carbon dioxide by thermal decarboxylation. Over 50% of the carboxylic acid functional groups are lost in the temperature range 100–250°C.

As the temperature of thermal treatment increases to the range 200–400°C, coal produces a number of lower molecular weight organic species (especially aliphatic compounds), which are believed to arise from components that are loosely bound to the more thermally stable part of the coal structure. At a higher temperature (375–700°C, depending on the heating rate), thermal destruction of the coal structure occurs, as reflected by the formation of a variety of hydrocarbons, including methane, other alkanes, polycyclic aromatics, phenols, and nitrogen-containing compounds. In this temperature range, bituminous coals soften and become plastic (thermoplastic) to varying degrees.

At still higher temperatures (600–800°C, depending on the heating rate), the plastic mass undergoes repolymerization, forming semicoke (solid coke containing significant volatile matter). At temperatures exceeding 600°C, semicoke hardens to form coke with the evolution of methane, hydrogen, and traces of carbon oxides. Pyrolysis of coal is essentially complete at approximately 1,000°C. The temperature at which the maximum devolatilization rate occurs depends on the heating rate. For a slow heating rate (about 5°C/s), the maximum rate occurs at around 400°C, whereas for a rapid heating rate (about 10°C/s), the maximum rate might not occur until 900°C [81, 82]. Maintaining the coal at the final temperature for an extended period normally increases the yield of volatile material, because char decomposition is a relatively slow process.

Because pyrolysis reactions do not occur at sharply defined temperatures, the heating rate has a marked effect on the nature and distribution of pyrolysis products, as summarized in Table 19.14. Solomon and coworkers conducted extensive work on the kinetics of coal devolatilization, and many reviews are available [83].

Coal Rank. The type of coal strongly influences pyrolysis behavior. Low-rank coals, such as lignite, contain oxygen functional groups that evolve water and carbon oxides upon pyrolysis. Higher-rank bituminous coals contain less oxygen; consequently, these coals produce significantly less water and carbon oxides when pyrolyzed. The nature of the tar produced is also dependent on coal rank. Bituminous coal tars tend to be more aromatic (and relatively more thermally stable) than the tars generated from lignites. When heated, bituminous coals soften, become plastic, and swell to varying degrees, whereas lower-rank coals generally do not become plastic. However, at a rapid heating rate (about 10°C/min) or elevated pressure, certain lower-rank coals may melt and demonstrate some plastic and swelling characteristics.

Other Factors. Several other factors influence, at least to some extent, the course of the pyrolysis process. These include particle size, bed configuration, pressure/vacuum during pyrolysis, nature of the coal ash, secondary reactions, etc. [84]. It is beyond the scope of this chapter to consider these items, but the interested reader can find additional information in the literature, including reports on pressure effects [45, 85], effect of vacuum [48], effect of inorganics [86], and effect of a reactive atmosphere [48, 87].

Utilization and Characterization of Pyrolysis Products

Efficient utilization of all the products, solid, liquid, and gaseous, is essential if favorable economics for a pyrolysis process are to be achieved. Products may require varying degrees of treatment before they are usable. The stream exiting the pyrolyzer requires separation of gas, liquid, and particulates, similar to the situation for a fixed-bed gasifier (i.e., cold gas cleanup). Possible operations include hot dust removal cyclones, quench/particulate scrubber towers, and venturi scrubbers to remove tar mist. In general, gas stream sulfur removal and wastewater treatment are also required. Finally, environmental factors, such as toxicity [88], carcinogenicity, and mutagenicity of the coal pyrolysis liquids (CPL), need to be considered. For example, the mutagenicity of CPL is strongly dependent on the conditions of pyrolysis (temperature, coal type, and atmosphere during pyrolysis).

Liquid fuels markets tend to have product specifications that do not vary widely. However, the characteristics of CPL can vary greatly, depending on coal type and operating conditions. In general, the H/C atomic ratio for CPL is lower than that for petroleum products, which is in the range 1.8–1.9 for No. 2 fuel oil and 1.7–1.8 for No. 6 fuel oil. In contrast, the H/C ratio for CPL is in the range 0.9–1.5, indicating that CPL is significantly more aromatic (50–70%) than petroleum-derived products.

The uses of CPL can be divided into two broad categories: (1) direct combustion, requiring little or no upgrading, and (2) transportation fuels and chemicals, requiring extensive upgrading. The high aromaticity of CPL indicates that conversion to gasoline is preferred to conversion to diesel fuel. Much of the attention in low-temperature tar processing has been devoted to hydroprocessing techniques, such as hydrotreating and hydrocracking, with the primary objectives of reducing viscosity, reducing polynuclear aromatics, and removing heteroatoms (sulfur, nitrogen, and oxygen) to produce usable fuels and chemicals.

The cost of H_2 is still the primary impediment to tar upgrading. Although the tar fraction can be used as a source for chemicals, such as phenolics, road tars, preservatives, and carbon binders, these uses do not constitute a large enough market to support a major industry. Likewise, the char has a range of potential applications, but the char must provide an economic return for pyrolysis processes to operate profitably. One obvious application is on-site combustion to generate heat for the pyrolysis process. Alternatively, the char can be combusted in an electric power plant. Char could also be gasified to produce hydrogen for use in hydropyrolysis or tar hydrotreatment. Other applications include production of activated carbon and molecular sieves.

Coal Liquefaction

This approach to the conversion of coal to liquids was developed in Germany by Bergius starting in 1913 [89]. It was made possible by the advances in utilizing high-pressure processing that was based on the commercialization of the ammonia synthesis in Germany beginning about 1910. The use of high-pressure processing for DCL continued in Germany during the 1900s.

DCL refers to the process of converting coal to liquid products by mixing ground coal with a recycled process solvent and/or petroleum-derived residual oil (coprocessing). The slurry reacts in a hydrogen atmosphere at 750–850°F (400–450°C) and 1,000–2,500 psig (7–17 MPa). Under these conditions, the coal structure breaks down into a variety of products, some of which are in the gas phase and some soluble in the liquid phase. The highly aromatic products are separated from the ash and unconverted coal by solvent deashing (SDA) or vacuum distillation and upgraded by conventional petroleum refining techniques to gasoline, kerosene, diesel and jet fuels, and other distillate fractions.

The success of any DCL is highly dependent on the type of coal used. Coal rank influences both overall conversion and product distribution. Lower-rank coals have been reported to give both higher and lower conversions than bituminous coals [90], with most evidence supporting the latter. Furthermore, the liquid products from lower-rank coals are generally more volatile and of lower molecular weight [91] than those from bituminous coals.

German organizations demonstrated commercial-scale operation of DCL during World War II, employing the Bergius process, which they had developed. After World War II, the United States constructed a 200-bbl/day DCL plant in Louisiana, Missouri, based on the German technology. This plant was operated by the Bureau of Mines from 1949 to 1954. However, as the extent of petroleum deposits in the Middle East and other locations became apparent and fears of an oil shortage subsided, interest in DCL diminished. For information on the early history of DCL, see Wu and Storch [92], Donath and Hoering [93], Ministry of Fuel and Power [94], and Perry's Chemical Engineering Handbook [95].

The 1973 oil embargo created a new surge of interest in DCL as one of a number of routes to produces petroleum substitutes from indigenous resources. From this renewed effort in the early 1970s, three major DCL processes emerged: Solvent Refined Coal (SRC-I and SRC-II), Exxon Donor Solvent (EDS), and H-Coal. These processes converted coal to liquids in a single reaction stage and, by the late 1970s, had been developed at pilot-plant scale, with designs underway for demonstration plants. However, in the early 1980s, these plants were canceled due to projected unfavorable economics. Interest then shifted to developing a two-stage process, which promised higher yields and better quality product with a concomitant improvement in economics.

Chemistry of Direct Coal Liquefaction

As coal is heated in a suitable solvent, it is rather rapidly (1–5 min) converted into soluble entities that have a high molecular weight and an average elemental composition relatively little changed from that of the moisture- and ash-free (maf) feed coal. As residence time is extended, hydrocracking converts the dissolved entities to lower-boiling liquids with reduced heteroatom content [96]. The exact product slate depends on many factors, including operating conditions, coal properties, and the solvent used. The dissolution reactions produce free radicals, which may undergo retrograde reactions to form insoluble coke or, if there is sufficient hydrogenation activity, be stabilized by reaction with hydrogen atoms. The stabilized free radicals can then undergo hydrocracking reactions to form a wide range of products from methane through heavy oil. Some hydrogen comes from the gas phase, but most is transferred from the solvent. The most effective hydrogen transfer agents (hydrogen donor solvents) are hydroaromatic compounds, which can lose hydrogen and revert to polynuclear aromatics and later to be rehydrogenated to hydroaromatics. The mechanism of these and other DCL reactions is not fully understood, and since every coal molecule is unique, it is unlikely that precise kinetics will ever be possible.

Because of the high concentration of ring structures in coal, coal liquids are highly aromatic, particularly the heavier fractions. This can be a problem, since some polynuclear aromatics are carcinogenic. Even though the carcinogens can be eliminated from the final product by limiting the endpoint of the distillate to about 370°C and recycling heavier components to extinction, care must be taken to prevent exposure of plant workers. Another problem is that the current trend is to limit the aromatics content of transportation fuels for environmental reasons. Therefore, production of specification fuels from coal liquids can require extensive (and potentially expensive) hydroprocessing.

The conversion of coal, measured as the fraction that is soluble, depends upon the coal rank as shown in Fig. 19.26 for subbituminous and lignite coals. For the lower carbon containing coals, it is the lower reactivity of the structure that limits short-time conversion; with increasing reaction time, the higher conversion is obtained. The solid line at about 85 MAF coal represents the coal chosen for the early work by Berguis; he was either well informed or very fortunate in his choice.

One of the great advances made in DCL was the improvement of catalysis and the lowering of the reaction temperature. This led to a very significant reduction in the amount of light gases found in the product slate and therefore an increase in the amount of distillate (Fig. 19.27) [97].

Keogh and Davis [98–100] converted a number of coals at various reaction conditions and came to the conclusion that the production of individual fractions (insoluble, preasphaltenes, etc.) and three lumped parameters (oil plus gases, preasphaltenes plus asphaltenes, and insoluble organic matter) as a function of a severity index, defined in terms of reaction temperature and time, defined a single reaction pathway. From ternary plots of the three lumped parameters, along with plots of individual fractions vs. a severity index, they deduced reaction pathways for a number of coals during either thermal and/or catalyzed pathways. From inspection of the data, they were able to show, for example, that *the addition of a catalyst affects the reaction rate but it does not alter the reaction pathway* (Fig. 19.28). It appears therefore that the catalyst generates the hydrogen donor by adding gaseous hydrogen to aromatic structures which then stabilize radicals generated by the thermal reaction that converts higher to lower molecular weight compounds. The large pilot plants normally did not generate data for a wide range of conversions; however, the reaction pathway defined by Keogh and Davis [99] in small reactors applied to the large pilot-plant data (Fig. 19.29).

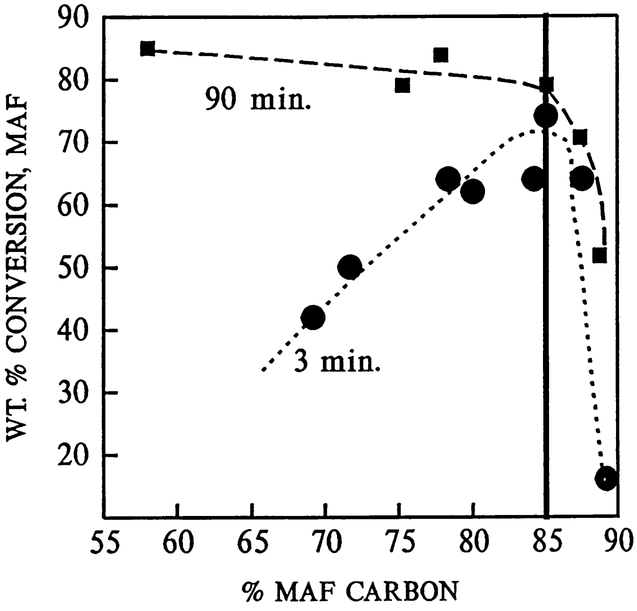


Fig. 19.26 Dependence of coal conversion upon moisture- and ash-free (MAF) carbon content (Adapted from Whitehurst DD, Mitchell TO, Farcasiu M (1980) Coal liquefaction: the chemistry and technology of thermal processes. Academic Press, New York)

The first DCL processes were single-stage processes, in which both coal dissolution and hydrocracking take place in the same reactor. Later, it became apparent that these two steps could not both be optimized at the same operating conditions. This led to the development of a two-stage process, in which a short contact time (SCT) thermal liquefaction reactor was followed by a catalytic and upgrading reactor to optimize conditions for both reactions, thus leading to higher yields, more efficient use of hydrogen, and a better product distribution. Other variations of two-stage processing, such as the integrated two-stage liquefaction (ISTL) process, were developed later. Tables 19.15 and 19.16 present typical compositions and yields, respectively,

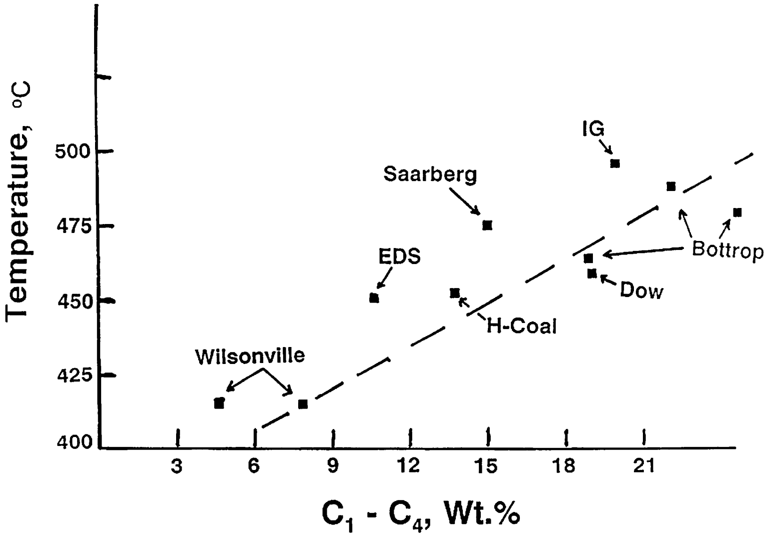
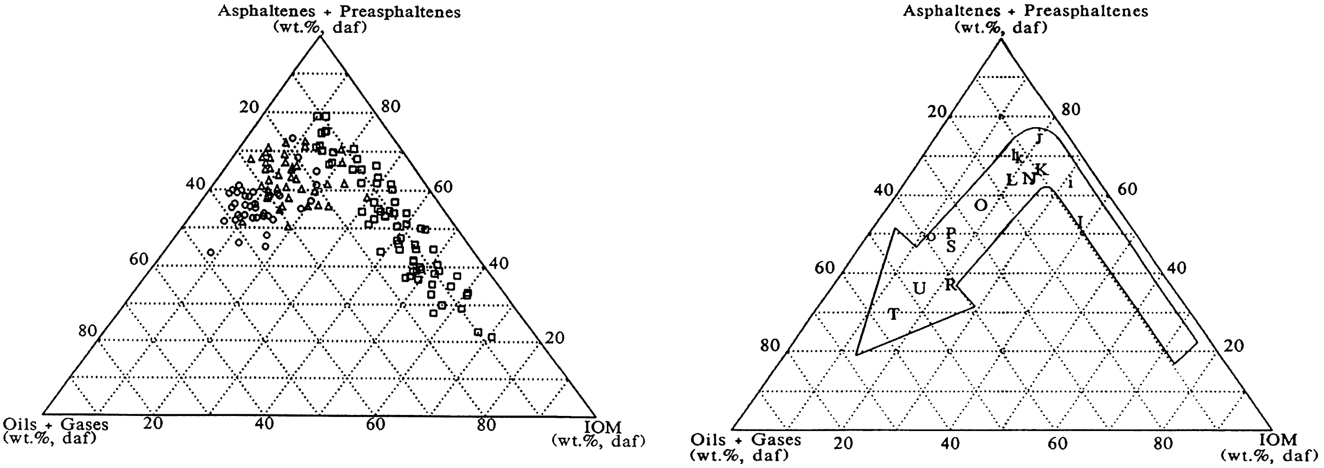


Fig. 19.27 Relation of light gas production to the operating temperature



Catalyst	Residence Time (min.)	Reaction Temperature (°K)		
		658	700	718
Shell 324	5	I	N	R
Shell 324	15	J	O	S
Shell 324	30	K	P	T
Shell 324	60	L		U
Molybdenum Naphthenate	5	i		
Molybdenum Naphthenate	15	j	o	
Molybdenum Naphthenate	30	k		
Molybdenum Naphthenate	40	l		

Fig. 19.28 (*Left*) Conversion of 69 bituminous coals in tetralin at 15 min at 385°C (*open squares*); 427°C (*open triangles*) and 4457°C (*open circles*) (98–100). (*Right*) Catalytic conversion of Western Kentucky coal (98–100)

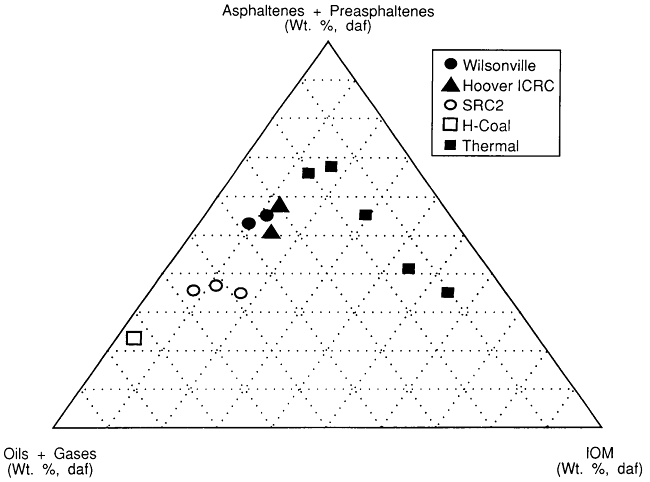


Fig. 19.29 Ternary plot of results from pilot plants (thermal and Hoover ICRC) and 6–200-t/day pilot plants

of distillates from various DCL processes. Table 19.17 provides a listing of large plants operated in the United States during the 1970–1980s period, and Table 19.18 lists major non-US plants. Table 19.19 presents a typical product distribution obtained at some of these plants.

Solvent Refined Coal (SRC-I and SRC-II)

The Solvent Refined Coal (SRC-I, process began by Spencer Chemical, later acquired by Gulf Oil, in 1962) was designed as a means of deashing coal [101]. It soon became apparent that the cost for the improvement in the quality of the coal that was achieved by the SRC-I process was too large for the process to be viable and SRC-I was abandoned. Gulf Oil continued to operate the Fort Lewis plant in the SRC-II mode which involved recycling the ash that contained iron, a catalyst for the liquefaction process, as well as operating with iron added to the feed. The SRC-II plant operated to provide sufficient SRC for combustion testing in a large power plant; however, the SRC-II had operating difficulties in maintaining an adequate production of recycle solvent. The smaller SRC-I plant at Wilsonville, Alabama, continued to operate and moved on to verify several processes that involved two-stage operations as described below.

Exxon Donor Solvent Process (EDS)

Research on the EDS process began in 1966. Exxon workers investigated a variety of process options before they settled upon a process scheme in which coal is mixed with a recycle

Table 19.15 Heteroatom content of products from direct liquefaction processes

	Wt.% in naphtha (C_5-390 °F)				Wt.% in distillate (400–975 °F)			
Process	O	N	S	H	O	N	S	H
H-Coal	0.6	0.3	0.07	13.19	1.1	0.5	0.08	9.52
EDS	2.42	0.23	0.28	12.03	2.0	0.6	0.3	8.97
SRC-II	3.51	0.42	0.26	10.46	3.63	0.99	0.25	9.27
ITSL	0.41	0.14	0.01	13.78	0.56	0.30	0.015	10.21

Table 19.16 Distillate yields from various direct liquefaction processes

Process	Configuration	Distillate (wt.% Maf. coal)	Yield (bbl/tons Maf. coal)	Distillate quality (gravity API)	Heteroatoms (wt.%)		
SRC-II (1982)	One-stage noncatalytic	41	2.4	12.3	1.0	2.33	0.33
H-Coal (1982)	One-stage noncatalytic	52	3.3	20.2	0.50	1.0	0.20
Wilsonville (1985)	ITSL thermal–catalytic	62	3.8	20.2	0.25	1.9	0.23
Wilsonville (1986)	CC-ITSL catalytic–catalytic	70	4.5	26.8	0.16	<1	0.11

Table 19.17 US pilot plants for coal liquefaction

Plants and location	Sponsor(s)	Capacity (t/day)	Conversion methods	Date
Project-Gasoline, Cresap, WV	Office of Coal Research/Consolidation Coal Co.	20	Extraction + catalytic hydrogenation	1967–1970
COED, Princeton, NJ	Office of Coal Research/FMC	36	Staged pyrolysis	1970–1975
H-Coal, Trenton, NJ	HRI, ARCO, Ashland Oil, Esso, Amoco	3	Hydrogenation in ebullated catalyst bed	1968–1983
Solvent refining (PAMCO), Ft. Lewis, WA	Gulf (PAMCO)/ERDA	50	Dissolution + filtration	1974–1983
Solvent refining, Wilsonville, AL	Southern Services	6–12	Dissolution + filtration, catalytic hydrogenation	1974–1992
Exxon donor solvent	Exxon	1	Extraction–hydrogenation	1975–1984
Exxon donor solvent	Exxon, US DOE	250	Extraction–hydrogenation	1980–1983
H-Coal, Ashland, KY	Ashland, and up to 12 other companies and US DOE	200–600	Hydrogenation in ebullated catalyst bed	1980–1983

Table 19.18 Non-US direct coal liquefaction processes (post-1950)

Plant and location	Sponsor(s)	Capacity (t/day)	Conversion methods	Catalyst	Date of operation
IGOR; Bottrop, Germany	German government, Veba Oel and Ruhrkohle	200	Extraction/hydrogenation	Red mud[a] NiMo–alumina[b]	1981–1987
Point-of-Ayr, UK	British Coal Corp.	2.5	Stirred tank-type[a] fluidized bed[b]		1983–1995
NEDOL	New Energy and Ind. Technol.	150	Extraction/hydrogenation	Natural pyrite	1996–1998
PYROSOL, Germany	Saarbergwerke AG	6	Countercurrent	None	1972–1988
BCL, Australia	Australia and Japan governments	50	Fixed bed	Iron based[a] NiMo–alumina[b]	1980s
CT-5, USSR		7		Mo	1986–1990
Shenhua I, China	China Shenhua Energy Co., Ltd.	6	Suspended bed	Iron	2002
Shenhua II, China	China Shenhua Energy Co., Ltd.	3,000	Suspended bed	Iron	2008

[a] First stage

[b] Second stage

Table 19.19 Products from direct coal liquefaction in large pilot plants

	SRC-II	H-Coal	EDS[a]	Integrated two-stage	Bergius	Run 254 G Ohio coal	Dow	Bottrop (Ill. #6 coal)
Yields (wt.% of MAF coal)								
C_1–C_4 gas	18.8	14.8	10.7	5.1	13.2	7.9[a]	22.1	20
Naphtha	11.6	18.6	33.6	7.4	22.0	23.8	13.2	15
Mid-distillate	20.3	24.7	13.1	38.8	10	26.1	28.6	37
Heavy distillate	9.0	5.3		9.0	~10	21.4	4.9	10
Efficiency No. distillate/no. H_2 added	8.2	8.8	8.7	12	8.0	7.3	9.0	–
% Distillate % C_2–C_4	2.2	3.3	4.4	11	3.2	9.0	2.1	3.1

[a] Bottoms recycle

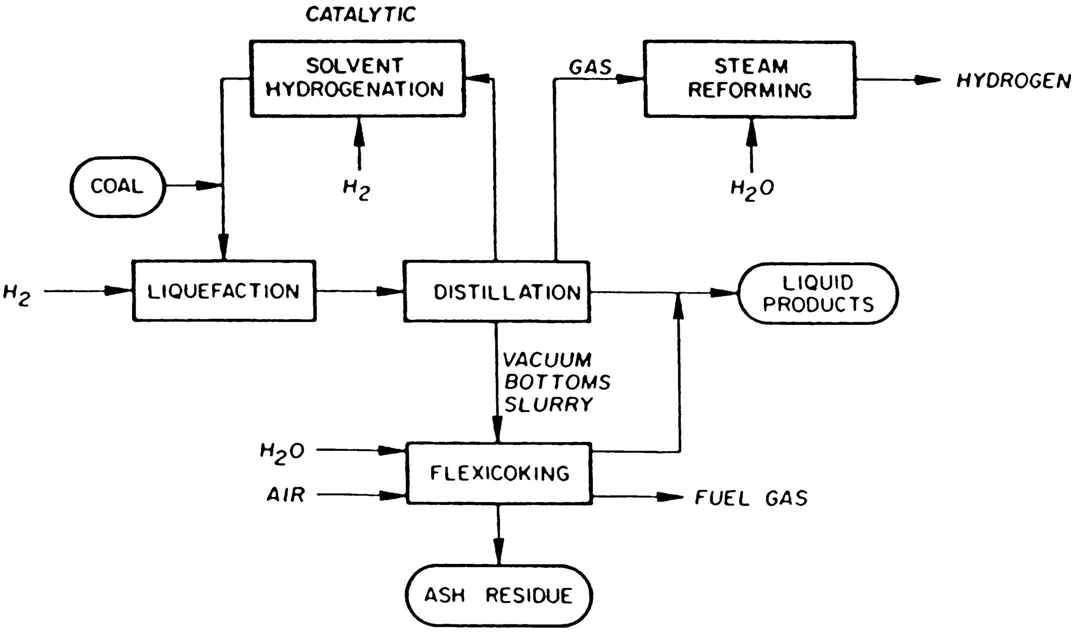


Fig. 19.30 Schematic of Exxon donor solvent process (Courtesy of The Pace Company, Denver, CO)

solvent that has been hydrogenated between passes to the liquefaction reactor and the mixture is then fed to the reactor together with high-pressure hydrogen (see Fig. 19.30). The recycle solvent is hydrogenated in a separate reactor before mixing with the coal. More than 2.6 barrels of a synthetic crude boiling below 525°C are produced from a ton of dry, high-volatile coal feed. The process development advanced to a 250-t/day pilot-plant stage that went into operation in June 1980 and operated during the next 3 years. At a smaller scale, it was demonstrated that the process could be used with a wide range of coals. Even though it was thought that the process was under control, it was learned that this was not the case. In spite of having a staff that was very familiar with operating large-scale refineries, they experienced many unexpected difficulties in operating the coal liquefaction plant. Operation of the plant indicated that there were a number of conditions that had to be overcome, and manageable solutions were found for these operating and mechanical problems. For example, it was learned that there was an optimum level of hydrogen content for the process solvent, and that over hydrogenation could be detrimental. Another important observation was that upon scaling up from the smaller plants to the 250-bbl/day plant, the coal conversion was 5–10% lower than in the smaller units. Radioactive tracer studies showed that the residence time distributions of the reactants change significantly as the reactor diameter was increased. The runs in the large ECLP reactor were terminated in 1983, and the pilot plant was dismantled. After 3 years of operation, it was reported that "the operating staff was now running the plant instead of the plant running the staff" [102].

Initially, the operation of the EDS process resembled the SRC-II process. A later modification of the EDS process utilized vacuum bottoms recycle, and this allowed for the production of a higher distillate yield. This processing scheme allowed for a longer contact time of the heavy, high-boiling bottoms recycle allowed for a higher yield of naphtha [103]. In addition to the longer residence time, the recycle bottoms contained ash which likely provided a catalytic impact [104]. The distillate yield was in the range 2.5–3.0 barrels of liquid per ton of coal.

Exxon continued research at intervals following the closure of the ECLP pilot plant. During this time, they developed the Exxon microcatalytic direct coal liquefaction technology (MCL) [105]. It was announced in 2009 that Accelergy Corp. had obtained a global leasing agreement that would allow them to produce liquid fuels and chemicals from coal resources using the MCL process. In one modification, the solvent hydrogenation could be eliminated because of the activity of the added catalyst, and this increased the thermal efficiency by 7%. In one case, the MCL step is integrated with an existing circulating fluid-bed power plant, and this provides a balanced plant efficiency of 74% (LHV) and produces 5.3 bbl of premium fuels per metric ton of coal [106].

H-Coal Process

The H-Coal process is an extension of the H-Oil process which was designed to upgrade heavy crude under high hydrogen pressure to lower-boiling products [107]. Both processes utilize an ebullated-bed catalytic reactor. A schematic of the H-Coal process with a single stage is illustrated in Fig. 19.31. Research began in 1964 at Hydrocarbon Research, Inc. (HRI, now a part of Headwaters, Inc.), using a bench-scale unit and then progressing to a 200–600-t/day pilot plant at Catlettsburg, Kentucky, that was completed in 1980 and operated through 1983. The time scale for the H-Coal process is typical of the other plants and shows that it would take about 20–30 years from inception to a pilot or commercial plant (Fig. 19.32).

The H-Coal process reactions are carried out at high temperature (ca. 450°C) and intermediate pressure (ca. 150 atm). The dried coal is mixed with a recycle solvent

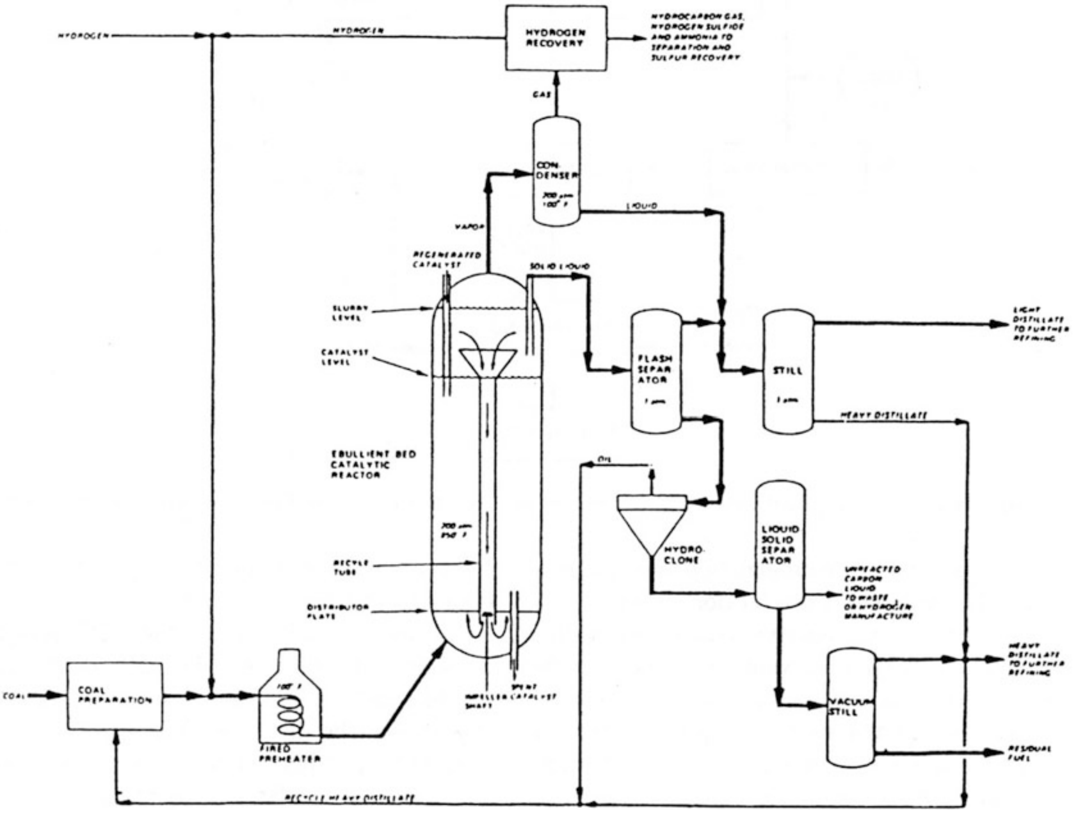


Fig. 19.31 Schematic of H-Coal process (Courtesy of The Pace Company, Denver, CO)

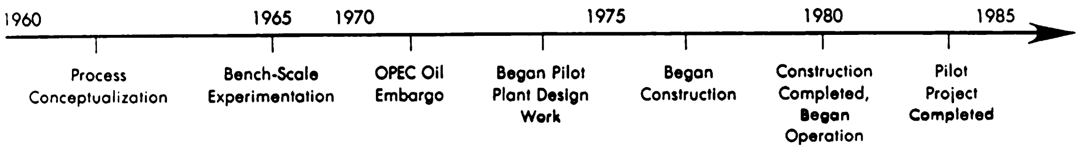


Fig. 19.32 History of development of the H-Coal process

and fed to the bottom of the reactor. In the H-Coal process, the catalyst was a typical hydrotreating catalyst (Co–Mo–alumina) that was maintained in an expanded bed by the upflow of hydrogen and the recycle solvent flow that was provided by an internal pump. The addition of fresh catalyst and the withdrawal of aged catalyst at the bottom of the reactor allowed for the maintenance of a constant catalyst activity allowing for a constant coal-conversion level and constant reactor productivity.

The H-Coal product has a high H/C ratio and a reasonably low heteroatom content. Because of the high temperature, the amounts of light gases that are formed are relatively high compared to the H-Oil process. Typically, between 14,000 and 20,000 standard cubic feet of hydrogen are required for each ton of coal processed. Approximately 3.0–3.5 barrels of liquid product were produced for each ton of coal. The high-volatile bituminous coal was found to be better suited for the process since the lower reactivity of lower-rank coals significantly reduced throughput and distillate yields. Ashland Oil completed the design for a 50,000-bbl/day plant to be located in Western Kentucky, but its construction was canceled due to the low price of petroleum.

German IGOR Process

In Germany, Ruhrkohle AG and Veba Oel AG operated a 200-t/day liquefaction plant that was initially based on the older German process (Deutsche Process, DT). The process was upgraded by adding the Integrated Gross Oil Refining (IGOR) stage and was operated during 1975–1993. The operations were supported by state and/or national government agencies. A 0.2-t/day plant was used in the initial work and was scaled to a 200-t/day plant located at Bottrop. The Bottrop plant utilized a dissolver reactor that contained a “red mud” iron catalyst as well as severe operating conditions (typically 470°C and 300 bar). Liquid yields for the process were in the 50–60% range (maf). A feasibility study for a 50,000-bbl/day plant was carried out, but the decrease in the price of petroleum crude eliminated support for the full-sized plant.

H-Tech

After the shutdown and dismantling of the IGOR/DT plant, key experts of the Kohleoel–Anlage Bottrop (KAB) pilot plant founded h-tech to pool the process know-how in engineering pilot plant, design, and operating plants and to further develop the technology [108]. It is claimed that the KAB process shows the highest conversion levels and liquid yields and thereby the lowest greenhouse gas emissions per barrel of product.

Pyrosol Process

A similar German process (Pyrosol process) for higher ash coals was developed by Saarbergwerke [109]. A 6-t/day pilot plant was operated between 1981 and 1984, but low oil prices prevented the expansion or even continuation of the program. A major difference between the two German processes was that the underflow is processed in a vacuum flash unit in the DT process, whereas in the Pyrosol process, it is carbonized in a hydropyrolysis step.

British Coal Liquid Solvent Extraction

Between 1973 and 1995, the Liquid Solvent Extraction (LSE) process was developed by the British Coal Corporation (Fig. 19.33) [110]. The early work led to the establishment of a 2.5-t/day plant at Point of Aye, North Whales, UK. The coal conversion is carried out in two or more thermal reactors followed by filtration. Following filtration, solvent recovery and hydrogenation is accomplished to produce additional products and recycle solvent. Virtually any coal can be utilized in the process. The process can be operated so only distillable is obtained with total distillate yields of 60–65 (daf), with virtually all boiling below 300°C. Essentially the total effluent from the reactor passed to a filtration drum and the reliable operation of this drum was never accomplished during the work.

Two-Stage Process

As DCL progressed, it was recognized that both coal dissolution and hydrocracking could not be optimized in a single reactor. The use of two reactors could enable the optimization to be made at different reaction conditions. This led to the introduction of a SCT thermal conversion followed by a catalytic upgrading reactor. This combination led to higher yields, more efficient use of hydrogen, and a better product distribution. The development involved both government and industrial organizations and proceeded in an evolutionary manner through a series of steps.

The first approach to separate the two stages was to utilize a SRC-I-type dissolution reactor and to place an upgrading reactor downstream. The operation of these two reactors was not integrated, meaning that the second stage did not contribute solvent to the first stage. Rather than use the typical fixed-bed reactor normally used for hydrotreating, an ebullated-bed reactor was used to avoid the plugging and rapid catalyst deactivation encountered by the fixed-bed reactor at the Cresap, WV plant.

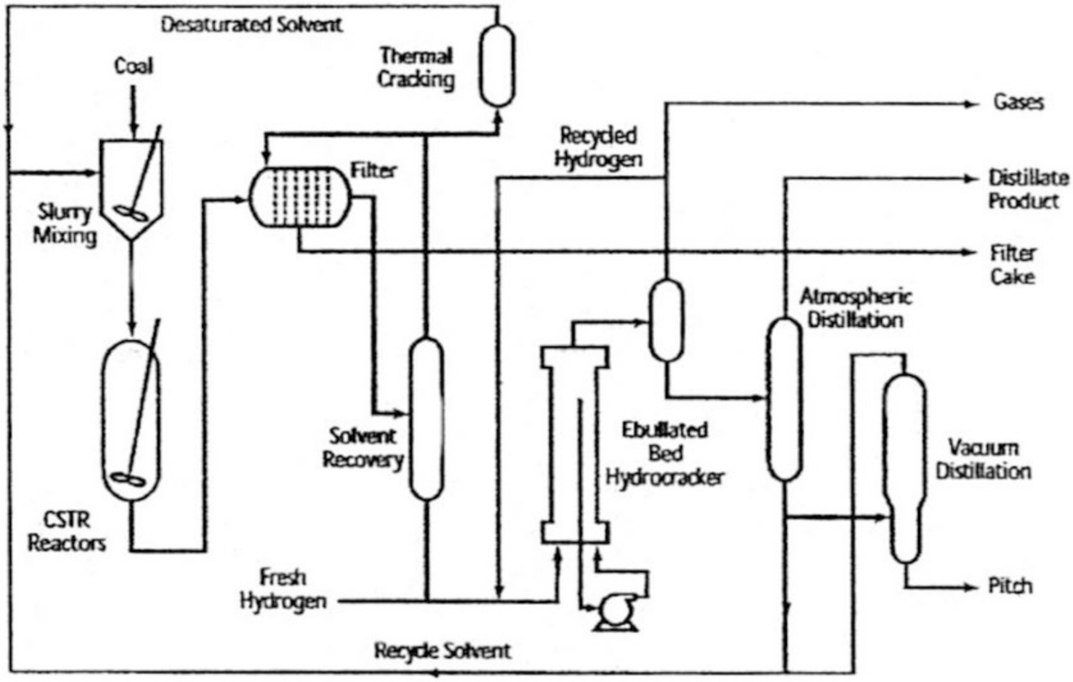


Fig. 19.33 Schematic of the LSE process (DTI, coal liquefaction. Clean coal technology program. Technology status report 010, London, UK, Department of Trade and Industry, 14pp, Oct 1999)

The ISTL process was developed next. In this process, hydrogenated resid is recycled to the first stage to act as solvent. SDA was used to remove ash from the recycled resid. Initially, the SDA was located between the first and second stages, but it was later located after the second stage to become the reconfigured two-stage liquefaction (RISTL). Contrary to expectations, the presence of ash in the second stage did not deactivate the catalyst. As an additional benefit, the product from the two-stage process was easier to upgrade than those from the single-stage process.

The next step in the continuing development was the introduction of the close-coupled integrated two-stage liquefaction (CC-ISTL) where the effluent from the first stage passes directly to the second stage, thereby improving thermal efficiency, reducing the retrograde reactions, eliminating the pressure letdown between reactors, and producing a higher-quality solvent [84, 104, 111–121]. Distillate yields of up to about 65% have been obtained for the conversion of Illinois No. 6 coal for CC-ISTL. A number of variations of the close-coupled process have been tried: thermal first stage/catalytic second stage (T/C CC-ISTL), catalytic ebullated bed for both stages (C/C CC-ISTL), and using a disposable small-particle catalyst rather than the pelleted catalyst.

In the most advanced version of the close-coupled two-stage process, disposable slurry catalysts (usually iron or molybdenum) were added with the coal with an integrated hydrotreated to process the distillate to produce a high distillate yield with a low concentration of heteroatoms.

The US and European units terminated large-scale activities in the early 1990s; following this period, DCL pilot and commercial plants were constructed and operated in the Far East [122].

In Japan, the New Energy and Industrial Technology Development Organization (NEDO) was established in 1980 to conduct activities that focus on research and development that is related to oil-alternative energy, technology for the efficient use of energy and industrial technology. Upon its formation, it assumed responsibility for three liquefaction processes: a solvent extraction process, a direct hydrogenation process, and a solvolysis process. By 1984, they had merged the three activities into a single concept and began the design for a 250-t/day pilot plant that was downsized to 150 t/day by 1988. Many days of testing in a 1-t/day PDU provided information for the design (about a 4-year period) and construction (about a 5-year period) of the pilot plant which was operated for about 3 years. The final NEDO plant has many features of the Exxon EDS process. During the past 10 years, Japanese officials have been working with other Far Eastern countries to move to the commercial stage but have not accomplished this to date. A yield of about 50% light and medium oils was obtained, but these were of poor quality and would need significant upgrading.

In the 1980s, the Japanese and Australians joined to operate a 50-t/day pilot plant for direct brown coal liquefaction in Australia. The coal was from Latrobe Valley in

Victoria and had low-ash content, but the moisture content was in the 55–65% range. Thus, the first step was a dewatering of the slurry followed by liquefaction using a "highly active" iron catalyst. An in-line hydrogenation process produced high-quality distillable oils. Solvent deashing separated ash from heavy oils, with the oils being recycled and the sludge removed for use as a low-grade boiler fuel. The oil yield was 50%, based on daf coal, and subsequent studies using a 0.1-t/day bench-scale plant led to increases in the oil yield as well as its quality and to reduce the amount of catalyst that was needed.

China, with the world's third largest coal reserves but little petroleum, began a massive planning effort for coal liquefaction with many projects in both direct and indirect liquefaction. In 2008, they terminated all of the major plants except for one for direct and one for indirect liquefaction.

The Shenhua Company is the leading Chinese coal producer. They investigated many of the western coal liquefaction operations but decided to develop a Chinese process. Based on results from smaller units, they built a 6-t/day pilot plant in Shanghai, and it has operated since 2004. One train of 20,000–25,000 bbl/day of a commercial plant was constructed and operated in 2009. The unit has two reactors for the coal dissolution and substantial downstream refining and upgrading to finished products. Runs of up to 100 days duration were accomplished with up to 70% of the design capacity being attained [123]. Carbon dioxide sequestration was included in the trial runs. Following the trial runs with the world's largest DCL plant, modifications are being made with start-up expected to occur during 2011.

Direct Coal Liquefaction and the Future

During the 1970s, DCL was the dominant CTL approach. Today, the restriction on the amount of aromatics in gasoline, limitations on sulfur in the fuel, and the difficulty of producing high-quality diesel fuels have almost eliminated the direct pathway from consideration. Only in Japan and China has the direct process continued to be developed, and it appears that the Japanese have downplayed this approach since the 2009 NEDO report does not emphasize their work on the direct CTL approach. However, one should remember that interest in these processes goes through cycles and conditions may soon change so that interest in direct CTL processes is revived [124–126].

Indirect Coal Liquefaction

For indirect coal liquefaction, a synthesis gas (syngas) composed primarily of carbon monoxide and hydrogen is produced in a suitable gasifier by reaction of coal with oxygen and steam. Prior to the synthesis reactor, the syngas is cooled and cleaned, which involves removal of particulates, H_2S, NH_3, and other impurities. Depending upon the catalyst being used, it may be necessary to adjust the H_2/CO ratio of the syngas. The Lurgi gasifiers used at Sasol in South Africa and Great Plains in the United States produce a syngas with a H_2/CO ratio of two or greater and can be employed without water–gas shift to produce the feed for the Fischer–Tropsch or methane synthesis. Most modern coal gasifiers typically produce syngas with a H_2/CO ratio in the range 0.45–1.0, and this must be adjusted upward to about 2.1, the usage ratio to produce the hydrocarbon products. Iron catalysts can be utilized with any syngas below the usage ratio, but a cobalt catalyst does not contain water–gas shift activity, and the feed should be equal to the usage ratio.

Any of the gasifiers mentioned earlier can be used to produce the syngas, and the one used will only impact the subprocesses that need to be placed between the gasifier and the syngas conversion reactor. The feature that defines the indirect coal liquefaction is the technology used to convert the syngas to the syncrude. Today, the Fischer–Tropsch synthesis is the dominant approach to accomplish this synthesis, but the synthesis of methanol, dimethyl ether, etc., can also be used.

Fischer–Tropsch

The Fischer–Tropsch (FT) synthesis was developed in 1925 by the Germans Franz Fischer and Hans Tropsch. The initial commercial processes in Germany were based on cobalt catalysts and utilized fixed-bed reactors. These reactors had a production rate of 15 bbl/day so that a plant would consist of the number of these small reactors that were needed to attain the design capacity of FT products. Because of the exothermosity of the reaction, various fixed-bed reactor designs were tried, but the most widely utilized design consisted of two concentric tubes with the catalyst located between the two tubes and the cooling water/steam located inside the inner tube and outside the outer tube. These reactors were fitted with cooling fins and bundled within a shell in a number needed to produce 15 bbl/day.

Soon after the end of the war, commercial operations were undertaken in the United States and South Africa. The US effort was led by HRI and utilized a fixed-fluid-bed reactor. The US effort was based on cheap natural gas, and the effort was terminated when the price of natural gas rose significantly. In addition to many problems with oxygen production, syngas generation, and syngas cleaning units, the scale-up of the fixed-fluid-bed reactor encountered many problems. The scale-up directly from a 2 to 7 in. laboratory to a commercial-scale reactor caused them to not recognize the formation of large gas bubbles and the much lower conversion that results from this. A circulating bed (HTFT), rather than a fixed-bed, reactor was used in South Africa, and problems were encountered during the initial reaction period. However, government support allowed continued development to overcome the operating problems. Sasol continued operating the HTFT reactors into the 1990s when they introduced a fixed-fluid-bed Sasol Advanced Synthol, SAS, reactor. The major advantages of the SAS reactor led to a rapid switch all of the HTFT reactors to SAS reactors. SAS reactors of 20,000 bbl/day have been placed in operation. In order to maintain the fluid bed, the FT products must remain in the gas phase at the reaction temperature. This requires the SAS reactors to limit the carbon number of the products to about 20–25; to accomplish this, the reactors are operated at high temperatures (300–350°C) and a suitable catalyst formulation. Because the cobalt catalyst produced predominantly methane at these high reaction temperatures, the fluid-bed operations are limited to an iron-based catalyst.

In the 1990s, Sasol brought on-line a 2,500-bbl/day slurry reactor wherein the small solid catalyst particles are suspended in the reactor wax products that are produced. Work using a similar reactor had been conducted in Germany during the 1940–1960 period that utilized an iron catalyst, but this work did not progress beyond the pilot-plant stage. Even so, the German work established much of the theoretical basis for the utilization of a slurry reactor for the Fischer–Tropsch synthesis.

For the slurry reactor, the catalyst particles are suspended in the high-boiling Fischer–Tropsch products; the catalyst slurry enters at the bottom of the reactor and flows upward. The products that are in the gas phase at the reaction temperature exit the reactor with the unconverted syngas through the top of the reactor. The nonvolatile products are removed from the reactor using an internal filter or in an external settling tank; wax removal is one of the challenging aspects of this reactor type. The reactor can be operated essentially in an isothermal mode, and this ensures that all catalyst particles contribute equally to the generation of products. Catalyst addition and withdrawal can be practiced so that a constant productivity can be maintained.

In South Africa, the fixed-bed reactor was also utilized. In the beginning, it was the same as those used in Germany except that the catalyst was iron based rather than cobalt based. These reactors were operated under low-temperature conditions so that waxes were a major product. Normally, these waxes would need to be cracked or hydrocracked to produce transportation fuels. However, the demand for waxes is sufficient for Sasol to process them for a variety of uses, including candles, chewing gum, fruit and vegetable coatings, etc.

The fixed-bed reactor provides for easy separation of the products from the catalyst. Special operating procedures are needed to minimize the uneven temperature along the length of the reactor due to the highly exothermic nature of the reaction. For both reactor types, operating temperature increase is used to offset the gradual deactivation of the catalyst.

Both the circulating and fixed-fluid-bed reactors are capable of nearly isothermal operations. Catalyst addition and withdrawal are easily accomplished with both fluid-bed reactors. Catalyst attrition and degradation due to carbon formation are problems that must be dealt with for the fluid-bed reactors.

Sasol has commercial operating experience with the four types of reactors utilized for Fischer–Tropsch synthesis (Fig. 19.34). Sasol operates with both the iron and cobalt catalysts at the commercial scale. Shell has concentrated on the fixed-bed reactor for their commercial operations that utilize the cobalt catalyst.

In the 1990s, Mossgas (now PetroSA) constructed and brought on-line a 40,000-bbl/day plant that utilized the circulating fluid-bed reactors such as those used by Sasol. During the past few years, PetroSA, working with Statoil, brought on-line and operated a 1,000-bbl/day slurry reactor. Today, this effort has produced the GTL.F1 technology, a combined effort of Statoil, PetroSA, and Lurgi. The PetroSA operation is based on natural gas rather than coal.

In a simple view, the FT synthesis can be viewed as a polymerization reaction in which the monomer is derived from, or is, the CO reactant. Thus, in this simplified model, the polymerization produces a product distribution as shown in Fig. 19.35 (left); in actual practice, the distribution that is obtained is better represented in Fig. 19.35 (right) plot.

The basic products of the FT reaction are paraffins, olefins, and alcohols, and the reactions may be represented as:

Paraffins

$$(2n+1)H_2 + nCO = C_nH_{2n+2} + nH_2O$$

Olefins

$$2nH_2 + nCO = C_nH_{2n} + nH_2O$$

Alcohols

$$2nH_2 + nCO = C_nH_{2n+1}OH + nH_2O$$

The following competing reaction, water–gas shift, may occur; in practice, it occurs to a

$$CO + H_2O = H_2 + CO_2$$

The FT reaction requires a syngas that has a H_2/CO ratio of 2.1–2.2 depending upon the fraction of light hydrocarbons that are produced. Since modern gasifiers produce a syngas with a H_2/CO ratio in the range of 0.5–1.2, water–gas shift or

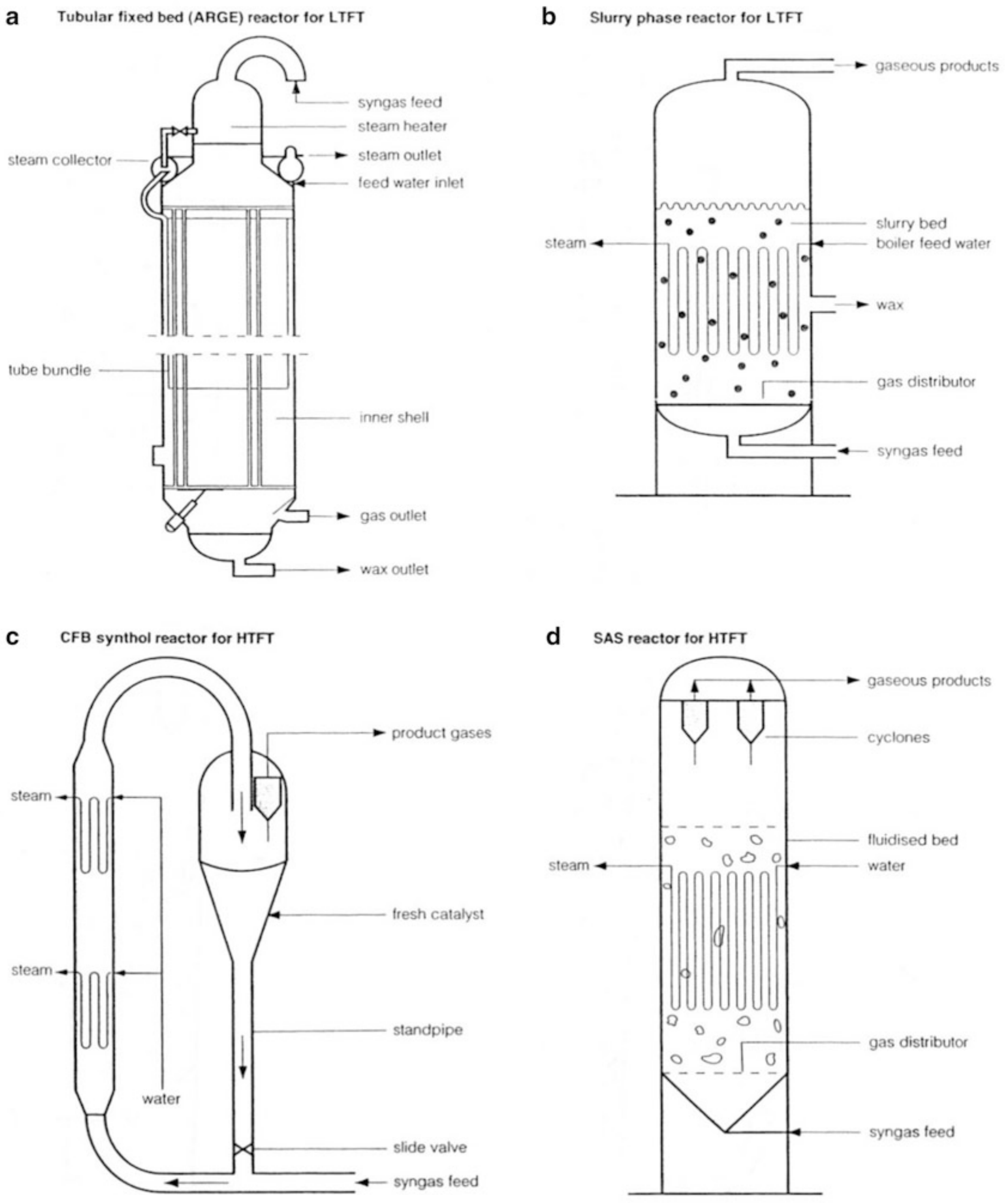


Fig. 19.34 The Sasol FT reactors [147]

another source of hydrogen is required. The near absence of WGS activity for the cobalt catalyst requires an upstream WGS reactor to be utilized with those gasifiers that produce a syngas with a H_2/CO ratio in the 0.5–1.2 range. When an iron catalyst is used, this will not be needed.

The FT synthesis products typically follow an Anderson–Schulz–Flory distribution:

$$W_n = n(1-\alpha)^2\alpha^{n-1}$$

where W_n is the weight fraction of produces with carbon number n, and α is the chain growth probability whose value depends upon the catalyst and operating conditions. When the value of α is zero, only methane is formed, and as the value of α approaches one, the products become predominantly high molecular weight waxes. In logarithmic form, the product distribution becomes

$$\log(W_n/_n) = n\log\alpha + \log[(1-\alpha)^2/\alpha]$$

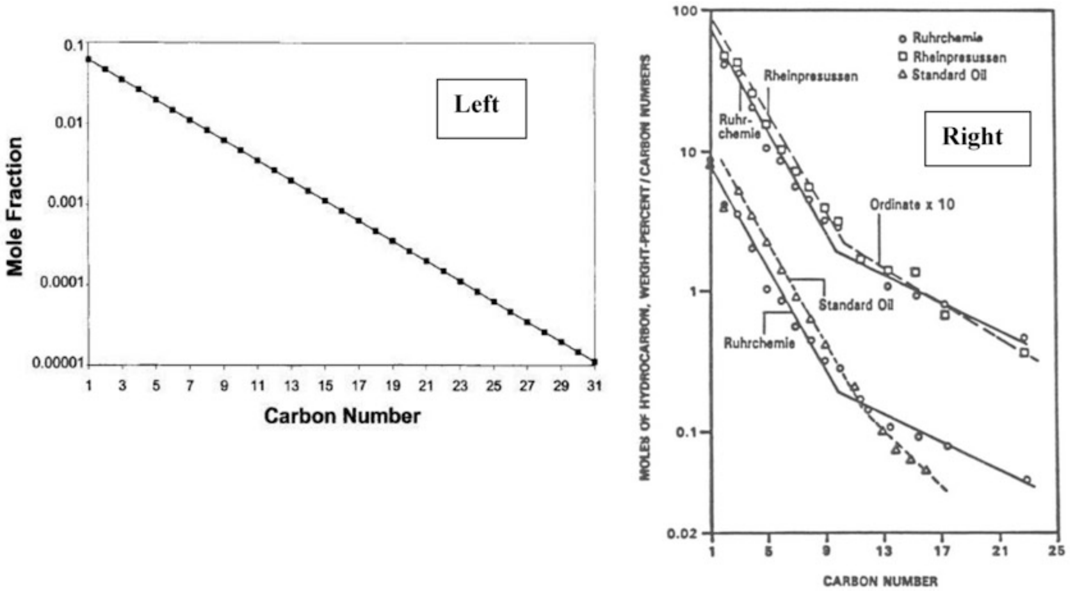


Fig. 19.35 (*Left*) Typical Schulz–Flory plot. (*Right*) Logarithmic plots of moles again carbon number. Hydrocarbons Schwarzheide tests compared with those from fluidized synthesis

A plot of log(W_n/n) vs. n is a straight line with a slope that depends on α (Fig. 19.35 left).

Indirect Coal Liquefaction and the Future

Today, the FT synthesis has replaced the DCL pathway as the preferred route. While Sasol currently is the only commercial coal-based FT operation, it is large and studies are underway that could lead to a doubling of the production from coal (from 150,000 to 300,000 bbl/day). In addition, significant expansion of FT production has occurred, based on natural gas. Thus, Shell/Qatar is expanding by adding 120,000 bbl/day to the current 15,000 bbl/day in Malaysia. Sasol is planning to increase the operation at Sasolburg by 12,000 bbl/day and is at various stages of planning for plants to be located in China, India, and Australia. Thus, it appears that there will be significant growth during the coming years.

Stranded gas and the need to reduce the carbon footprint of fossil fuels offer promise of the need for smaller scale plants. One potential advance that could speed developments in these areas is the introduction of microchannel reactors. In addition to the ability to construct much smaller reactors (a meter length rather than 100), Velocys reports up to at least a tenfold increase in space–time yield using these reactors [127–129].

The technology is practiced in two temperature regions: a high- and low-temperature synthesis. An iron catalyst is used for the high-temperature operation where only a small fraction of the products are higher boiling than the diesel fraction so that downstream processing is essentially limited to a distillation step. Either iron or cobalt can be used for the low-temperature process. An iron catalyst was utilized for the 1950s German slurry reactor investigations and the early Sasol work for their low-temperature slurry reactor. However, today, most of the emphasis is on the cobalt catalyst for both the fixed-bed and the slurry reactors.

In any event, the situation is in a state of flux. A listing was presented for proposed CTL plants in the United States in 2007; an examination of the list at the beginning of 2011 shows that nearly all of these plants have either been canceled or put on hold. The US DOE has caused a baseline technical and economic assessment for FT to be made [130] as well as for coal and biomass for producing FT fuels and electricity [131].

Methanol. Methanol is a water-soluble, low molecular weight alcohol that may be of increasing importance as a low-sulfur fuel, a chemical feedstock, and perhaps as a gasoline additive or an intermediate in gasoline production. The synthesis of methanol is accomplished by the catalytic conversion of synthesis gas containing two moles of hydrogen for each mole of carbon monoxide. Methanol synthesis is widely practiced in industry on a commercial scale. See Chap. 10 for a discussion of methanol manufacture.

Methanol to Gasoline: The Mobil Process. Mobil Research and Development Corporation has developed a process that catalytically dehydrates and polymerizes methanol to produce a high-octane gasoline. The catalyst is one of a family of synthetic zeolites designated ZSM-5 by Mobil. These zeolites

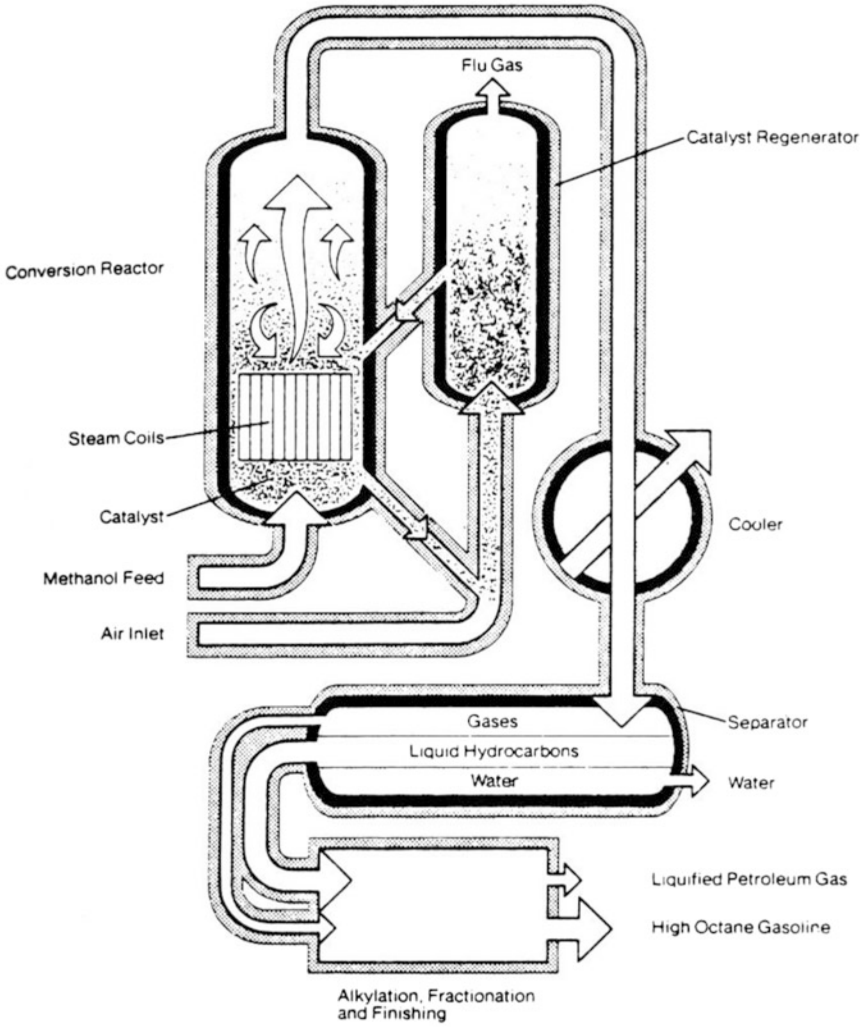


Fig. 19.36 Schematic of Mobil's fluid-bed MTG process, which uses a unique zeolite catalyst to convert methanol to high-octane, unleaded gasoline (Courtesy of The Pace Company, Denver, CO)

have a unique channel structure, and one of these new zeolites converts methanol into a mixture of hydrocarbons corresponding to high-quality gasoline. The mass balance shows 0.44 lb of gasoline and 0.56 lb of water from each pound of methanol; however, the gasoline produced contains 95% of the heating value in the methanol feed.

A schematic diagram of one version of the process is shown in Fig. 19.36. Crude methanol is vaporized at the bottom of the reactor and passes through a dense fluidized bed at 410–415°C and 25 psig. The methanol is converted to hydrocarbons and water. The catalyst is separated from reaction products in a disengager section at the top of the reactor, the reactor effluent is condensed, and the water and hydrocarbon products are separated. To make additional gasoline, propene and butenes can be alkylated with isobutane by conventional petroleum technology (see Chap. 18).

Portions of the powdered catalyst are removed from the reactor, regenerated with air, and returned to the reactor. Small amounts of carbon monoxide, carbon dioxide, and coke are formed as by-products. The process and potential opportunities for the use of MTG to produce fuels from coal have been discussed recently (2008) [132].

The process was operated at the 15,000-bbl/day scale in New Zealand in the 1980s. Multiple fixed-bed reactors were used in the New Zealand operation. The first reactor was used to convert methanol to an equilibrium mixture of methanol, dimethyl ether, and water. This equilibrium mixture is mixed with recycle gas and passed over the ZSM-5 catalyst in a fixed-bed reactor. The first reactor to produce the alcohol, ether, and water mixture releases about 20% of the overall heat of reaction. ExxonMobil claims to have made advances since then that reduce both capital investment and operation expense [133].

Methanol can also be converted to olefins and produce essentially only ethene and propene. While the ZSM-5 catalyst can be adapted for this purpose, usually a less acidic catalyst is utilized, frequently a silicoaluminophosphate (e.g., SAPO-34) [134]. A process (UOP/HYDRO MTO) is

Table 19.20 Crude oil price (Brent) for a profitable DCL project [136]

Market	Crude oil net back price if no carbon capture	Crude oil net back price at CO_2 = 40 \$/tons
China	50–60 \$/bbl	70–80 \$/bbl
United States	75–85 \$/bbl	95–105 \$/bbl
Europe	90–100 \$/bbl	110–120 \$/bbl

described that claims up to 80% ethene/propene and greater than 95% conversion to C_2–C_4 alkenes [135]. The UOP process is at or near the commercial scale.

CTL Summary

The economics for CTL are complex and are very site dependent. This is illustrated by the prices for three locations shown in Table 19.20 for DCL [136]. The outlook should be similar for indirect coal liquefaction. For indirect CTL, the avoidance of CO_2 is relatively low since CO_2 is readily available from the syngas purification step. Based on the cost shown in Table 19.20, projects should be profitable in China and the United States but not in Europe; however, the price in 2011 is not adequate to lead to the initiation of major activities in either country.

The situation with CTL is very unstable. Typical is the case in China. Five years ago, there were many direct and indirect processes at various stages of development. Then, in 2008, the National Development and Reform Commission decided that, apart from two projects operated by the Shenhua Group, none could go ahead before receiving official approval. This was because CTL is "a technology-, talent- and capital-intensive project at an experimental stage with high business risks" [138]. Since Sasol was making significant profits for their CTL operations in South Africa and was involved in a CTL project in China, this decision must have been based on conditions specific to China.

Military needs have provided impetus but little funding for coal to liquids development [138, 139]. The current outlook for CTL has been presented by several groups in recent years [140–147].

Petrochemical Feedstocks

Feedstocks for the petrochemical industry are produced mainly from crude oil (Chap. 18) and natural gas (Chap. 20). About 90% of all organic chemicals are manufactured from just six feedstocks: synthesis gas, ethylene, propylene, butadiene, benzene, and p-xylene, with synthesis gas accounting for over half of these feedstocks. Although a lot of syngas for the manufacture of ammonia, methanol, and other chemicals is produced by steam reforming of natural gas or light oils, as discussed above, syngas can also be produced by coal gasification.

Several sources discuss and evaluate processes for the production of chemicals from coal [148–155]. Most basic chemicals currently produced from oil or natural gas can be produced from coal using demonstrated technology. The broadest range of synthesized products is available by conversion of coal to synthesis gas and utilization of existing processes for production of chemicals, such as ammonia, methanol, and acetyl derivatives. Advances in the chemistry of synthesis gas are rapidly expanding the array of chemicals obtainable from coal by this route [156–162]. Many other chemicals can be separated from the by-products of gasification and pyrolysis processes. In fact, the foundation of the modern chemical industry was built on processes utilizing coal hydrogenation and pyrolysis to obtain a variety of chemicals, including dyes, solvents, fuels, and pharmaceuticals [163]. However, in the 1950s, the wide availability and low cost of crude oil made petroleum-derived materials the preferred source for chemicals production.

Presently (2011), the choice of feedstock for chemicals production depends on complex technical, economic, environmental, and political factors. Clearly, not all chemicals are suitable for production from coal with current technology. Some factors to be considered in the evaluation of the appropriate feedstock for a particular chemical product are (1) the relationship between the carbon/hydrogen ratio in the chemical product and the feedstock, (2) the delivered cost of alternative raw materials, (3) capital costs, (4) environmental protection, and (5) the reliability of supply. Recently, except for special situations, such as that for Sasol in South Africa, the manufacture of chemicals from coal at coal prices relative to the prices of petroleum and natural gas has not been economically attractive.

Chemicals from Coal

The use of coal as a raw material for the production of chemicals started with the discovery of the first coal tar dye by Sir William Perkins in 1856. Prior to that time, the liquid by-products of coking were regarded as wastes. Perkins' discovery revealed that valuable materials could be isolated or made from coal tar and ultimately led to the establishment of the modern chemical industry. Coke ovens rapidly came to be important as sources of chemicals, such as benzene, naphthalene, ammonia, and hydrogen. Coal tar became an important source for aromatic and heterocyclic chemicals for the dye and pharmaceutical industries.

Coke production was formerly the most important demonstrated technology associated with the direct production of chemicals from coal. Industrial chemicals currently obtained in significant amounts as coke by-products include benzene, toluene, xylene, naphthalene, anthracene, phenanthrene, phenol, ammonia, ammonium sulfate, sulfur, and carbon dioxide. The vast majority of aromatics production from coal occurs in Eastern Europe, India, China, and Japan [164].

Chemical by-products from coke manufacture amount to roughly 4% of the coal feed. Although this yield initially supplied adequate quantities of chemicals for developing end uses, the growth of the industry eventually required additional sources of chemicals. The advent of technology for producing calcium carbide and acetylene in 1892 provided a means for conversion of coke to chemical products. Also, this technology stimulated the growth of an independent chemical industry by allowing the manufacture of chemicals to be pursued apart from the manufacture of coke for steel production. More recently, independent production of chemicals from coal has gone through the synthesis gas route.

Chemicals from coal via gasification of coal to synthesis gas as a stand-alone chemical process have been difficult to justify economically because the alternative routes through crude oil or natural gas have been relatively cheap, but that has recently changed. With special circumstances such as large-scale, high final product value, low coal cost, or intangibles such as security of raw material supply, these types of plants have been successfully built and operated on a limited basis. However, as crude oil and natural gas pricing has increased, interest in coal to chemicals has dramatically increased.

Economics of coal gasification plants can often be improved by coproduction of "secondary" chemicals that may be low volume, but high value. This concept of coproduction to improve overall economics is illustrated by the efforts at Dakota Gasification's Great Plains Synfuels plant in North Dakota and at Sasol in South Africa. At Great Plains, the primary product is synthetic natural gas, but other high-value products include naphtha, phenol, fertilizers, krypton–xenon from the air separation plant, and CO_2 exported for enhanced oil recovery. Research and development efforts are ongoing to enable profitable recovery of other by-products. The types of by-products produced depend, in part, on the type of gasification technology employed. Low-temperature gasification such as the Lurgi gasifier used at Dakota Gasification produces more tars and oils that can be separated and refined to other products. Sasol has utilized the tars from gasification and the Fischer–Tropsch products to become a chemical company. At times, chemicals have produced more than half their profits even though chemicals account for less than one-fourth of their production. The high-temperature gasification processes such as GE (formerly Texaco) produce very few chemicals directly other than carbon monoxide, hydrogen, and carbon dioxide.

There is also much interest in coal gasification from power producers to produce electricity from coal more cleanly and efficiently than traditional pulverized-coal steam plants. This concept is commonly referred to as IGCC (integrated gasification combined cycle). Economic justification of stand-alone power plants via IGCC is also difficult due to the high capital cost of these types of plants. There could be significant synergies between power production and chemical coproduction that could improve the economics of both. Power plants bring the large scale necessary for unit cost reductions, but the electricity price in many markets is low and variable (especially in off-peak hours). Chemicals bring a steadier, higher-value demand.

In recent years, several commercial plants have been constructed for conversion of coal to synthesis gas for chemical manufacturing. These include the Eastman Chemical's acetic anhydride plant, the Ube (Japan) ammonia plant, the SAR (Germany) oxo chemicals plant, and several coal to ammonia plants in China (e.g., Weihe, Huainan, and Lunan). The Ube plant and the SAR plant have since converted to lower-cost opportunity fuels (petroleum coke and residues). The Eastman plant is still operating exclusively on coal. Feedstock changes at the other plants illustrate the vulnerability of coal-conversion processes to a changing economic climate. The fact that the Eastman process remains competitive under changing conditions is due to a set of special circumstances that favor a coal-based process. The success of the Eastman chemicals from coal complex demonstrates that synthesis gas from coal is a viable feedstock for some industrial chemicals under certain conditions.

Clearly, the number of chemicals produced from coal-derived synthesis gas can expand as new technologies are developed and favorable economics exist. The most likely such chemicals are those for which processes have been demonstrated but which presently are uneconomic. Relatively small improvements in technology, shifts in feedstock availability and/or cost, decreased capital costs, or political factors could enhance the viability of coal-based processes for the production of methanol, ethanol, and higher alcohols, vinyl acetate, ethylene glycol, carboxylic acids, and light olefins.

New technologies are being developed for the partial oxidation of coal to organic acids, for plasma pyrolysis of coal to produce acetylene and other small molecules [165], and for the manufacture of additional chemicals such as acetaldehyde from synthesis gas [166]. Also, the use of coal tar for the manufacture of high-temperature engineering plastics and bioconversion of coal to aromatics, organic acids, alcohols, and methane [164] have been reported. Innovative strategies are being considered that offer the potential for reduced costs; for example, by co-gasifying coal with waste, waste disposal credits might be obtained.

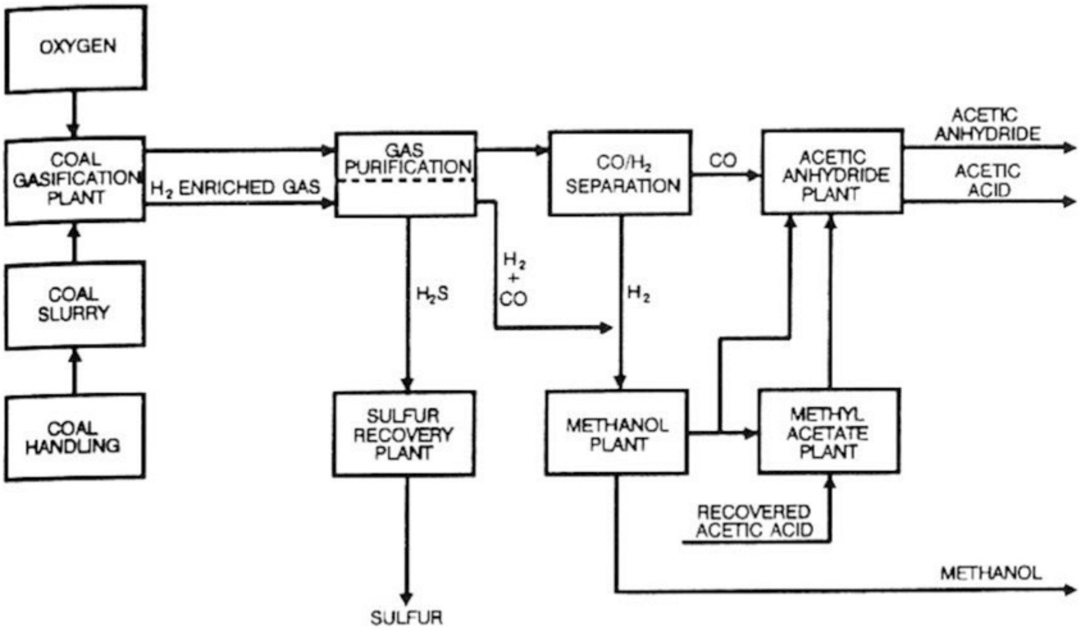


Fig. 19.37 Overall block flow diagram for coal gasification–acetic anhydride complex

Examples of Chemicals Production from Coal

A pioneer US manufacturer to produce a slate of industrial chemicals from coal is Eastman Chemical Company. Commercial facilities include a coal gasification plant for synthesis gas manufacture, raw gas cleanup and separation facilities, a sulfur recovery unit, and chemical plants to produce methanol, methyl acetate, acetic acid, and acetic anhydride. A flow diagram of Eastman's chemicals from coal plant is shown in Fig. 19.37.

A GE quench coal gasifier provides synthesis gas at elevated pressure from local coal. Medium- and low-pressure steams for use elsewhere in the complex are produced by recovering waste heat. After the gasifier product is scrubbed with water to cool the gas and remove ash particles, a portion of the syngas is sent to a water–gas shift reactor to increase its hydrogen content. A Rectisol unit, using a cold methanol wash, then removes hydrogen sulfide and carbon dioxide from the product gas streams. Cryogenic separation in a Linde "cold box" provides a carbon monoxide stream for the acetic anhydride plant and a hydrogen-rich stream for methanol production. The recovered hydrogen sulfide is converted to elemental sulfur in a Claus unit followed by a SCOT tail gas treating unit.

Methanol is produced from carbon monoxide, carbon dioxide, and hydrogen. The proper feed composition is achieved by combining the hydrogen-enriched syngas from the shift reactor and the hydrogen–carbon monoxide stream from the gas separation unit. The feed stream goes to both an energy-efficient, low-pressure, gas-phase, catalytic Lurgi methanol process and an Air Products liquid-phase process. Methanol is reacted with acetic acid to form methyl acetate. In the final step of the process, purified carbon monoxide from the gas separation plant is reacted with methyl acetate to form acetic anhydride. Part of the acetic anhydride is reacted with methanol to coproduce acetic acid. The methyl acetate resulting from this reaction is recarbonylated in the acetic anhydride reactor. The acetic anhydride is used to make various cellulose acetate plastics and fibers and acetate solvents. The reaction produces by-product acetic acid that is recycled to the methyl acetate plant to react with methanol.

Another example of large-scale production of chemicals from coal is the Sasol facility in South Africa. Sasol produces many products from coal-derived syngas, including ethylene, propylene, α-olefins, alcohols, and ketones. They have also increased their production of methanol, synthetic lubricants, detergent alcohols, acrylic acid and acrylates, oxo alcohols, styrene and polystyrene, propylene oxide, and propylene glycol.

Today, the Shenhua Broton Coal Chemical project is a leader in restoring coal as a feedstock for chemicals production, and it will be one of the first new world-scale coals to chemicals project. The methanol plant will be one of the world's largest plants, and the methanol that is produced will be a feedstock for a variety of large-scale chemicals production. Other Chinese plants are converting coal to simple chemicals, and their expansion in this area is very rapid [167–171].

As the above examples illustrate, the potential exists for producing larger quantities of commodity chemicals from coal. Some factors that could facilitate the increased use of coal for this purpose include price increases for petroleum and increased use of coal gasifiers for power production. As IGCC plants are built, there will be an opportunity to incorporate in the plant design the coproduction of high-value chemicals.

Acknowledgment The authors are indebted to the authors of the chapters in previous editions. Without the use and guidance that these afforded us, this section would have been much more difficult.

References

1. Sýkorová I, Pickel W, Christanis K, Wolf M, Taylor GH, Flores D (2005) Classification of huminite – ICCP System 1994 Int J Coal Geol 62:85–106
2. International Committee for Coal and Organic Petrology (1998) The new vitrinite classification (ICCP System 1994). Fuel 77:349–358
3. International Committee for Coal and Organic Petrology (1994) The new inertinite classification (ICCP System 1994). Fuel 80:459–471
4. Dryden IGC (1963) Chemical constitution and reactions of coal. In: Lowry HH (ed) Chemistry of coal utilization, supplementary volume. Wiley, New York, pp 223–295
5. Hatcher PG, Clifford DJ (1996) The organic geochemistry of coal: from plant materials to coal. Org Geochem 27:251–274
6. Wender I, Heredy LA, Neuworth MB, Dryden IGC (1981) Chemical reactions and the constitution of coal. In: Elliott MA (ed) Chemistry of coal utilization, second supplementary volume. Wiley, New York, pp 425–522
7. Gavalas GR (1982) Coal pyrolysis. In: Anderson LL (ed) Coal science and technology, vol 4. Elsevier, Amsterdam
8. Lin YA (ed) (1982) Physical cleaning of coal—present and developing methods. Marcel Dekker, New York
9. Leonard JW (ed) (1992) Coal preparation, 5th edn. American Institute of Mining, Metallurgical and Petroleum Engineers, Inc., New York
10. Hower JC (1998) Interrelationship of coal grinding properties and coal petrology. Miner Metall Process 15(3):1–16
11. Osborne DG (1989) Coal preparation technology. Graham & Trotman, London
12. Gala HB, Srivastava RD, Rhee KH, Hucko RE (1989) An overview of the chemistry of molten-caustic-leaching process. Coal Prep 7:1–28
13. Sakkestad BA (ed) (1998) The proceedings of the 23rd international technical conference on coal utilization & fuel systems. Coal & Slurry Technology Association, Washington, DC
14. http://www.our-energy.com/coal.html. Accessed 4 Feb 2006
15. Stultz SC, Kitto JB (eds) (1992) Steam—its generation and use. The Babcock & Wilcox Company, Barberton, OH
16. Singer JG (ed) (1991) Combustion fossil power—a reference book on burning and steam generation. Combustion Engineering, Inc., Windsor, CT
17. Bartok W, Sarofim AF (eds) (1991) Fossil fuel combustion—a source book. Wiley, New York
18. Bartok W, Lyon RK, McIntyre AD, Ruth LA, Sommerlad RE (1988) Combustors: applications and design considerations. Chem Eng Progr 84(3):54
19. Ceely FJ, Daman EL (1981) Combustion process technology. In: Elliott MA (ed) Chemistry of coal utilization. Wiley-Interscience, New York
20. McCarthy K, Laux S, Graska J, Rosin T, Hausman GL (1999) Improved low NO_x firing systems for pulverized coal combustion. In: 24th international technical conference on coal utilization and fuel systems, Clearwater, FL, 8 March 1999
21. Department of Energy, Clean Coal Technology compendium. www.lanl.gov/projects/cctc
22. Shuster E (2011) Tracking new coal-fired power plants. NETL, Pittsburgh, PA
23. www.sourcewatch.org/index.php?title=Existing_U.S._Coal_Plants
24. Characterization of the U.S. Industrial Commercial Boiler Population, submitted by Energy and Environmental Analysis, Inc. to ORNL, May 2005
25. Geisbrecht R, Dipietro P (2009) Evaluating options for U.S. coal fired power plants in the face of uncertainties and greenhouse gas caps: the economics of refurbishing, retrofitting, and repowering. Energy Procedia 1:4347–4354
26. Schilling H-D (2006) Efficiency of coal-fired-power stations—evolution and prospects. EurActiv (http://www.euractive.com). Accessed 25 April 2006
27. Roberts BF, Goudarzi L. www.econsci.com/euar9801.html
28. Corcoran E (1991) Cleaning up coal. Sci Am 264(5):106
29. Haggin J (1991) Coal. Chem Eng. News 69:32–36
30. Environmental Protection Agency (1997) Mercury study report to congress volume VII—an evaluation of mercury control technologies and costs. EPA-452/R-97-0 I 0, Dec 1997
31. Brown T, Smith D, Hargis R, O'Dowd W (1999) Mercury measurement and its control: what we know, have learned, and need to further investigate. J Air Waste Manag Assoc 49:1–97
32. Berlin K, Sussman RM (2007) Global warming and the future of coal: the path to carbon capture and storage, May 2007. www.americanprogress.org
33. Interdisciplinary MIT Study (2007) Geological Carbon Sequestration in The Future of Coal: Options for a carbon-constrained world, 43–62 ISBN 13: 978888-0-615-14092-6
34. Beaty HW, Mahrous H, Seidman AH (2001) Handbook of electric power calculations, 3rd edn. McGraw-Hill, New York
35. Schobert HH (1987) Coal: the energy source of the past and the future. American Chemical Society, Washington, DC
36. Interdisciplinary MIT Study (2007) Coal-based Electricity Generation in The Future of Coal: Options for a carbon-constrained world, 17-42 ISBN 13: 978888-0-615-14092-6
37. Wall T, Liu Y (2008) Chemical looping combustion and CO_2 capture: status and developments, Technology Note 32, The University of Newcastle, Feb 2008
38. He F, Li H, Zhao Z (2009) Advancements in development of chemical-looping combustion: a review. Int. J Chem. Eng. Article ID 710515, 16 pages doi:10.1155/2009/710515
39. Independent assessment of the potential of the OSU chemical looping concept in a coal-to-liquids configuration, DOE/NETL-2008/1307, Dec 2007
40. Cost and performance baseline for fossil energy plants, volume 1. Bituminous coal and natural gas to electricity, DOE/2010/1397 final report, Nov 2010
41. Current and future technologies for gasification-based power generation, volume 2. A pathway study focused on carbon capture advanced power systems R&D using bituminous coal, DOE/NETL-2009/1389, Revision 1, Nov 2010
42. Wolfe RA (2007) Mild gasification process for the production of smokeless coal, coke and coal liquids. In: Proceedings of the annual international Pittsburgh coal conference, 24th, 344/1-344/25, Johannesburg, S. Africa, September 10-14, 2007

43. Wolfe RA, Henkelman MR, Davis CS, Wolfe EA (1995) Production of high-quality coke by continuous mild gasification. Trans Am Foundrymen's Soc 103:777–782
44. Shamsi A, Shadle LJ, Seshadri KS (2004) Study of low-temperature oxidation of buckskin subbituminous coal and derived chars produced in ENCOAL process. Fuel Proc Tech 86(3):275–292
45. Sondreal EA, Ness RO, Timpe RC, Knudson CL, Jha MC, Berggren MH, Sinor JE (1989) The EMRC-AMAX mild gasification process for midwestern bituminous and western low-rank coals. Proceedings – annual Intl. Pittsburgh Coal Conf 6(2):785–794, vol 2, 25–29 Sept 1989, pp 785–794
46. Higman C, van der Burgt M (2003) Gasification. Elsevier, Amsterdam
47. Synthetic Fuels Association, Inc. (1983) Coal gasification systems: a guide to status, applications and economics. EPRI AP-3109, Project 2207, June 1983
48. SFA Pacific, Inc. (1993) Coal gasification guidebook: status, applications, and technologies. Research project 2221-39, prepared for EPRI, Final report, Dec 1993
49. Walker PL Jr (1986) Carbon gasification ubiquitous reaction of great importance to mankind, NATO ASI sires, series E. Appl Sci 105:3–23
50. Breault RW (2010) Gasification processes old and new: a basic review of the major technologies. Energies 3:216–240
51. Shinner R (1987) Fourth technical meeting of COGARN. In: Penner SS et al (eds) Coal gasification: direct applications and syntheses of chemicals and fuels; A research needs assessment, presented for US DOE, Office of Energy Research, DOE Contract No. DE-ACO I -85ER30076, NTIS-PR-360
52. Ludvig MM, Gard GL, Emmett PH (1983) Use of controlled oxidation to increase the surface area of coal. Application to bituminous and semianthracite coal. Fuel 62:1393–1396
53. Dry ME, Steynberg AP (eds) (2004) Studies in surface science and catalysis, vol 152. Fischer-Tropsch Technology, Elsevier, Amsterdam
54. Newman SA (ed) (1985) Acid and sour gas treating processes. Gulf, Houston
55. Spencer DE, Alpert SB, Gluckman MJ (1985) Integrated coal gasification combined cycles (IGCC) an emerging commercial option for the power industry. In: Paper presented at 1985 ACS meeting, CA
56. Spencer DE (1986) The commercial implications of the cool water project for the electric power industry. EPRI, Palo Alto, CA
57. Zhuang QZ (2006) Commercial experience in China…GE's Gasification Technologies Conference, Washington, DC, October 3, 2006
58. Amick P (2010) E-Gas^{TM} technology, 2010 outlook. In: Gasification technologies conference, Washington, DC, 1 Nov 2010
59. www.netl.doe.gov/technologies/coalpower/gasification/gasifipedia/4-gasifiers/4—1-2-5.mhi.html
60. Jager H (2008) Compact gasifier 90% smaller and half the cost of conventional units. GTW 2008 IGCC Reference Guide 28-32
61. Klimenko AY (2009) Early ideas in underground coal gasification and their evolution. Energies 2:456–476
62. Couch G. Underground coal gasification, Coal Online, http://www.coalonline.info/site/coalonline/content/browser/82189/Underground-coal-gasification
63. Ricketts B, Hotchkiss R, Livingston B, Hall M (2002) Technology status review of waste/biomass co-gasification with coal. In: IChemE fifth European gasification conference, Noordwijk, The Netherlands, 8–10 April 2002
64. Bain RL (2009) Status of U.S. biomass gasification; Task meeting. IEA Bioenergy, Task 33, NREL, 14 May 2009
65. McDaniel J (2002) Biomass co-gasification at Polk Power Station, final technical report, May 2002. http://www.osti.gov/bridge/servlets/purl/823831-bTDn1G/native/823831.PDF
66. Affordable, low-carbon diesel fuel from domestic coal and biomass, NETL, DOE/NETL-209/1349, 14 Jan 2009
67. Tennant JB (2010) Overview of DOE's gasification program, NETL, Version 1.1, Dec 2010
68. Childress J (2010) World gasification industry: status, trends & drivers. In: Gasification workshop, Terre Haute, Indiana, 30 June to 1 July 2010
69. Abadie LM, Chamorro JM (2009) The economics of gasification: a market-based approach. Energies 2:662–694
70. SRI Consulting Process Economics Program Reports, 154 (Jan 1983), 154A (Dec 2006) and 154B (Dec 2007)
71. Khan MR, Kurata T (1985) The feasibility of mild gasification of coal: research needs. DOEIMETC-85/4019, NTRS/DE850 13625
72. Probstein RE, Hicks RE (1982) Synthetic fuels. McGraw-Hill, New York
73. Seglin L, Bresler SA (1981) Low-temperature pyrolysis technology. In: Elliott MA (ed) Chemistry of coal utilization, second supplementary volume. Wiley, New York, pp 785–846
74. Lee S (2007) Clean liquids from fuels. In: Lee S, Speight JG, Loyalka SK (eds) Handbook of alternative fuel technologies, CRC Press LLC, Boca Raton, FL pp 81–123
75. Miura K (2006) Achievement of the coal devolatilization and char gasification working group under the BRAIN-C program. J Japan Inst Energy 85(1):12–19
76. Kajitami S, Watanabe H, Zhang Y, Hara S, Miura K (2008) Modeling of morphology of char during coal gasification. J Japan Inst Energy 87(1):48–55
77. Howard JB (1981) Fundamentals of coal pyrolysis and hydropyrolysis. In: Elliott MA (ed) Chemistry of coal utilization, second supplementary volume. Wiley, New York, pp 665–784
78. Solomon PR, Serio MA (1987) Evaluation of coal pyrolysis kinetics. In: Lahaye J, Prado G (eds) Fundamentals of physical chemistry of pulverized coal combustion. Martinus Nijhoff Publishers, Leiden
79. Lee S (1996) Alternative fuels (coal pyrolysis for liquid fuels). Taylor & Francis, New York, pp 166–198
80. Höök M, Aleklett K (2010) A review on coal to liquid fuels and its coal consumption. Int J Energy Res (Wiley Interscience) 34:848–864
81. Khan MR (1985) Thermoplastic properties of coal pyrolysis at elevated pressures: effects of experimental variables, inorganic additives, and preoxidation. PhD thesis, The Pennsylvania State University, p 333
82. Anthony DB, Howard JB (1976) Coal devolatilization and hydrogasification. AIChE J 24(2):625–656
83. Solomon PR, Serio MA, Suuberg EM (1992) Coal pyrolysis: experiments, kinetics rates and mechanisms. Prog Energy Combust Sci 18(2):133–220
84. Arthur JR, Wadsworth KD (1950) The effect of inorganic compounds on the behavior of coals and cokes. BCURA Prog Coal Sci 396–410
85. Gryaznov NS, Petrov VK (1960) Coke and chemistry, pp 27–32. USSR (Coal Tar RA Translation)
86. Given PH, Yarzab RE (1978) Analysis of the organic substance of coals: problems posed by the presence of mineral matter. In: Karr C Jr (ed) Analytical methods for coal and coal products, vol II. Academic, New York, pp 3–41
87. Walker PL Jr, Davis A, Verina SK, Rivera-Utrilia L, Khan MR (1984) Interaction of gases, vapors, and liquids with coals and minerals at ambient temperatures-their use to characterize porosity. The Pennsylvania State University, DOE-300 13–19, Under Contract No. DE-AC22-80PC300 13
88. Gray RH, Drucker H, Massey MJ (eds) (1988) Toxicology of coal conversion processing. Wiley, New York
89. Speight JG (2008) Synthetic fuels handbook: properties, process and performance. McGraw-Hill, New York, pp 9–10

90. Neavel RC (1976) Liquefaction of coal in hydrogen-donor and non-donor vehicles. Fuel 55:237–242
91. Derbyshire F, Stansberry P (1987) Comments on the reactivity of low-rank coals in liquefaction. Fuel 66:1741–1742
92. Wu WRK, Storch HH (1968) Hydrogenation of coal tar. U.S. Bureau of Mines Bulletin No. 633
93. Donath EE, Hoering M (1977) Early coal hydrogenation catalysis. Fuel Process Technol J 1:3–20
94. Ministry of Fuel and Power (1947) Report on the petroleum and synthetic oil industry of Germany. Ministry of Fuel and Power, His Majesty's Stationery Office, London
95. Green DW (ed) (1997) Perry's chemical engineers' handbook, 7th edn. McGraw-Hill, New York
96. Derbyshire F (1968) Catalysis in coal liquefaction: new directions for research. IEACR/08, IEA Coal Research, London
97. Keogh RA, Davis BH (2009) Coal liquefaction: the reaction pathway and the politics. Paper presented at the 237th ACs National meeting, Salt Lake City, Utah, 22–26 March 2009
98. Keogh RA, Tsai K, Xu L, Davis BH (1991) Liquefaction pathways of U.S. bituminous coals. Energy Fuel 5:625–632
99. Keogh RA, Davis BH (1994) Comparison of liquefaction pathways of a bituminous and subbituminous coal. Energy Fuel 8:289–293
100. Zhang C, Lee CW, Keogh RA, Demirel B, Davis BH (2001) Thermal and catalytic conversion of asphaltenes. Fuel 80:1131–1146
101. DOE (1989) Coal liquefaction: a research and development needs assessment. DOE Coal Liquefaction Research Needs (COLIRN) panel assessment, Final report, vol 2, OE/ER-0400, March 1989
102. Vick GK, Epperly WR (1982) Status of the development of EDS and liquefaction. Science 217(4557):311–316
103. Neavel RC (1981) Exxon donor solvent liquefaction process. Philos Trans Roy Soc Lond A 300:141–156
104. Tomlinson GC, Gray D, Neuworth MB, Talib A (1985) The impact of rank-related coal properties on the response of coals to continuous direct liquefaction processes. Sandia Contractors Report, SAND85-7238, Oct 1985
105. Maa PS, Mraw SC, Veluswamy LR, Wen MY, Mein GA, Spencer DC (1997) Exxon's Microcat Coal Liquefaction (MCL) process. In 213th ACS national meeting, San Francisco, 13–17 April 1997, PETR-024
106. Bauman R, Zaczepinski S (2011) 2nd Generation microcatalytic coal liquefaction, In: world CTL 2011 conference, Paris, France, 1–3 March 2011
107. Holmes DF Jr, Mitchell J, Farthing S (1983) A final report of Kentucky's role in the H-coal pilot plant project. Kentucky Energy Cabinet, Aug 1983
108. Hecka C (2011) Competences in unconventional energy resources. In: World CTL 2011 conference, Paris, France, 1–3 March 2011
109. Wanzl W (2008) Pyrosol—a process for coal hydrogenation. In: Schmolfeld J (ed) Die Veredhung und Umwandlung von Kohle. DGMH, Hamburg, ISBN 978-3-936418-88-0, pp 687–691
110. Review of worldwide coal to liquids: R,D&D activities and the need for further initiatives within Europe, Contract #RFC2-CT-2008-00006, June 2009
111. Derbyshire F (1986) Coal liquefaction. In: Ullmann's encyclopedia of industrial chemistry, vol A7. VCH Publishers, New York
112. Maa PS, Trachte KL, Williams RD. (1981) Solvent effects in exxon donor solvent coal liquefaction. In: Presented at the ACS national meeting, New York, Aug 23–28
113. Neavel RC (1981) Exxon donor solvent liquefaction process. Philos Trans Roy Soc Lond A 300:141–156
114. Holmes DF Jr, Mitchell L, Farthing S (1983) A final report of Kentucky's role in the H-coal pilot plant project. Kentucky Energy Cabinet, Aug 1983
115. EPRI (1991) Two-stage coal liquefaction integrated configurations: the advanced coal liquefaction R&D facility, Wilsonville, AL. EPRI GS-7293, May 1991
116. Gollakota SV, Lee LM, Davies OL (1989) Process optimization of close-coupled integrated two-stage liquefaction by the use of cleaned coals. Fuel Process Technol 22:205–216
117. Nalitham RV, Lee LM, Lamb CW, Johnson TW (1987) Two-stage coal liquefaction process performance with close-coupled reactors. Fuel Process Technol 17:13–27
118. Rosenthal JW, Dahlberg AL, Kuehler CW, Cash DR, Freedman W (1982) The Chevron coal liquefaction process (CCLP). Fuel 61 (10):1045–1049
119. El Sawy A, Gray D, Talib A, Tomlinson G (1986) A techno-economic assessment of recent advances in direct coal liquefaction. Sandia Contractor report SAND86-71 03, June 1986
120. Derbyshire FJ (1989) Role of catalysis in coal liquefaction research and development. Energy Fuel 3:273–277
121. Burke FP, Brandes SD, McCoy DC, Winschel RA, Gray D, Tomlinson G (2001) Campaign of the late twentieth century: topical report. DOE/PC93054-94, July 2001
122. Shui H, Cao Z, Xu C (2010) Recent advances in direct coal liquefaction. Energies 3:155–170
123. Sun Q (2011) Private communication
124. Robinson KK (2009) Reaction engineering of direct coal liquefaction. Energies 2:976–1006
125. Lepinski J, Lee T, Tam S (2009) Recent advances in direct coal liquefaction technology. In: ACS national meeting, Salt Lake City, Utah, 26 March 2009
126. Winslow J, Schmeta E (2009) Direct coal liquefaction overview. Presented to NETL, 23 March 2009
127. Kilanowski DR (2009) Microchannel Fischer-Tropsch approaches market entry phase, 4th BTLtec. Graz, Austria, 24–25 Sept 2009
128. Tonkovich AL, Mazanec T, Jorosch K, Fitzgerald S, Yang B, Taha R, Kilanowski D, Lerou J, McDonald J, Atkinson D, Dritz T (2010) Gas-to-liquids conversion of associated gas enabled by microchannel technology. Oxford Catalyst Group, Oxford, pp 1–7
129. LeViness S, Tonkovich AL, Jarosch K, Fitzgerald S, Yang B, McDaniel J (2011) Improved Fischer-Tropsch economics enabled by microchannel technology. Velocys 1–7
130. Shuster E, Haslebeck J, Rutkowski M, Olson S, Kramer S (2007) Baseline technical and economic assessment of a commercial scale Fischer-Tropsch liquids facility. DOE/NETL-2007/1260, 9 April 2007
131. Liu G, Larson ED, Williams RH, Kreutz TG, Guo X (2011) Making Fischer-Tropsch fuels and electricity from coal and biomass: performance and cost analysis. Energy Fuel 25:415–437
132. Zhao X, McGhion RD, Tabak SA (2008) Coal to clean gasoline. Hydrocarbon Eng 13(9):39–40, 42, 44, 46, 48–49
133. Tabak SA, Zhao X, Brandl A (2008) An alternative route for coal to liquid fuel: Exxon mobil methanol to gasoline (MTG) process. In: First world coal-to-liquids conference, Paris, 3–4 April 2008
134. Barger PT (2002) Methanol to olefins (MTO) and beyond. Catalytic science series (Zeolites for Cleaner Technologies), vol 3, pp 239–260
135. Barger PT, Vora BV, Pujado PR, Chen Q (2003) Converting natural gas to ethylene and propylene using the UOP/HYDRO MTO process. Stud Surf Sci 145:109–114
136. Appert O (2011) Coal conversion to fluid hydrocarbons: main stakes and issues. World CTL 2011 congress, Paris, France, 1–3 March 2011
137. Qi Z (2008) China Daily, 9 Oct 2008. www.chinadaily.com.cn/bizchina/2008-10/09/content_7090441.htm
138. Bartis JT, Comm F, Ortiz DS (2008) Producing liquid fuels from coal: prospects and policy issues, Rand Corporation, prepared for U.S. Air Force and NETL, U.S. Department of Energy

139. Bartis JT, Van Bibber L (2011) Alternative fuels for military applications. Rand Corporation, Santa Monica
140. Sun Q, Flecher JJ, Zhang Y, Ren X (2005) Comparative analysis of costs of alternative coal liquefaction processes. Energy Fuel 19:1160–1164
141. Couch GR. Recent developments in coal liquefaction, IEA Clean Coal Centre, www.iea-coal.org.uk
142. Williams RH, Larson ED (2003) A comparison of direct and indirect liquefaction technologies for making fluid fuels from coal. Energy Sustain Dev 7(103):129
143. Cicero DC (2008) Coal-to-liquids in the United States: status and roadmap, CTLtec America 2008, 23–24 June 2008
144. Bezdek RH, Wendling RM, Hirsch RL (2006) Economic impacts of U.S. liquid fuel mitigation options, DOE/NETL-2006/1237, 8 July 2006
145. Miller CL (2007) Coal conversion—pathway to alternate fuels. In: 2007 EIA energy outlook modeling and data conference, Washington, DC, 28 March 2007
146. (2009) Liquid transportation fuels from coal and biomass: technological status, costs, and environmental impacts. The National Academy Press, Washington, DC
147. Couch GR (2008) Coal to liquids. IEA Clean Coal Centre, London
148. O'Hara JB, Klumpe HW, Bela A, Jentz NE (1978) Project POGO-A coal refinery. Chem Eng Prog 74(8):46–63
149. Zoellar MR. Evolving production of the acetyls (acetic acid, acetaldehyde, acetic anhydride and vinyl acetate): a mirror for the evolution of the chemical industry. In: Flank WM, Abraham MA, Matthews MA (eds) Innovations in industrial and engineering chemistry. ACS symposium series, 1000 2009, pp 365–387
150. Song C (2009) Clean fuels and chemicals from coal with CO_2 capture for sustainable energy development. In: Preparation of symposium, ACS, Division of Fuel Chemistry, vol 54(2), p 622
151. Song C, Schobert HH, Anderson JM (2005) Premium carbon products and organic chemicals from coal. IEA Clean Coal Centre, Aug 2005
152. Report on applying coal gasification technology in China's coal chemical industry, UNESCO, 2007
153. Zhufeng Y (2007) Development and policy study of coal=based chemicals in China. In: China NEET workshop, Nov 2007
154. Coal to olefins, PERP report 05/06 S5, Nexant Chem Systems, April 2007
155. Jia H (2008) China revives coal chemicals drive. RSC Adv The Chem Sci, Chemistry World, RSC Advancing the Chemical Sciences, 9 April 2008
156. Wender I (1987) Chemicals from coal: new processes. In: Payne KR (ed) Critical reports on applied chemistry (Chapter 5), vol 14. Wiley, Chichester
157. Cornils B (1987) Chemicals from coal: new processes. In: Payne KR (ed) Critical reports on applied chemistry (Chapter 5), vol 14. Wiley, Chichester
158. Pruett RL (1981) Synthesis gas: a raw material for industrial chemicals. Science 211:11–16
159. Han AP, Vang Christensen P (2007) Ammonia plant safety & related facilities. 48:211–219
160. Song C (2009) Clean fuels and chemicals from coal with CO_2 capture for sustainable energy development, preprints. ACS Div Fuel Chem 54(2):662
161. Zoeller JR (2009) Eastman chemical company's "Chemicals from coal" program: the first quarter century. Catal Today 140:118–126
162. Moock DN (2001) Proc Ann Int Pittsburgh Coal Conf 18:83–100
163. Austin GT (ed) (1984) Shreve's chemical process industries, 5th edn. McGraw-Hill, New York
164. Haggin J (1989) Chem Eng News 67(7):28
165. Anderson LL, Tillman DA (1979) Synthetic fuels from coal: Overview and assessment Wiley-Interscience, New York
166. Moloy KG, Wegman RW (1989) Rhodium-catalyzed reductive carbonylation of methanol. Organometallics 8:2883
167. Campoy A, Oster S (2007) Coal-to-chemicals projects boom in China. Wall Street J 31
168. Trifiró F (2009) From syngas to methanol and dimethyl ether, presented at Summer School, Bologna, Sept 2009. www.ics.Trieste.it/media/139813/df6496.pdf
169. Sutton M, Roberts P. Methanol production—a key step in the return of chemicals from coal. Davy Process Technology. www.davyprotech.com
170. Siemens to Supply Coal Gasification Technology to Shenhua Group for DME Project. www.greencarcongress.com/2007/01/Siemens_to_supp.html 28 January 2007
171. Walter P (2009) Coal: king of chemicals? Chem Ind (London, UK) 18:16–17

Natural Gas

20

James D. Idol, Mahmood Moshfeghian, and Arland H. Johannes

Characteristics

Natural gas is a naturally occurring mixture of simple hydrocarbons and nonhydrocarbons that exists as a gas at ordinary pressures and temperatures. In the raw state, as produced from the earth, natural gas consists principally of methane (CH_4) and ethane (C_2H_4), with fractional amounts of propane (C_3H_8), butane (C_4H_{10}), and other hydrocarbons, pentane (C_5H_{12}) and heavier. Occasionally, small traces of light aromatic hydrocarbons such as benzene and toluene may also be present.

The ethane and heavier hydrocarbon components of natural gas are rather easily separated from the gas stream and liquefied under moderate pressure.

Most of these components are separated and recovered because of their added value when they are sold as separate products. Unprocessed natural gas is described as "rich" ("wet") or "lean" ("dry"), depending on the amounts of liquefiable heavier components present in the unprocessed gas. (The dividing line is not specific).

In addition to hydrocarbon components, raw natural gas from some reservoirs may contain varying amounts of nonhydrocarbon contaminants, or diluents, such as hydrogen sulfide (H_2S), carbon dioxide (CO_2), water (H_2O), nitrogen (N_2), and/or helium (He). Natural gas containing the first two of these compounds, hydrogen sulfide and carbon dioxide, is termed "sour" and the contaminants are referred to as "acid" gases. Natural gas that contains low enough concentrations of the acid gases to meet sales specifications is termed "sweet."

Although some natural gas is suitable for use as a fuel as produced, virtually all natural gas undergoes some processing to produce a merchantable gas the composition of which is principally methane and ethane. Water and acid gas components are removed to prevent freezing, corrosion, or other operating problems in transmission or utilization. Diluents that exist in significant quantities usually are removed to improve the combustion and/or heating properties of the gas. Much of the ethane and propane and essentially all of the butanes and heavier hydrocarbons are removed because of their greater value when sold as separate liquid products.

Table 20.1 gives typical raw gas compositions. "Casinghead" gas is typically collected from liquid–gas separators in oil production facilities. Because the separators generally operate at low pressures, the gas from this source tends to be richer in liquefiable hydrocarbons content. "Gas well" gas and "condensate well" gas generally are produced at higher pressure and with little or no accompanying liquid, so their compositions show less liquefiable hydrocarbon content. These characteristics describe "traditional" natural gas or gas from commonly thought of sources. Today gas is available or becoming available from nontraditional sources. This "unconventional gas" tends to have a composition that is much more like the "gas well" gas than "casinghead" gas.

Table 20.2 shows the physical properties of the principal constituents of natural gas as produced. Table 20.3 presents typical specifications for pipeline quality, or "merchantable," natural gas.

The common unit of measurement for natural gas is the "standard" cubic foot in the English system and the "standard" cubic meter in the SI system. Each of these "standards" is expressed at pressures and temperatures commonly used as standard to the system in the geographical area of concern. In the United States, where standards frequently vary from state to state, the cubic foot is frequently expressed in the English system at standard conditions of 14.69 pounds per square inch absolute (psia) and 60° Fahrenheit (60°F),

---

J.D. Idol (✉)
Department of Materials, Science and Engineering, Rutgers University,
607 Taylor Road, Room 203, Piscataway, NJ 08854, USA
e-mail: jdidol@rci.rutgers.edu

M. Moshfeghian
John Campbell and Company, 1215 Crossroads Blvd., Suite 230,
Norman, OK 73072, USA

A.H. Johannes
Oklahoma State University, Stillwater, OK 74078, USA
e-mail: aj@okstate.edu

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_20,

Table 20.1 Typical raw gas compositions (water-free basis)

	Casinghead (wet) gas		Condensate well gas		Gas well (dry) gas	
	(Mol. %)	(Gal/mscf)	(Mol. %)	(Gal/mscf)	(Mol. %)	(Gal/mscf)
Carbon dioxide	0.63					
Nitrogen	3.73		1.25		0.53	
Hydrogen sulfide	0.57					
Methane	64.48		91.01		94.87	
Ethane	11.98		4.88		2.89	
Propane	8.75	2.399	1.69	0.463	0.92	0.252
iso-Butane	0.93	0.303	0.14	0.046	0.31	0.101
n-Butane	2.91	0.914	0.52	0.163	0.22	0.069
iso-Pentane	0.54		0.09		0.09	
n-Pentane	0.80		0.18		0.06	
$iC_5 + nC_{5+}$		0.777		0.203		0.103
Hexanes	0.37		0.13		0.05	
Heptanes plus	0.31		0.11		0.06	
	100.00	4.393	100.00	0.875	100.00	0.525

Source: furnished in cooperation from the Gas Processors Suppliers Association Engineering Data Book, 12th Edition, 2004

Table 20.2 Properties of components in natural gas and natural gas liquids

	Chemical formula	Molecular mass	Boiling point @14.496 psi (°F)	Vapor pressure @100°F (psia)	Gas relative density air = 1	Heating value BTU/scf gas
Methane	CH_4	16.043	−258.73	(5,000)[a]	0.554	1010.0
Ethane	C_2H_6	30.070	−127.49	(800)[a]	1.038	1769.6
Propane	C_3H_8	44.097	−43.75	188.64	1.523	2516.1
i-Butane	C_4H_{10}	58.123	10.78	72.58	2.007	3251.9
n-Butane	C_4H_{10}	58.123	31.08	51.71	2.007	3262.3
i-Pentane	C_5H_{12}	72.150	82.12	20.45	2.491	4000.9
n-Pentane	C_5H_{12}	72.150	96.92	15.57	2.491	4008.9
Hexane	C_6H_{14}	86.177	155.72	4.96	2.976	4755.9
Heptane	C_7H_{16}	100.204	209.16	1.62	3.460	5502.5
Octane	C_8H_{18}	114.231	258.21	0.54	3.944	6248.9
Carbon dioxide	CO_2	44.01	−109.26		1.520	0
Hydrogen sulfide	H_2S	34.08	−76.50	394.59	1.177	637.1

Source: furnished in cooperation from the Gas Processors Association Standard 2145

[a] Extrapolated. Above critical temperature

although there are a number of other "standard" conditions specified by various regulatory and government agencies at various governmental levels. In metric units, the cubic meter is commonly defined as being at one atmosphere of pressure (101.325 kPa) and 15 degree Celsius (15°C).

Another common unit of measurement for natural gas is by use of its heating value, expressed in British thermal units (BTU) per standard cubic foot in the English system, and in Joules (or calories) per unit volume in the metric system. Commercially used natural gas, after processing, yields the equivalent of about 950–1,050 BTU/sft^3. Also used as a unit of measurement is the therm, equivalent to 100,000 BTU, or the nominal heat content of 100 standard cubic feet of lean, processed natural gas.

The most common unit of measurement in the English system is the mscf or thousand (10^3) standard cubic feet. Larger volumes, used to express production or pipeline volumes, are noted as bscf, or billion (10^9) standard cubic feet. Even larger volumes, such as reserve figures, usually are expressed as tcf, or trillion (10^{12}) standard cubic feet.

Occurrence of Natural Gas

Natural gas occurs in a number of different forms, depending primarily on the source: associated gas is gas produced "in association" with crude oil; nonassociated gas is gas produced from a gas well with little or no accompanying liquid;

Table 20.3 Typical pipeline quality natural gas [23]

	Minimum	Maximum
Major and minor components, mol %		
Methane	75	–
Ethane	–	10
Propane	–	2
Butanes	–	5
Pentanes and heavier	–	0.5
Nitrogen and other inerts	–	18
Carbon dioxide	–	3
Hydrogen	–	5
Total unsaturated hydrocarbons	–	0.5
Carbon monoxide	–	0.1
Trace components		
Hydrogen sulfide	–	0.25 gr/100 scf
Mercaptan sulfur	–	0.50 gr/100 scf
Total sulfur	–	1.0 gr/100 scf
Water vapor	–	7.0 lb/mm scf
Oxygen	–	10 ppmv
Other characteristics		
Heating value, BTU/scf		
Gross saturated	950	1,150
Gross dry	967	1,170
Relative density	0.530	0.740
Liquids: free of liquid water and hydrocarbons at delivery temperature and pressure		
Solids: free of particulates in amounts deleterious to transmission and utilization equipment		

gr/100 = grains/100

Source: furnished in cooperation from the Gas Processors Suppliers Association Engineering Data Book, 12th Edition, 2004

gas produced in combination with liquid condensate which is much lower in molecular weight and lighter in color than crude oil; and "coal bed methane (CBM)," which is gas produced from relatively shallow (3,000–5,000 ft) formations containing primarily coal but that also contain marketable quantities of high methane content natural gas. Typical compositions for the first three of these categories are shown in Table 20.1. Figure 20.1 is a simplified schematic layout of the flow of natural gas from the producing well through various stages of processing.

Associated gas is found in crude oil reservoirs, either dissolved in the crude oil or in conjunction with crude oil deposits. Dissolved or associated gas provides the driving force or reservoir pressure necessary to produce oil from a reservoir. Dissolved gases are produced from oil wells along with the oil. The gas separates, or is separated from the crude oil at the well head. Such gas also may be called "casinghead gas" or "oil well gas." In the early days of the industry, virtually all the natural gas usefully utilized was associated gas from oil wells. Much of the very large gas reserves of several Persian Gulf countries and the Prudhoe Bay reserves of Alaska are associated gas.

Nonassociated gas occurs in reservoirs separate from those of crude oil. It is commonly referred to as "gas well gas," and contains much less of the heavier, or condensable, hydrocarbons that are found in associated gas.

Recently huge reserves of "nonassociated" or "gas well" gas have been discovered in the Persian Gulf area.

Today, because of the wide demand for natural gas, most utilized gas is nonassociated gas.

The third form in which conventional natural gas reserves occur might be called a single-phase fluid, because it is neither a true gas nor a true liquid. It is not a gas or liquid because no surface boundary exists between gas and liquid. Reservoirs of this type are called "gas condensate" reservoirs, and usually are found in moderately deep formations, have very high pressures, and pose special problems in production and processing.

In recent years gas wells have been drilled at increasing depths. This greatly increases drilling costs and new wells may cost many millions of dollars to drill and bring into production. At least a partial remedy for this has been the rapid development of "CBM" as a major source of natural gas in the United States. The wells are relatively shallow—around 5,000 ft maximum at present—and much less expensive to drill, complete, and bring to onstream production.

In addition, CBM appears to be widely distributed across the lower 48 states of the United States.

Early on the subsurface coal mining industry learned that "swamp gas" or "firedamp" was prevalent in underground coal mines. The gas is largely methane formed from the decay and rotting of carbonaceous materials to ultimately form coal. When pressure is released (as by drilling into the coal seam) the methane-rich gas migrates to the low pressure and is produced. The composition and properties of several CBM streams are shown in Fig. 20.2 [1].

CBM has rapidly gained a major position in the lower 48 states natural gas production. Figure 20.3 [2] shows, by year, the number of producing CBM wells in the lower 48 U.S. states. The total number producing for the year 2000 was almost 14,000, and the total gas production as shown in Fig. 20.4 [2] was almost 1.4 tcf. This works out to an average per well production in the order of 40 mcf per day per well.

Evolution of the U.S. Natural Gas Industry

Although natural gas is the predominant product (98%) of the gas industry today, its utilization evolved out of the production and use of manufactured gas, and its major expansion came about through utilization of early "town gas" systems originally installed to distribute manufactured gas.

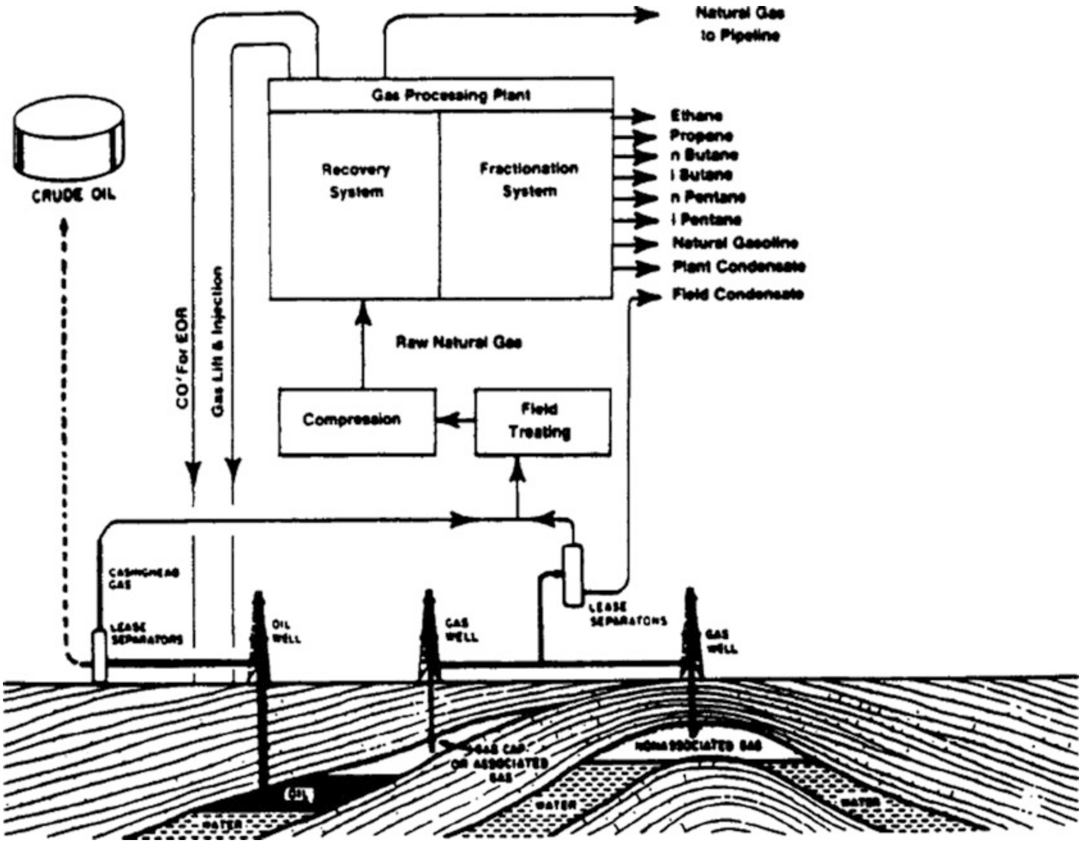


Fig. 20.1 Schematic flow of natural gas from well through processing

BASIN	STATE	DEPTH	BTU	N_2	CO_2	C_1	C_2	C_3	nC_4	iC_4	nC_5	iC_5	Other
SAN JUAN BASIN	NM	2792	914	3.500	6.600	89.500	0.200	0.000	0.000	0.000	0.100	0.010	0.090
SAN JUAN BASIN	CO	2576	930	1.500	7.500	90.100	0.500	0.300	0.000	0.000	0.000	0.000	0.100
SAN JUAN BASIN	NM	2580	1029	0.900	3.500	91.100	2.600	1.100	0.500	0.000	0.200	0.000	0.100
WARRIOR BASIN	AL	2162	995	0.900	0.160	98.200	0.020	0.000	0.000	0.000	0.000	0.000	0.720
GREEN RIVER BASIN	WY		957	1.000	4.290	94.600	0.110	0.000	0.000	0.000	0.000	0.000	0.000
PICEANCE BASIN	CO	6344	1001	0.370	10.470	81.710	4.050	1.800	0.410	0.580	0.120	0.220	0.270
POWDER RIVER BASIN	WY	402	969	3.630	0.470	95.860	0.010	0.000	0.000	0.000	0.000	0.000	0.030
ARKOMA BASIN	OK	837	1002	0.900	0.900	96.400	1.400	0.000	0.000	0.000	0.000	0.000	0.400
CHER. & FOREST CITY	KS	983	976	2.160	1.170	96.090	0.380	0.110	0.020	0.010	0.000	0.000	0.060
SAN JUAN BASIN	NM	2843	996	1.060	4.950	90.690	1.880	0.810	0.200	0.150	0.050	0.070	0.140

Key for Items in Table:
Depth, feet
Heating value, BTU/Standard Cubic Foot at 14.73 psia and 60°F
N_2 = Nitrogen
CO_2 = Carbon Dioxide
C_1 = Mehane
C_2 = Ethane
C_3 = Propane
iC_4 = Iso Butane
nC_4 = Normal Butane
iC_5 = Iso Pentane
nC_5 = Normal Pentane
Other---To obtain “other” add nitrogen through nC_5 in each row and subtract from 100.00. The difference is “other.”

Fig. 20.2 Composition of several coal bed methane streams [1]

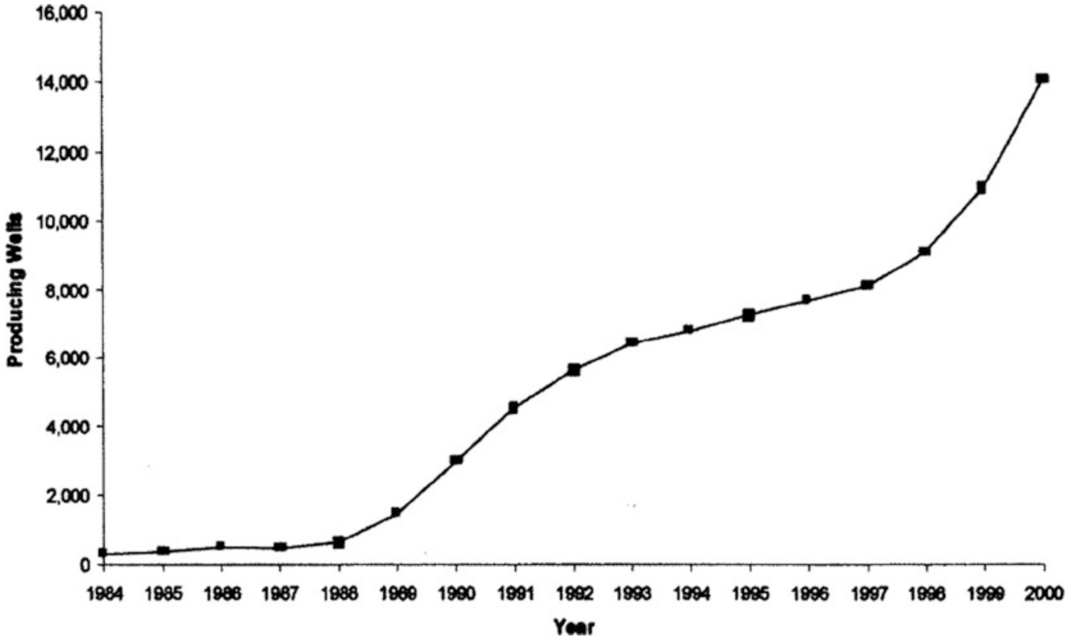


Fig. 20.3 Number of producing coal bed methane wells in lower 48 states (USA) [2]

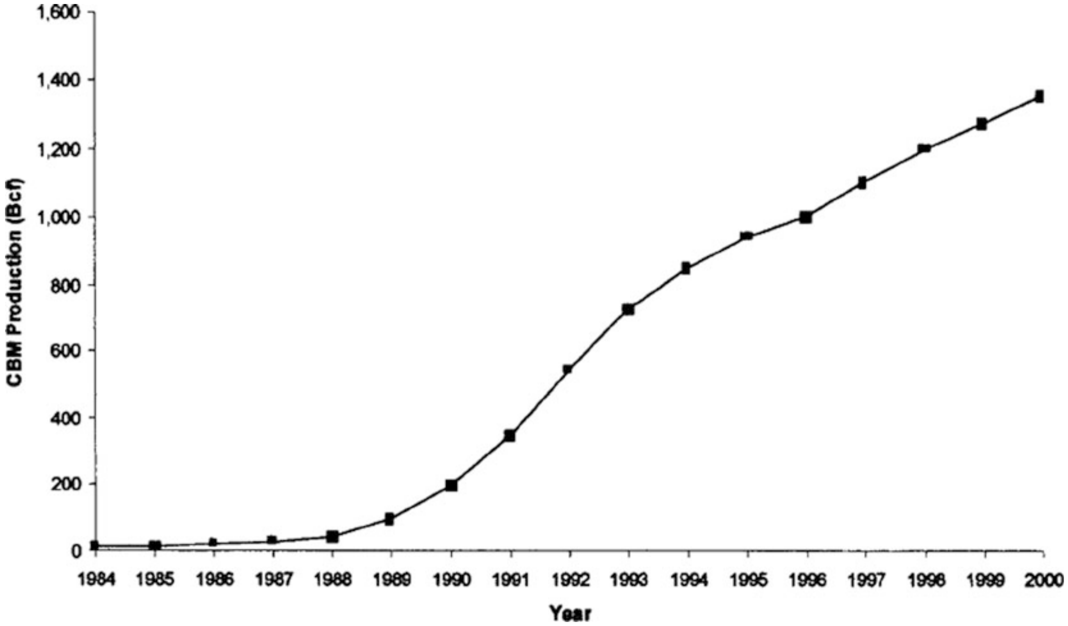


Fig. 20.4 Annual coal bed methane production for lower 48 states (USA) [2]

Manufactured gases are those obtained by destructive distillation of coal, by the thermal decomposition of oil, or by the reactions of steam passing through a bed of heated coal or coke. Prior to the development of large natural gas reserves, manufactured gas was used widely for illumination in the "gaslight" era of the United States.

Following the discovery of major oil and gas reserves in the early twentieth century, associated gas became plentiful in oil-producing areas, and natural gas began to displace manufactured gas in town systems. However, early natural gas utilization was confined largely to areas at or very near the source of supply because methods and materials for transporting the gas had not yet been developed.

Natural gas is relatively difficult to transport and store and is extremely bulky compared with other forms of energy. At atmospheric pressure, one cubic foot of natural gas contains about two-tenths the energy content of a cubic foot of solid or liquid fuel. This is the primary reason natural gas is transported through pipelines at very high pressures, typically 900–1,000 psia. The development of high tensile strength, thin wall pipe, and modern welding techniques in the mid-1930s permitted the construction of numerous long-distance, high-pressure gas transmission lines from the large and growing gas reserves of the southwestern United States to the major industrial and population centers of the country.

In rapid succession, a number of long-distance pipelines were laid during the period from 1930 to 1940, including three pioneer 1,000-mile lines from areas containing large natural gas reserves in the Southwest to major markets in the North and the Northeast. Spurred by abundant and inexpensive gas reserves, the U.S. natural gas industry expanded rapidly after 1945, and natural gas virtually replaced the more expensive manufactured gas in city distribution systems.

In 1930, marketed production of natural gas totaled 1.9 tcf; following early pipeline expansion, total consumption by 1940 had risen to 2.7 tcf. Then, in the decade between 1940 and 1950, gas consumption more than doubled to 6.3 tcf, and it doubled again to 12.8 tcf by 1960. Marketed production in the United States peaked at 22.7 tcf in 1973, followed by significant declines resulting from conservation efforts. By 1990 the marketed production of natural gas was about 18 tcf per year, and demand was increasing steadily because of the convenience of natural gas and its qualities as an environmentally desirable fuel.

U.S. Marketed Production [3–5]

Natural gas accounts for well over one-half of total U.S. petroleum energy production. Its major market segments are residential and commercial fuels, industrial fuel and chemical feedstock, and electric power generation.

Residential consumption is predominantly for space heating, cooking, and water heating. Commercial use includes space heating and other fuel needs for commercial establishments such as stores, hotels, and restaurants. Industrial use, accounting for about 35% of total gas consumption, includes fuels for manufacturing operations, industrial boiler fuel, and process heat. Oil refineries are major consumers of natural gas for refining operations. In addition, natural gas is the prime feedstock for the manufacture of ammonia fertilizers, methanol, and other basic chemical derivatives. Figure 20.5 is a simplified schematic that shows some of the principal chemical and commercial products derived from natural gas.

Fuel for electric power generation accounts for about 20% of marketed natural gas production. This market segment is extremely price-sensitive and competes directly with fuel oil and coal. Given the current environmental emphasis on reduction of air pollutants, the use of cleaner-burning natural gas is expected to increase significantly in the foreseeable future.

In addition to what is termed “marketed production” for these uses, substantial quantities of natural gas are consumed annually as lease, processing plant and pipeline fuels.

At current production levels, a majority of U.S. petroleum energy production consists of the gaseous fuels natural gas and natural gas liquids. Moreover, since reaching peak production in 1985, U.S. crude oil production has consistently declined while gaseous fuel production has steadily increased. These trends suggest that the U.S. petroleum economy will become even more dependent on the gaseous fuels in the future.

Natural Gas Liquids

Included in the gaseous fuels segment of U.S. petroleum energy production are the natural gas liquids extracted from natural gas as produced and before sale. These natural gas liquids, commonly termed “NGLs,” are ethane, propane, butanes, pentanes, and higher molecular weight hydrocarbons. The recovery and separation of natural gas liquids from raw natural gas is a relatively simple operation, which relies principally on the differences in boiling points and vapor pressures of the various components. (These differences are shown in Table 20.2.) The two most prevalent recovery processes are the absorption process and the cryogenic turbo-expander process. Together, they account for approximately 90% of total U.S. natural gas liquids production. The bulk of U.S. gas liquids is produced as a raw mix stream in some several hundred gas-processing plants located in or near major gas-producing provinces. Typically, the raw mix streams are transported through an extensive pipeline network to large central fractionator complexes where the mixed liquid stream is fractionated into its components ethane, propane, butanes, and pentanes.

The total U.S. recovery of natural gas liquids is about 30–35 barrels per million cubic feet of marketed natural gas. The total NGL supply is supplemented by about 200–250 million barrels per year of refinery production, which equates to about 5% of the total crude oil charge to refineries. Some of the petrochemical products produced from natural gas and NGLs are shown in Fig. 20.5.

Major demand segments for NGLs indicate that around 35–40% of gas liquids, principally ethane and propane, are consumed as cracking feedstocks for ethylene manufacture. Some of the chemical derivatives obtained from ethane and propane are shown in Fig. 20.6.

Traditionally, a major demand for natural gas liquids, mainly butanes and pentanes, has been in the manufacture of motor gasolines. However, the recent phasing out of

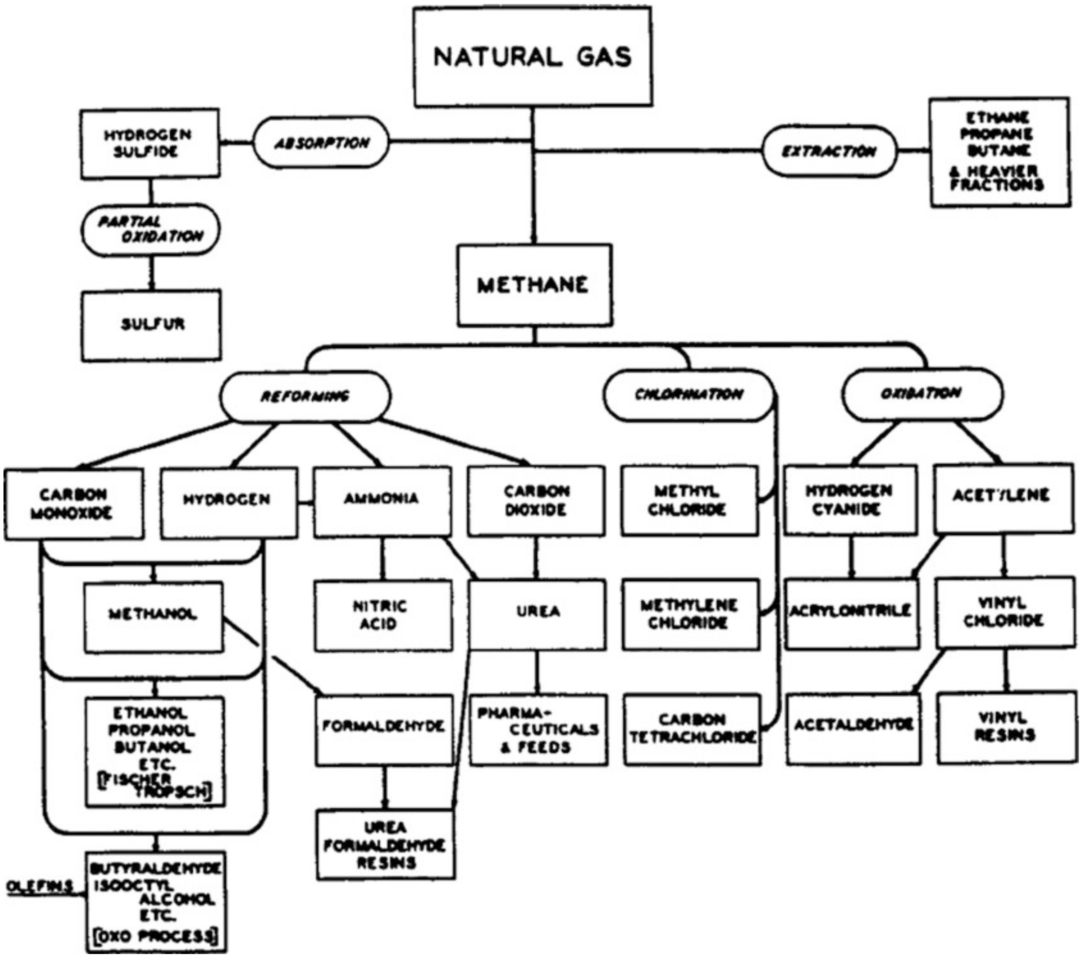


Fig. 20.5 Some petrochemical products derived from natural gas (Crockett and Widgery [35]. Reprinted by permission of the Gas Processors Suppliers Association)

leaded gasoline and the current, ongoing reduction in motor gasoline volatility have drastically reduced demand for butanes and pentanes as gasoline blending stocks. On the other hand, these same measures, enacted to reduce air pollutants from the automobile, have resulted in an increased demand for these products as prime feed stocks for the formulation of clean-burning reformulated gasolines.

A third major market is consumption of propane as both a residential and commercial fuel, principally in rural areas beyond the reach of natural gas distribution systems. Additional major uses indicated in the "other" category of Fig. 20.6 include consumption as internal combustion engine fuels and numerous agricultural uses, such as crop drying.

U.S. Natural Gas Reserves

At any given time proven U.S. natural gas reserves seem to remain at approximately 10 years of indigenous supply at the then current consumption rates. Proven reserves estimates are subject to constant change, either up or down, depending on reservoir operating experience, field extensions resulting from additional drilling, and revised estimates.

Not included in reserves data are undiscovered but probably recoverable reserves in unexplored or unproven provinces. Currently, these probable reserves are estimated by the U.S. Department of Energy at 307–500 tcf, much of which almost certainly will become proven reserves when they become economically feasible.

Whatever the ultimate recoverable reserves may be, current estimates of proven reserves of natural gas comprise about 50% of U.S. proven reserves of total petroleum energy.

Structure of the U.S. Natural Gas Industry

The natural gas industry in the United States is made up of four major segments: production, processing, transmission, and distribution. The industry is physically interconnected by a

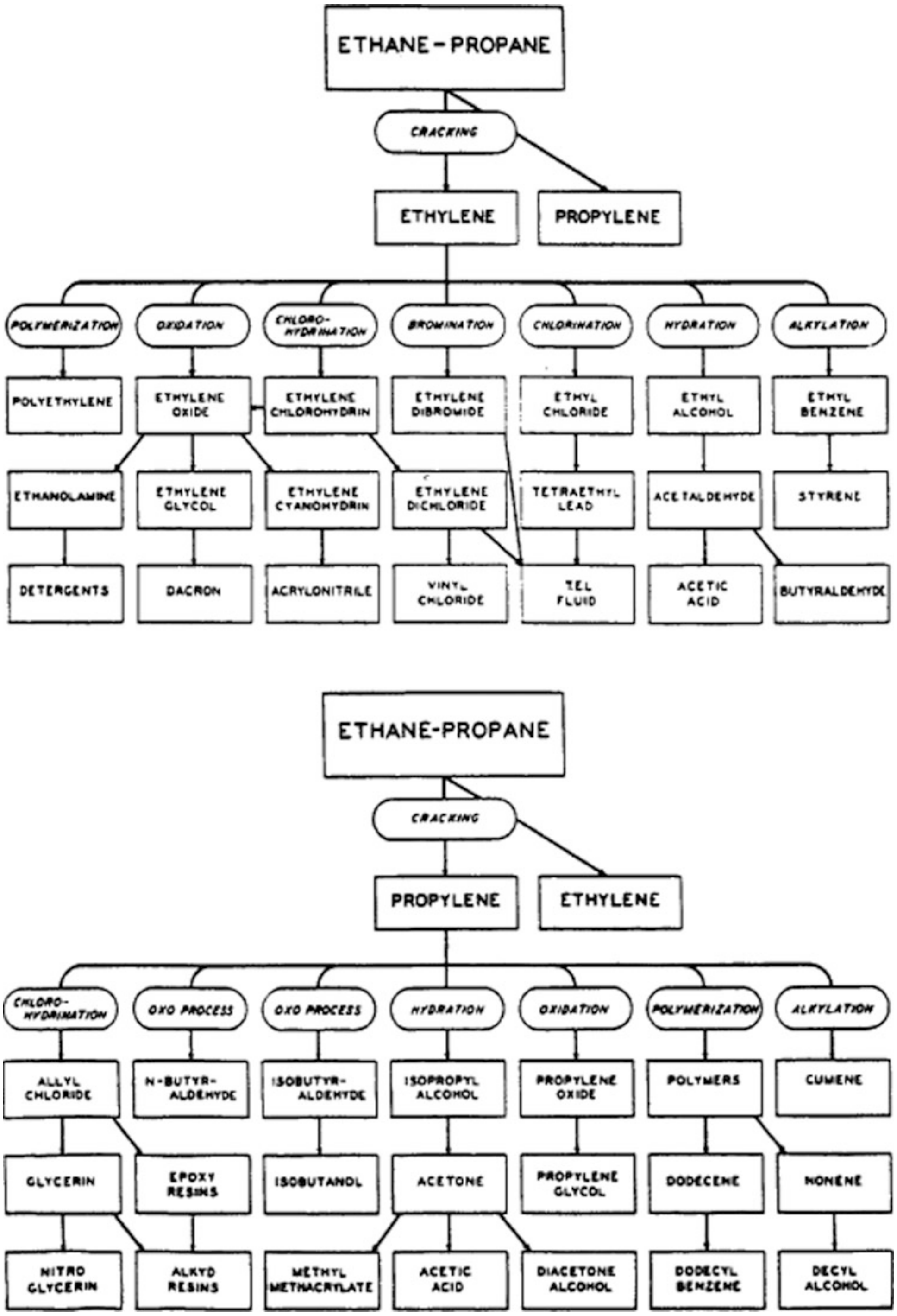


Fig. 20.6 Some products derived from ethane and propane (Crockett and Widgery [35]. Reprinted by permission of the Gas Processors Supplier Association)

pipeline network that extends through the lower 48 states and across international borders into Mexico and Canada.

The production function is a large part of the oil-producing industry, and most of the same firms are dominant in both oil and gas. Exploration and drilling technologies for both oil and gas are essentially identical. Production from both oil well gas and gas well gas is introduced into large field gathering systems for delivery to the processing facilities. A typical gathering system may include 2,000 or more miles of gathering lines connected to 1,000 or more producing wells.

The gas-processing function includes the gathering and delivery of gas streams from a field or fields into a central processing facility. There are currently several hundred gas-processing plants in operation in the United States. In these plants the raw natural gas is dehydrated to remove moisture, treated for removal of contaminants and, if necessary, compressed to pipeline pressure. During treating, most raw gas also is processed for recovery of liquid products, including ethane, propane, butanes, and natural gasoline. These liquid products are then sold separately to be used for petrochemical and gasoline feedstocks and other fuel uses.

The transmission function covers transport of the processed pipeline quality gas from the gas-processing plant to major markets. About two-thirds of U.S. marketed gas production reaches the ultimate consumer through a network of some 250,000 miles of pipelines. Other deliveries may be made directly to industrial consumers or to city distribution systems.

The distribution function receives natural gas from the transmission pipeline at the city gate of the local distribution system and delivers it to the ultimate consumer through a network totaling some 600,000 miles of distribution pipelines.

World Natural Gas

Natural gas, in recent years, has become an international commodity fuel. Current world reserves of natural gas are estimated at about 6,500 tcf [6], the equivalent of about 40% of world petroleum energy reserves. About 35% of these reserves are in remote regions of the Former Soviet Union (FSU) and about 40% are located in the Persian Gulf countries of the Middle East.

The location of these reserves, far from the major populations of consuming countries, suggests that the world natural gas industry is still in a developmental stage and that further extension and expansion can be expected. Gas from the Groeningen field of the Netherlands, from Norway and the North Sea, and from the FSU is flowing by pipeline into industrialized Western Europe. In addition, natural gas is piped from the FSU into Eastern Europe.

Additional international movements of natural gas are in the form of liquefied natural gas (LNG), principally from Northern Africa and the Middle East into Japan, South Korea, and the United States.

It is certain that world use of natural gas will increase dramatically in the near future as industrialized countries replace coal-fired facilities with cleaner-burning natural gas. In addition, a number of countries remote from major markets are in the process of installing world-scale plants for utilization of natural gas and gas liquids for production of fertilizers, methanol, premium gasoline blending stocks, and other basic petrochemical derivatives that will result in higher-value products from natural gas and gas liquids for which there are no local markets.

Already, the gaseous fuels account for approximately 40% of world petroleum energy production and nearly one-half of world petroleum energy reserves. Moreover, the growth of world natural gas consumption is already increasing rapidly.

Despite a world surplus of crude oil capacity for the foreseeable future, it seems certain that the world has entered the “Age of Gaseous Energy,” and that natural gas and gas liquids will continue to move into a position of petroleum energy dominance in the next decade.

Gas-to-Liquids Technology

In contrast to the United States, many oil- and gas-producing areas of the world do not have a well-developed gas distribution system and large numbers of consumers want a clean-burning, environmentally friendly industrial and residential fuel. Also, many of these areas have the potential capability for producing large quantities of natural gas for long periods of time. One way of solving this problem is to market LNG. Atmospheric pressure tankers are the popular choice, and they must operate at or near the atmospheric boiling point of methane, which is ~161.52°C or ~258.74°F. Huge quantities of compression horsepower are required to produce the LNG, and huge quantities of heat are required to gasify the LNG after it has reached port and before it is suitable for consumer use.

Despite these difficulties and obstacles, there are several successful LNG operations in various parts of the world. Japan, which has no petroleum or natural gas reserves, has been particularly active in trying to import LNG and other products. As fuel costs rise, LNG inherently becomes more competitive. Additionally, LNG liquefaction plants are becoming more efficient and cost effective. The combination of these two effects indicate a bright future for tankered LNG.

Before World War II, German scientists developed processes for converting gas primarily formed by burning coal to hydrocarbons and other chemicals with molecular weights in the motor fuel range. Processes using these reactions and catalysts were used for some fuel and chemicals production during World War II. After the hostilities a good share of the Fischer–Tropsch “know-how” and technology was liberated and put within the purview of the U.S. Bureau of Mines. In the early 1950s the South African government made a political decision to produce gasoline and chemicals from their plentiful internal supply of coal and not be dependent on imported petroleum. This led to the design and construction of multiple integrated chemical–motor fuel plants to serve

Table 20.4 Summarized data for the first three major GTL complexes built

Plant site	Date of commissioning	Design capacity (Mbpd)	F-T unit technology	Investment/daily capacity barrel ($/bbl)
Montunui, New Zealand	1985	14.5	Mobil	102,000
Mossel Bay, South Africa	1993	27.5	Sasol	127,000
Bintulu, Malaysia	1993	12	Shell	125,000

Based on information in Samsam Bakhtiari [7]. Copyright 2001 by Gulf Publishing Company, all rights reserved

Table 20.5 Summarized data for the three second-generation GTL complexes presently under construction

Plant site	Scheduled startup	Design capacity (Mbpd)	F-T unit technology	Investment estimate/daily capacity barrel ($/bbl)
Australia, NW Shelf	2004	11.5	Syntroleum	52,000
Nigeria, Escravos	2005	34	Sasol	23,500
Qatar, Ras Laffan	2005	34	Sasol	23,500

Based on information in Samsam Bakhtiari [7]. Copyright 2001 by Gulf Publishing Company, all rights reserved

the internal requirements for motor fuel and simultaneously provide a variety of chemicals for both export and internal consumption. Derivatives of the original Fischer–Tropsch and later Sasol processes are currently being used to try to make marketable natural gas discoveries in areas where there is no industrial or home-based demand for clean-burning natural gas fuel.

As crude oil exploration spread to cover more of the globe, discoveries of huge natural gas reserves spread to more inhospitable places and thoughts of plants to produce longchain liquid hydrocarbons became more enticing. In the early 1980s planning for such plants began in earnest and in the space of a few years three separate plants utilizing three different liquidation technologies were well under way. Table 20.4 [7] shows some of the key parameters for the three plants. All the investment numbers are too large for the projects ever to be profitable. Estimates now are that about $50,000 investment per daily capacity barrel is about the maximum the traffic can bear.

Plans for several more liquidation plants had developed but the cost, construction, and operating problems of the first three slowed investor ardor and caused reassessment of the possibilities and goals. Expectations had been that by the mid 1990s a dozen or more plants would be in operation producing more than a million barrels per day of automotive fuel range hydrocarbons. Instead, the New Zealand plant was switched to producing methanol and the other two were continuing to encounter operating problems.

Three second-generation gas-to-liquid (GTL) plants are known currently to be underway. Some details for each are shown in Table 20.5 [8]. Investment per daily capacity barrel is much lower than for the first three. But the total capacity of the three new plants is only 80,000 barrels per day. If GTL technology proves out, it will be at a much more sedate pace than originally envisioned.

Preparing Natural Gas for Transmission and Sale

To make a product suitable for sale, the various contaminants and undesirable constituents contained in the raw gas must be removed. The first step is to separate the gas from any liquids and/or solids. This initial separation usually is carried out in near proximity to the producing well. The separators used depend primarily upon the difference in density between the gas and the liquid for separating the two phases. These separators may be classified as vertical, horizontal, or spherical, depending on their geometry and physical orientation [9].

The liquids produced are sold as oil or condensate, depending upon the type of reservoir from which the product is obtained. The natural gas stream after this first stage of separation contains only gaseous components, though some of these may be liquefied by later processing.

Typical compositions for gas streams at this point are shown in Table 20.1. (The compositions in Table 20.1 are on a water-free basis.) As typically produced, each of the gas streams would be saturated with water vapor. Developing the technology for handling gas streams that contain significant quantities of water vapor was one of the keys necessary to the development of the natural gas industry as it exists today.

In the presence of water, natural gas can, under the proper conditions of temperature and pressure, form hydrates. To the naked eye the hydrates are similar in appearance to snow; and they can result in plugging of pipelines and flow systems. Hydrates can form at temperatures well above the freezing point of water. In the early days of pipelining natural gas under pressure, shutdowns of the lines during the winter months frequently were caused by plugging with hydrates.

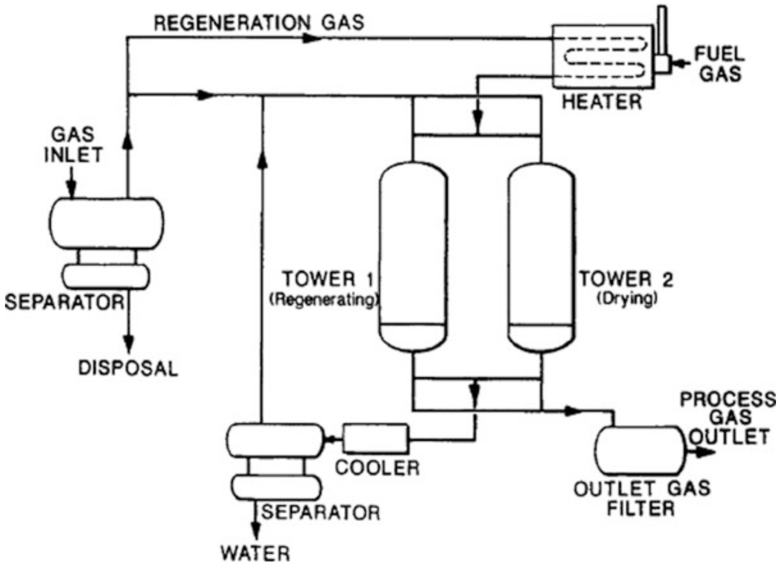


Fig. 20.7 Two-tower adsorption dehydration unit

A hydrate is a type of chemical compound called a clathrate, defined as a solid molecular compound in which one component is trapped in the cavities of cage-like crystals of another component. In the natural gas hydrate, water molecules form the "cage," and hydrocarbon molecules are the trapped component(s).

Hydrate formation can be prevented or avoided in two ways: (1) by dehydration, in which water is removed from the gas stream so that under conditions of processing or transporting, no liquid water will condense; and (2) by inhibition, which involves injecting into the gas stream a component that will dissolve the water and thus interfere with the ability of the water to form hydrates with the gas. Both procedures are widely used in the gas industry.

Of the wide variety of materials that can effectively depress the hydrate-forming tendency of a natural gas, three—methyl alcohol, ethylene glycol, and diethylene glycol—are used almost exclusively. Ethylene glycol ($C_2H_6O_2$) and diethylene glycol ($C_4H_{10}O_3$) are both polyhydroxy alcohols. They are high-boiling-point materials that mix in all proportions with water and both remain predominantly in the liquid phase. At temperatures approaching the freezing point of water, solutions containing more than 50% by weight ethylene glycol (or diethylene glycol) have viscosities approaching 10 cp. This limits their utilization in some low-temperature applications.

Methyl alcohol (CH_3OH) has a much lower boiling point and consequently a much higher vapor pressure than either ethylene or diethylene glycol. Hence procedures used for estimating the amount of methyl alcohol required must take into account the amount of methyl alcohol that vaporizes into the gas phase [10–12]. Although the necessity for vaporization requires the injection of greater quantities of methanol, it does have potential advantages. If the gas phase were cooled in further processing or transmission, one would expect some methanol to condense along with additional water and/or hydrocarbons. This would provide a safety factor for hydrate prevention that does not exist with the glycols.

Dehydration of the gas can be accomplished through the use of either a liquid or a solid desiccant. The solid desiccants most often employed include alumina, silica gel, and molecular sieves. The liquid desiccant most frequently used is triethylene glycol. Any of the solid materials used as desiccants must be regenerated periodically, so there must be at least two adsorption towers if dehydration of the gas is to be continuous. A schematic of a typical two-tower adsorption plant is shown in Fig. 20.7. The adsorbent must be heated to high temperature (about 400°F, 200°C) to remove the water and regenerate the solid desiccant. This usually is accomplished by heating a slip stream of the gas and circulating the heated gas through the off-stream adsorbent bed.

Figure 20.8 shows a schematic flow diagram for a typical triethylene glycol dehydration system. The lean glycol is pumped to the top of an absorber column and flows downward in countercurrent contact with the water–wet gas entering the bottom of the tower.

The absorber may contain either plates or packing. Dehydrated gas leaves the top of the absorber and the glycol diluted by the absorbed water flows from the bottom into a regenerator. After the absorbed water has been fractionated from the triethylene glycol solution, the solution is recirculated to the absorber. The concentration of the glycol leaving the reboiler will be approximately 98.5–98.7% by weight triethylene glycol if the reboiler operates at 400°F (205°C) and one atmosphere pressure.

Normal contract specification for gas to be transmitted through a high-pressure pipeline is a water content of 7 lb of

Fig. 20.8 Glycol dehydration unit

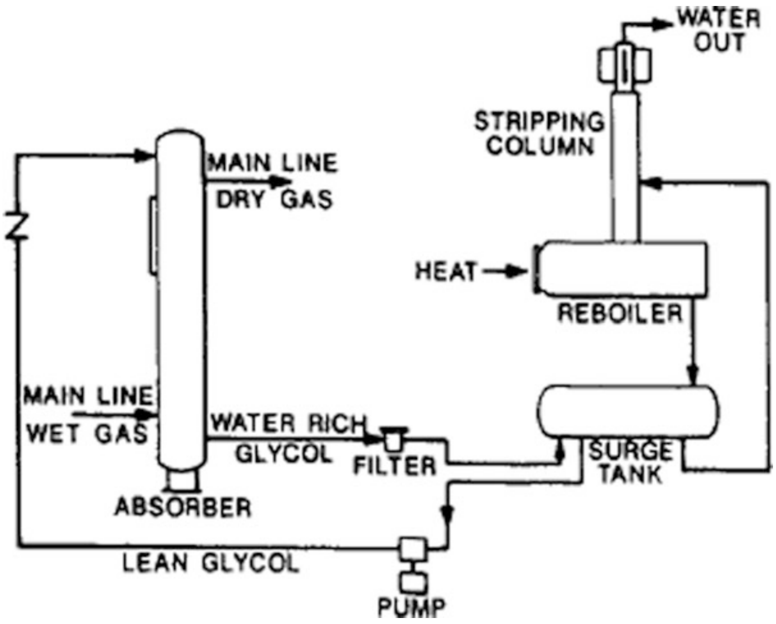
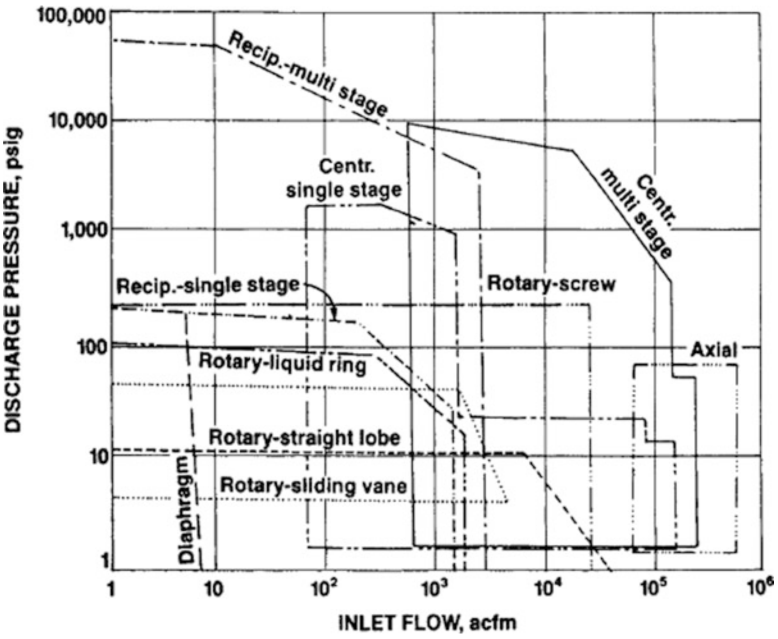


Fig. 20.9 Approximate range of application for compressors (GPSA Data Book, 12th Edition, 2004; Copyright Gas Processors Suppliers Association, and reprinted by permission of the copyright owner)



water per million standard cubic feet of natural gas. This is approximately the water content of natural gas in equilibrium with water at the freezing point (32°F) when the gas is under a pressure of 1,000 psia.

The solid desiccants offer much lower water contents for the dehydrated gas. With proper design and operation, molecular sieve dehydrators can satisfactorily prepare gas for total liquefaction at temperatures as low as −263°F (−165°C). Lower outlet water dew points can be obtained with triethylene glycol if the concentration of glycol going to the absorber is increased. This can be accomplished by using a vacuum in the reboiler or by introducing another material into the reboiler to lower the effective partial pressure of the water vapor above the glycol. Proper design and operation of such glycol systems can produce outlet water dew points in the −40°F (−40°C) range, or lower.

Natural gas pipelines normally operate at elevated pressures. In many instances the gas is available at low pressures so must be pumped or compressed to higher pressure. If the volume of the gas stream is small, reciprocating compressors must be used. For larger gas volumes, rotary compressors of various types are available. Figure 20.9 [13] shows pressure and flow rate ranges in which different types of compressors currently available are used.

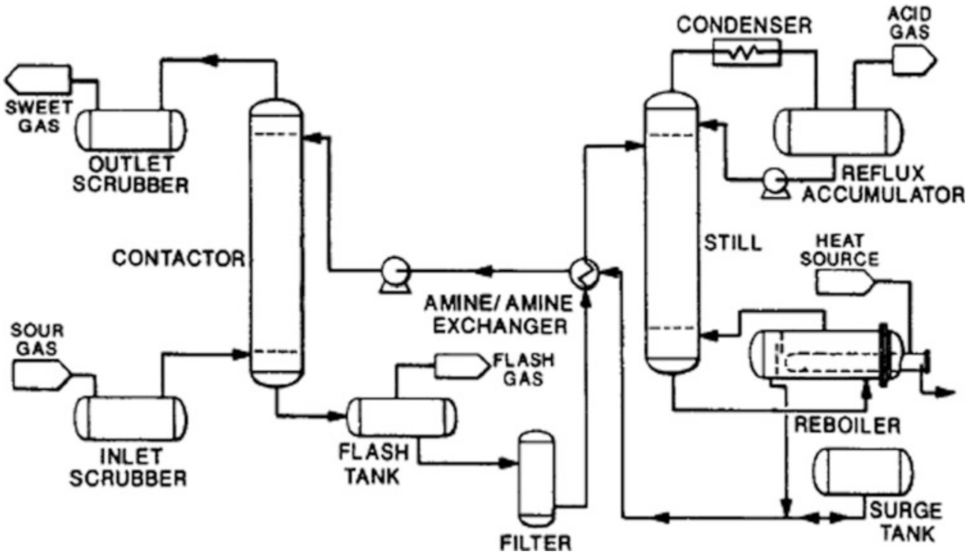


Fig. 20.10 Typical ethanolamine sweetening unit

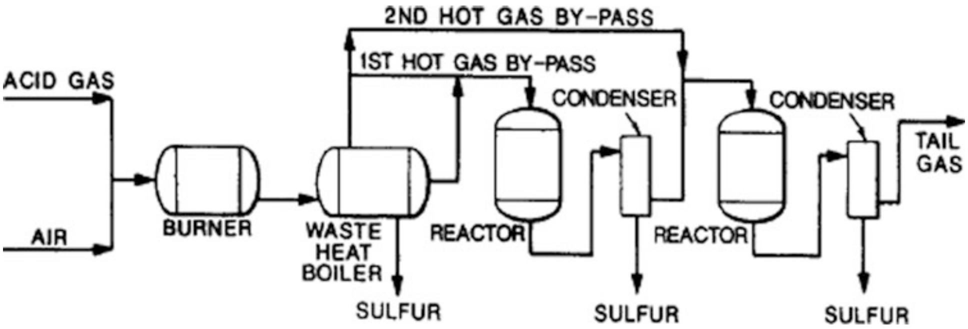


Fig. 20.11 Once-through Claus sulfur unit

If the natural gas stream contains unacceptable quantities of hydrogen sulfide and/or carbon dioxide, they must be removed in order to make the gas suitable for transmission and sale. The details of removal of H_2S and CO_2 from natural gas streams are beyond the scope of this chapter, but excellent discussions are available [13, 14]. There are many different processes available, depending upon the contaminants to be removed and their concentration in both the sour gas available and the sweetened gas to be produced. The dominant treating process is still the use of an alkanolamine. A typical flow diagram for an amine sweetening installation for removal of hydrogen sulfide and carbon dioxide from a natural gas stream is shown in Fig. 20.10 [14].

Environmental restrictions and controls normally will not permit the release of the hydrogen sulfide and/or its incineration to sulfur dioxide. As pointed out by Maddox [14], the least undesirable alternative is conversion of the hydrogen sulfide to sulfur. Again, there are many different processes available, several of which are discussed in the book by Maddox. The dominant method is by use of the Claus process, which involves partial combustion of the hydrogen sulfide to sulfur dioxide and catalytic combination of the hydrogen sulfide and sulfur dioxide to produce sulfur. Detailed discussions of the Claus technology as well as other techniques of converting hydrogen sulfide to sulfur are available [13, 14].

Figure 20.11 is a sketch of a “once-through” Claus process using a hot gas bypass for reheat. Sufficient air is used to burn one-third of the H_2S and all hydrocarbons. Steam usually is generated in the waste heat boiler, where the combustion products are cooled, and sulfur is condensed. The gas must be reheated before introduction to the reactor to prevent sulfur condensation in the reactor bed. Tail gas usually is sent to a tail gas cleanup unit to minimize sulfur emissions.

Processing for Liquids Recovery

The first time that liquids were recovered from a natural gas stream probably happened more by accident than by design. Any time a gas stream from a liquid–gas separator is either cooled or compressed, liquids will form. This simple statement still provides the basis for nearly all hydrocarbon liquid recovery systems from natural gas. Figure 20.12 [15] shows

Fig. 20.12 Liquid hydrocarbon recovery by compression and cooling (*adapted from Huntington* [15])

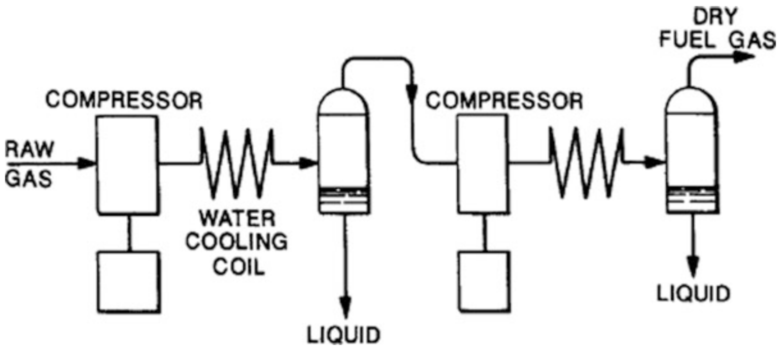
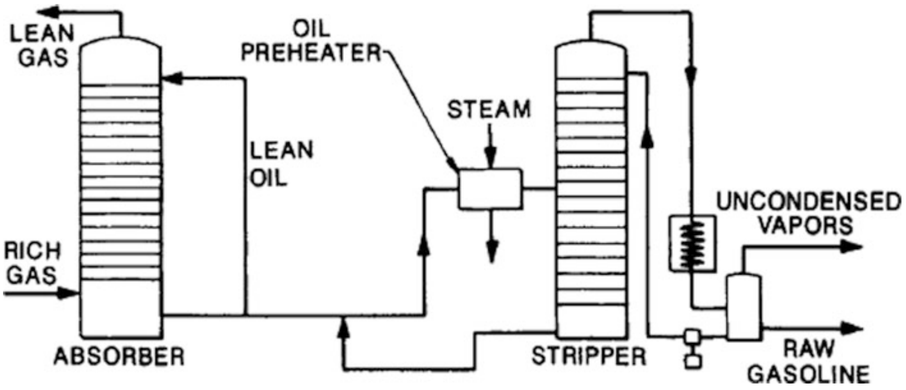


Fig. 20.13 Absorption-stripping unit (*adapted from Huntington* [15])



schematically the flow through a simple compression liquids-recovery facility. Replacing the water cooling after the second stage of compression with refrigeration would increase liquids recovery. If the gas had not been dehydrated, there would be danger of hydrate formation at this point unless the water had been removed.

As the natural gas industry developed and expanded, so did the demand for liquids recovered from natural gas. One way of recovering additional liquids is through the use of a heavy absorption oil that has good solubility characteristics for the propane and heavier hydrocarbon components in the gas stream. As in the case of dehydration, the solvent must be regenerated. Figure 20.13 [15] shows the absorption-stripping section of an absorption-type hydrocarbon liquids-recovery facility. The absorbent is circulated to the top of the absorber and flows downward, absorbing liquefiable components from the gas stream. It then is heated in the stripper, with the stripped components being cooled and condensed while the stripped absorption oil is recirculated. The liquid hydrocarbon stream typically would flow through a fractionation train for separation into the individual components or mixtures that could be sold.

Combining refrigeration with oil absorption allowed for additional liquid recoveries and greater economy of operation. Using temperatures as low as −40°F (−40°C), refrigerated oil absorption plants were capable of recovering 70% or more of the propane present in the natural gas stream. Recoveries of the butane and heavier constituents were essentially 100%.

Development of the turbo-expander process allowed the design and construction of plants for recovery of liquid ethane, as well as the heavier hydrocarbon components. The turbo-expander extracts useful work from the gas during expansion from a high pressure to a lower pressure. Because of the work extraction the gas is cooled more than in a simple pressure expansion, and, by means of suitable heat exchange, temperatures as low as −150°F (−100°C) can easily be achieved. The separation at low pressure gives higher relative values of the vapor–liquid equilibrium constant for nitrogen and methane than for ethane and heavier hydrocarbons. The result is that the process shown in Fig. 20.14 [11] can recover as much as 75% or more of the ethane contained in the gas as the demethanized product. In some cases the process will be run to produce a de-ethanized product that contains essentially all the propane in the feed gas.

The liquid product from the facility either can go to on-site fractionation into salable products or, more typically, be introduced into a pipeline and transported to a central fractionation facility near the point of ultimate use for the liquid product.

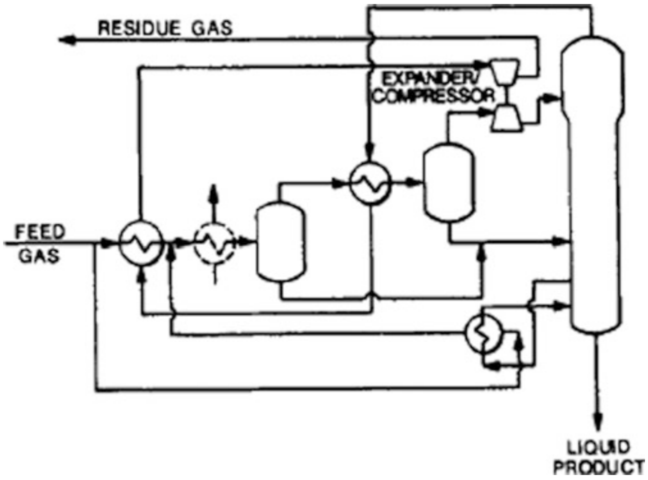


Fig. 20.14 Turbo-expander process flow

A New Potential Source for Natural Gas

Under proper conditions of temperature and pressure, water molecules can form ice-like structures in which there are cavities. If small molecules such as methane and nitrogen are present, they can be held in the cavities and there serve to stabilize the crystalline structure. The resulting crystalline structures are called clathrates [16, 17]. Although each crystalline form has a unique composition, there are no chemical bonds per se. Thermodynamically the structures are solid solutions. Although several different hydrate structures are known, the Structure 1 (S-1) hydrate typically dominates in natural gas systems. The general formula for type S-1 hydrates is $8X{\cdot}46H_2O$, where X can be molecules of argon, krypton, xenon, nitrogen, oxygen, hydrogen sulfide, carbon dioxide, methane, ethane, and/or propane. When gas containing more than one component is involved, the hydrate typically will contain more than a single gaseous constituent.

The stability of the hydrate structure depends heavily on the size of the entrapped guest molecules. Ultimately the molecules can become so large (pentane for example) that a stable hydrate cannot form. As noted earlier, many materials can interfere and prevent hydrate formation.

In recent years deposits of methane hydrate have been discovered in all parts of the world. The methane apparently forms in the decomposition of bio mass, primarily in water-containing areas, and the stabilized hydrate accumulates.

The storage of methane as hydrates offers a potentially vast natural gas resource. As to the question of how much hydrate there is right now, there is no definitive answer. However, the worldwide amount of carbon bound in gas hydrates has been *estimated to total twice the amount of carbon to be found in all known fossil fuels originally on Earth*. Additionally, -conventional gas resources appear to be trapped beneath methane hydrate layers in ocean sediments [18].

In the United States, deposits have been confirmed on all areas of the continental shelf and under Alaskan permafrost. Similar deposits have been confirmed in many locations throughout the world. But little is known about the actual size and location of even most hydrates that are presumed to exist.

Estimates of the amount of gas sequestered in hydrates varies markedly. Today, most engineers and scientists estimate that at an absolute minimum there is approximately 100,000 tcf. They also estimate there is a maximum of 270,000,000 tcf, or more.

The numbers for the United States are as follows. The U.S. Geological Survey (USGS) released a report in 1995 evaluating the U.S. hydrate resource base. It categorized estimates by the level of certainty that they exist [19].

- At a level of certainty of 95%, resources ~112,785 tcf
- At a level of certainty of 50%, resources ~276,119 tcf
- At a level of certainty of 5%, resources ~676,110 tcf
- The mean of these estimates ~320,222 tcf

Obviously, hydrates have tremendous potential as a future source for natural gas. The concluding section of this chapter will serve as a summary of what is currently being used and of promising technology for future development.

Methane Conversion Processes

Methane conversion as treated here encompasses chemical transformation only, although in the broadest interpretation purification, compression, and liquefaction operations are sometimes included, especially in relation to methane consumption as a fuel.

Methane is the simplest, most abundant, and geographically most widely distributed hydrocarbon. It therefore receives constantly increasing attention as an alternate energy source to coal and petroleum from both the world fuels industry and from the science and engineering community to broaden its utility and enhance its transportability by energy-efficient conversion to liquid hydrocarbons and functional chemical raw materials.

Apart from combustion as fuel, the largest chemical conversion of methane is steam reforming to synthesis gas (syngas) mixtures of CO and H_2 in ratios varied to fit the particular process. Syngas is the raw material base for methanol and its many derivatives including gasoline range hydrocarbons and methyl-t-butyl ether (MTBE), Fischer–Tropsch hydrocarbons, hydroformylation (oxo reaction) products, formaldehyde, hydrogen source for ammonia, and a further wide spectrum of industrial chemicals [20]. These are regarded as "indirect" methane conversion products since methane is consumed as the raw material for the final plant products and syngas accounts for over 80% of all methane conversion to commercial products.

The industrially important direct methane conversion processes comprise oxidative coupling, reductive coupling including pyrolysis reactions, partial oxidation, halogenation and oxyhalogenation [21], and ammoxidation. Other direct conversions include alkylation, electrophilic substitution, and C–H bond activation over various complex and super acid catalysts [22]. Several of these direct conversion technologies remain to be exploited to achieve their full commercial potentials.

Indirect Conversion via Syngas

Steam reforming of methane—and low-to-middle range hydrocarbons—is widely practiced to generate syngas for a number of basic, high-volume fuel stocks, such as gasoline range and middle distillate hydrocarbons, and also basic industrial chemicals such as methanol and ammonia.

The catalytic reaction of steam with methane at elevated temperatures (300−400 + °C) over various catalysts: copper or nickel/molybdenum oxide/alumina—can be made to yield CO and H_2 in desired ratios. The generalized reaction for hydrocarbons with steam is:

$$C_nH_m + CO + H_2O \rightleftharpoons nCO + [m/2 + n]H_2 \rightleftharpoons CO_2 + H_2$$

With adjustment of the steam/methane ratio, the reactor can produce a synthesis gas with $CO/H_2 = 1/2$, the stoichiometric proportions needed for methanol production. This mixture at approximately 200 atm pressure is fed to the methanol unit where the reaction then proceeds at 350°C. Per pass conversions range from 30 to 50 over the catalyst—typically a supported copper oxide with a zinc, chromium, or manganese oxide promoter: [23]

$$CO + 2H_2 \rightleftharpoons CH_3OH$$

The methanol unit operating regime, process parameters, and sometimes the catalyst may be adjusted to yield a significant coproduction of higher mixed alcohols typically ranging from C_2 to C_6.

Methanol to gasoline units following the Mobil fixed-bed design with a ZSM-5 molecular sieve catalyst are capable of producing high-grade gasoline with an iso-paraffinic hydrocarbon front end and a methlyated aromatic hydrocarbon back end. Dilution with other source alkylate to a lower aromatic content gives a premium grade mid-90s octane gasoline.

Alternatively, high-quality gasoline range distillates are produced directly from syngas via the Fischer–Tropsch synthesis. The product from all the operating units is a mixture of paraffins and olefins having a straight chain structure and molecular weight distribution predicted by the Schulz–Flory–Anderson theory. Isomerization for improved octane rating yields superior automotive fuels. Recent reviews present a position of concern that GTL processes based in syngas intermediate unavoidably entail capital costs exceeding $25,000–30,000 per daily barrel of capacity, and in most cases are not competitive with petroleum-based liquid fuel processes. The 30+ percent consumption of originally contained energy in the methane consumed to power the syngas units—compression and heat costs—must also be reckoned a process driving energy component.

Indirect Conversion via Nonsyngas Intermediates

The oxidative conversion of methane to reactive intermediates such as methyl sulfate has been reported recently [22]. The reaction, catalyzed by Group V and VI metal complexes in concentrated sulfuric acid media, utilizes molecular oxygen (air or industrial gas source) selectively to break and substitute the C–H bond in sequential oxidative scission and sulfate ester formation steps. Methanol may then be generated via hydrolysis with subsequent sulfuric acid recovery. While not yet industrially established as a methanol source for liquid fuels manufacture, this technology provides a methane C–H bond-breaking technique that is restricted to one C–H bond only and conserves much of the CH_3 group energy content that is expended in driving the commercial syngas-based methanol process alternative.

Direct Methane Conversion to Hydrocarbons and Chemical Derivatives

The direct, one-step conversion of methane to higher hydrocarbons and chemical derivatives, after several decades, continues to be intensely researched worldwide as a possible source of liquid fuels and chemical raw materials based on natural gas—independent of petroleum and coal. For clarity and convenience the following "direct conversion" categories [20] are recognized:

- Oxidative coupling to higher hydrocarbons
- Partial oxidation to chemical derivatives

- Pyrolysis or cracking
- Other direct conversion processes

Oxidative Coupling to Higher Hydrocarbons

The catalyzed gas phase reaction of methane with oxygen, wherein a hydrogen atom is abstracted from two methane molecules which are subsequently joined or "coupled" to form ethane with the coincidental formation of gaseous water, was reported [24] in the early 1980s:

$$2CH_4 + \frac{1}{2}O_2 \rightleftharpoons C_2H_6 + H_2O$$

Immediate interest and continuing, increased research by additional investigators has shown that variations in reaction conditions (typically 500–800°C, 1–5 bar, high space velocities) in combination with homo and heterogeneous, fixed and fluid bed catalysts, enables the formation of higher alkanes and olefins reaching to C_6–C_8 chain length, although C_2–C_3 dominates. With optimal methane/oxygen ratio, reaction temperature, and catalyst parameters, further oxydehydrogenation to ethlyene can occur: [25]

$$CH_3 - CH_3 + \frac{1}{2}O_2 \rightleftharpoons CH_2 = CH_2 + H_2O$$

Methane-based commercial production of ethylene via oxidative coupling has been investigated, but to date the lower per pass conversions required for acceptable ethylene selectivities combined with purified oxygen costs make this process noncompetitive with thermal cracking of ethane from natural gas liquids.

Oxidative coupling of methane nonetheless may be basically regarded as a polymerization reaction. Catalyzed oxidative coupling reactions are in commercial use today for high polymer production, for example, conversion of 2,6-dimethylphenol to poly (phenylene oxide) with molecular oxygen. However, the lower C–H bond energy and resulting greater liability of the hydrogen atoms involved in the coupling reaction allow operation in the liquid phase at far lower temperatures (below 100°C).

Methane-coupling reaction conversions and yields less than 25% initially were—and still are—below those acceptable for commercial fuel and chemical feedstock production. But worldwide research and development in more recent years continue to suggest that variations in process parameters, reactor design, and catalyst composition and structure may bridge this gap [26]. Lower reaction temperatures—in the 300–400°C range may be needed. Yields of 90% coupled with hydrocarbon products at 10% methane conversion have been reported at temperatures below 600°C where the activation energy for ethane and ethylene formation is below 90 kcal [27].

The major obstacle to high-yield oxidative methane coupling remains the high energy necessary to dissociate the C–H bond with its 94 kcal bonding energy. The rigorous reaction conditions and temperatures required to accomplish C–H bond scission in the methane molecule are essentially the same as those favoring further oxidation and degradation of the methyl radical or related short-lived highly reactive intermediates to waste products such as CO and CO_2. Research in recent years has identified catalysts, particularly including some based on transition metals, such as iridium and palladium complexes, which are able to dissociate the C–H bond at much lower temperatures, even as low as below 200°C.

Industry experts today suggest conversions of 40–50% and selectivities above 80% based on methane and oxygen as the minimum needed for commercial consideration after fixed and variable costs are added. Nonetheless, methane oxidative coupling holds the most promising combination of process simplicity, product slate versatility and low cost, and worldwide raw material availability not offered now by practiced fuel and chemical feedstock technologies.

An important aspect of direct methane oxidative coupling that merits further attention is optimization of the process to focus on the simplest products of the highest value and which are the easiest to obtain, namely the C-2 and C-3 products. Of these, the C-3 products, now commercially derived from light hydrocarbon and naphtha cracking or refinery streams, propane and particularly propylene, are in increasingly high demand and offer a versatility of existing end uses and potential utility not matched by the higher range C-4 and up aliphatic hydrocarbons, namely

1. Petrochemical raw materials for a multiplicity of plastics, films, fibers, elastomers, and chemical industry intermediates.
2. Specialty (LPG) and commodity automotive fuels–the research octane rating of 115 for propane compels industry interest.
3. High energy content, low vapor pressure (200 psi), and liquid state at ambient temperatures favors low-cost liquid pipeline transportation vs. high-pressure compressed gas—1,000+ psig if methane conversion is done near the production site.

Methane oxidative coupling process simplifications and advances—especially combinations of reactor design and catalyst features, product separation, and recycle economies—are areas marked for technological innovation with accompanying economic opportunity. C-3's handling and transportation advantages compared with remote natural gas, their potential as alternate fuel, and their demand and versatility as chemical feedstocks compared with higher liquid homologues should drive intensive methane-to-C-3's research and development for the next decade. Significant areas of exploration and development [28] will focus on:

1. Reactor design—fixed vs. fluid/moving bed variations.
2. Catalyst discovery research—metal oxides and supports, shape selective and hetero metal substituted molecular sieves, pillared clays, biomimetic, methanotropic and other bio systems; and combinatorial catalytic screening techniques, liquid phase homogeneous systems.
3. Reactor–catalyst combinations with emphasis on porous substrates.
4. Unique absorption–desorption product recovery systems.

Partial Oxidation to Chemical Derivatives

Oxygen-deficient cool flame partial oxidation of methane to methanol and/or formaldehyde has long been known [29], but methane-based selectivities of 71% methanol and 14% formaldehyde at 2% conversions require a recycle ratio of 200 to 1. Higher conversion levels led to flame temperature increase with markedly lower yields. Constraining factors were controlling the reaction stream adiabatic temperature rise within narrow ranges (25°C) to minimize waste product formation, and desirable narrow tube reactor design which raises the reactor pressure to undesirable levels—above 60 bar—and also lowers yields. At 5% conversion, the adiabatic temperature increase was near 250°C.

An alternative approach to partial oxidation of methane is oxyhydrohalogenation—usually oxyhydrochlorination [30]. In this alternative chemistry modification, methane is converted first to methyl chloride in a reactor fed a mixture of HCl, steam, and methane. The intermediately formed methyl chloride then is converted in a second catalytic (zeolite) reactor to low-to-medium range hydrocarbons. The net reaction is:

$$\mathrm{nCH_4 + 2HCl + nO_2 \rightleftharpoons C_nH_{(2n+2)} + Cl_2 + nH_2O}$$

In this process, the strategy of selective chlorination of methane to intermediate methyl chloride in effect provides for much lower temperatures and milder C–H bond-breaking energetic regimes than those so far reported in the higher temperature direct oxidative processes. The lower-temperature, halogen-based stoichiometry further sidesteps the problematic degradative oxidation of the extremely fragile methyl radical and derived intermediate species to waste products (CO and CO_2) experienced in most reported direct oxidative coupling work. This innovative approach has received careful attention as an alternative to the other "direct" oxidative coupling processes. Conversely, the handling, recovery, and recycling of chlorine and hydrogen add complexity and possible cost to the technology.

Pyrolysis or Cracking

Acetylene production via high-temperature gas phase pyrolysis is the method of choice for the production of acetylene in large volume usage. It is essentially confined to on-site dedicated units to avoid the expense and hazards of transporting acetylene long distances by pipeline, truck, or rail. In the classic Wulff process and variations thereon, pure methane is fed to a pipe pyrolysis reactor operating slightly above atmospheric pressure—not more that a few bars—at residence times of a few seconds and temperatures from near 1,000°C to much higher. The reactor effluent after heat exchanger cooling is passed to an absorber where acetylene is solvent scrubbed from the reactor stream comprising the major product hydrogen and unconverted/methane. Following hydrogen separation, methane is recycled. Small amounts of carbon may carry through to the absorber where they are removed from the solvent raffinate after acetylene recovery.

Methane high-temperature, in situ cracking to carbon and easily disposed of gaseous by products is practiced in the carbon/graphite composites industry where additional layerings of amorphous or microstructured carbon deposits are fabricated on preformed scaffolds. These are utilized primarily in the aerospace industry. Thin diamond coatings are also fabricated by the low-vacuum, in situ chemical vapor deposition from methane decomposition.

Other Direct Conversion Processes

Chlorination of methane to methyl chloride, methylene chloride, chloroform, and carbon tetrachloride is practiced industrially worldwide on a large scale. Bromination and iodination are similarly practiced on a smaller scale to prepare the corresponding bromine and iodine derivatives which are important pharmaceutical, synthetic, and fireproofing raw materials.

Methane Ammoxidation

Hydrogen cyanide (HCN) is a widely used high-volume raw and intermediate material in the world chemical industry. For the last several decades it has been increasingly supplied as a by-product of propylene ammoxidation—by which acrylonitrile, a major chemical, plastics fiber, and rubber ingredient, is made. However, transportation of liquid HCN (boiling point 26°C) over significant distances is considered hazardous because of its high toxicity and volatility and on-site production for immediate use is usually desirable. Thus, methane catalytic oxidation with oxygen (air) in the presence of ammonia, or "ammoxidation," the Andrussow process [31], is the long-established most practiced industrial technology for HCN manufacture. The singlestep reaction is straightforward and both conversions and yields are high:

$$CH_4 + NH_3 + \frac{11}{2}O_2 \rightleftharpoons HCN + 3H_2O$$

$$CH_4 + NH_3 + 1\frac{1}{2}O_2 \rightleftharpoons HCN + 3H_2O$$

As can be seen from the above equation, formation of HCN is in reality a *heterobimolecular oxidative coupling* reaction of methane with ammonia. The ammoxidation reactor construction is a simple fixed-bed multitube and the catalyst is usually a platinum or sometimes a Group V or VI metal oxide on a silica or alumina support. The HCN product is recovered by condensation and fractionation. With the reaction simplicity and yield, and widespread availability of starting materials, in situ HCN generation is an ideal industry solution to HCN supply. (See Chap. 29 for more details).

Of coincidental interest, ammoxidation of propylene [32], itself an oxidative coupling product of methane as noted above, is the commercially practiced route to acrylonitrile, which is produced in high conversions and yields:

$$CH_2 = CH - CH_3 + 1\frac{1}{2}O_2 \rightleftharpoons CH_2 = CH - CN + 3H_2O$$

Acrylonitrile is a commodity monomer raw material for many fibers, plastics, rubbers, and chemical intermediates. Economically recoverable quantities of by-product HCN may be produced which are purchased by industrial consumers or, as necessary, disposed of by incineration.

Methane Reductive Nitrilization

Vapor phase co-pyrolysis of methane–ammonia mixtures at temperatures in the 1,200–1,400°C range at modest pressures, the BMA Process [32], is also an industrial source of HCN and is carried out in several international locations where raw material costs are favorable. Another methane reductive nitrilization technology of considerable interest that has been carried out in several facilities internationally, is the Shawinigan fluohmic reactor process. Mixed methane and ammonia gases are reductively co-pyrolized to HCN and hydrogen in a fluidized bed of carbon particles maintained at pyrolytic temperatures in the 1,400°C range. Propane is reported to be a preferred feedstock [33], but methane is sometimes the economic choice. The carbon particle bed temperature is achieved by the transmission of low-voltage electric alternating current from immersed electrodes through and between carbon particles with accompanying interparticle electric arcing. Reaction stoichiometry differs from the Andrussow given the absence of oxygen in the reactant stream, and of water in the reactor effluent stream.

Space restrictions do not allow a complete survey of the countless methane conversion examples available or in the published literature. This overview is intended to highlight those methane conversion technologies which are or have been used industrially and might impact conversion of methane from hydrate deposits.

The authors are firm in their belief that methane conversion(s) hold much greater promise for the future exploitation of huge hydrate reserves than does GTL technology.

A Different View of the Origins of Natural Gas

As an essential transportation and heating fuel, and principal raw material for many process industries the future supply of natural gas is an always critical issue. Because it is more widely geographically distributed than petroleum, maintaining an adequate supply at stable pricing is a national and world economic and societal concern. In concluding this section the authors feel an overview of the status and future outlook for the natural gas supply from the standpoint of methane and light hydrocarbon origins and sourcing is appropriate. Simultaneously, much industry attention is focusing on methane and natural gas hydrates and affordable recovery and production technologies.

The two prevailing concepts on the origin and sources of natural gas and methane, its major constituent, are based on either biological or nonbiological reactions and mechanisms. Both address petroleum and, coincidentally, natural gas because the methane and other light hydrocarbons in natural gas are also constituents of most petroleum as produced. The mounting evidence for nonbiological methane and natural gas is briefly discussed below.

Natural Gas from Biological Origins

The traditional biological concept, still by far the most widely discussed and accepted across the petroleum and gas industries, is based on the widely studied biologically assisted decay and transformation of extensive deposits of vegetable-derived carbohydrate matter—trees, shrubs, etc.—into petroleum hydrocarbons. Deposits of such matter built up over millions of years were captured and compressed at hundreds of kilobars pressure under continuously building sedimentary rock strata in the earth's crust. The theory reconciles how the limestone or sandstone rock burden and at-depth geothermal temperatures combined to provide a favorable environment for indigenous bacterial strains. It then explains how their known metabolic capabilities on carbohydrate substrates accomplished scission of carbohydrate structures, with plausible reaction paths for subsequent recombinant synthesis of hydrocarbons yielding, in combination with subjection to pressure and temperature extremes during geological processes, a variable product plate of kerogen, petroleum, and natural gas, commonly known as "fossil fuels." The concept of a geological region deeper in the earth's crust, i.e., 10–30 km, well below the deepest well bores and where hardy extremophilic bacteria might survive and biocatalyze conversion of methane to higher hydrocarbons has been proposed by Gold [34]. Bacteria from very deep sea volcanic vents which thrive at 130°C and 100 bars pressure are well known. Such conditions exist at ocean depths around the world. They metabolize carbon nitrogen-, oxygen, and sulfur-containing aquatic life to degraded metabolites, and some convert vented mercaptans and H_2S to elemental sulfur and degraded products. Searches are underway to identify strains and higher temperature functionality limits that might oxidatively couple—and thus polymerize—methane to natural gas and petroleum range hydrocarbons.

Natural Gas from Nonbiological Origins

The more recently proposed nonbiological natural gas origin concepts expertly presented and summarized by Gold [34], are based on recognizing that natural gas and petroleum building blocks were constituents of, or derived from, the primordial materials and processes that originally formed planet Earth. Primordial matter and dust in the pre-Earth planetary orbit consisted of H_2 (the most abundant element in the universe) Si, O_2, C, CH_4, C_2H_6, S, and substantial quantities of metals such as Al, Na, K, Ca, Fe, Ni, V, Cr, Mo, radioactive U, Th, and others. Planet Earth formation began about four billion years ago and proceeded through the gradual agglomeration of the planetary orbit materials. As earth mass increased, gravity-driven densification and compaction caused a consequential temperature rise. Metal oxide formation and reduction by hydrogen or carbon (yielding by-product H_2O) led to formation of a dense, semi-solid, mostly metal core. This was followed by the gravitational and density graded formation of the lower and upper mantles (and associated volcanism). The latter was constituted from lighter medium to metals and their oxides, silicious or nonmetal, oxides, sulfate, silicate, and carbonate species. Primordial nitrogen, methane and ethane, light petroleum liquids derived from them, and other volatiles were squeezed out of the more dense lower strata and diffused or otherwise migrated to the crustal strata above. This uppermost earth crust was constituted of basaltic, silica/silicate, sulfate, and carbonate materials via sedimentation processes yielding features including domes, faults, strati graphic traps, and other petroleum and gas reservoir structures, accompanied by magma intrusions, and over- and under-thrusts. The least dense outer crustal materials stabilized gradually into the land masses and oceans. Over time, gravity attracted and captured from the planetary orbit and outer space a gaseous atmosphere with contributions of upward migrating gases from the lower crust and mantles. Earth's depth vs. density gradient and its geotherm or thermal gradient, were established during these events.

Metal and metal oxide catalysts with this capability were reported in the 1980s. Several of these metals occur in widely distributed petroleum samples and are among the above listed as constituents of primordial Earth planetary orbit dust. In the metallic and compounded state some have the ability to catalyze hydrogenation of carbon to kerogen-like high viscosity hydrocarbons. Kerogen, a heavier petroleum-like hydrocarbon mixture occurs in tar sands and porous shales (oil shale). Thus the question of conceivable prehistoric or more recent petroleum from methane generation must be considered.

Natural Gas Release from Lower Crust and Mantle Domains

Methane, the most stable hydrocarbon, and some others, in the absence of oxidants can endure temperatures of 1,000°C and higher. With some mantle and lower crustal rock porosity, gas law behavior would require that during planet earth formation and the long gradient-densification process, most of the primordial methane component (which) escaped oxidation, should have diffused upward to be trapped in the sedimentary crustal reservoirs and likewise for ethane and other light hydrocarbons, with little loss. If this upward migration continued from mantle or lower crustal depths, ca. 300 km, depleted modern natural gas upper crustal reservoirs, given time, should substantially refill. This has been observed in numerous cases, offering convincing support for the nonbiological natural gas concept.

Of particular note is that methane, and natural gas, if from great depths, should contain He from radioactive decay of the very dense uranium and thorium in the lower crust and mantle domains. This has been documented in most of the above instances. Natural gas reservoirs producing gas containing over 5%, sometimes as much as 10%, helium are long known, and they provide this purified gas for many industrial uses and lighter-than-air transportation craft.

Methane and a few other light hydrocarbons, C_2 and C_3 are capable of forming solid clathrate compositions with water. These are literally ice-like structures with methane molecules occupying cavities formed by solid-state hydrogen-bonded H_2O molecules. A ratio of six H_2O units to one CH_4 unit forms the stable methane hydrate at temperatures below 7°C and pressures higher than 50 bars. A typical mostly methane natural gas with about 10% minor components, including ethane, other light hydrocarbons, CO_2 and N_2 forms a stable hydrate at temperatures below 10°C and pressures higher than 21 bars. Such hydrates are often nuisances (and worse) in cold-weather gas pipeline operation, but have great significance to natural gas technology and outlook. Pressure–temperature combinations spanning the above example are common at various ocean depths, and vast gas hydrate deposits exist beneath ocean floors where upwelling natural gas forms hydrate caps with subocean floor seawater and forms very large subterranean free gas reservoirs. The distribution and size of such hydrate deposits offer further persuasive evidence confirming the deep earth gas theory. Analogous in-earth deposits are believed to exist in land sites at temperatures compliant with the geotherm, pressure at depth, and compatible rock structure. Gas reserves in known ocean floor hydrate reservoirs are estimated at 10^{15} kg.

Summary for Methane and Natural Gas Future Sourcing

As for future sourcing, methane generation from biological sources such as coal beds and composted vegetation is well known, industrially exploited, and of particular recent interest as a potentially renewable source. Methane is also an important constituent of numerous petroleum grades that contain hopanoid and optically active hydrocarbons, which are associated with biological and biochemical processes.

Beginning, however, in the 1990s, compelling factors and reasons that support the origin of nonbiologically derived methane have come to worldwide attention. These are rooted in its having been a substantial component of the planetary orbit dust from which the primordial earth mass agglomerated; and that it now appears it is being continuously released from the mantle and lower crustal domains via upward diffusion or other migratory processes to shallow crustal and surface levels. Moreover, methane (with its good thermodynamic stability) and higher hydrocarbons (with somewhat less stability depending on molecular weight and structure) could, with subsequent core and mantle temperature and pressure rise, well have been, and could continue to be, formed within these core and mantle domains by the metal catalyzed combination of primordial hydrogen and carbon.

At least three established major factors strongly reinforce these nonbiological methane scenarios. First is the presence of significant amounts (occasionally up to 10 volume%) helium in widely spread natural gas production sites. Second is the observed gradual refilling of previously depleted natural gas reservoirs by upwelling. Third is the existence of vast methane/natural gas hydrate deposits below ocean floors and land masses, some of which act as caps for larger gaseous methane reservoirs beneath. These gas hydrates are able to form at temperatures above the freezing point of water and at pressures near 5,000–6,000 psi. The estimated worldwide total carbon in these natural gas formations is 10- to 100-fold greater than presently estimated worldwide petroleum reserves.

As a result of these discoveries in recent years, the search for cost-effective production and applicable extraction, drilling, or other recovery technologies is an industry and multinational priority.

Acknowledgement The authors wish to acknowledge the contribution of Dr. Robert N. Maddox, late Emeritus Professor, Oklahoma State University, who first created this chapter in an earlier edition of the Handbook.

References

1. Swindell G (2002) Personal Communication with R. N. Maddox, March 2002
2. North American Coalbed Methane Resource Map (2001) Gas Technology Institute, GTI-01/0165, Des Plaines, IL 60018
3. International Energy Outlook (2011) U.S. Energy Information Administration, DOE/EIA-0484, U.S. Department of Energy
4. U.S. Crude Oil, Natural Gas and Natural Gas Liquids Reserves. Energy Information Administration, U.S. Department of Energy
5. Sales of Natural Gas Liquids and Liquefied Refinery Gases. American Petroleum Institute, Gas Processors' Association, and National Propane Gas Association
6. British Petroleum Company. Natural Gas Reserves (2011) BP Website: http://www.bp.com/ 2011
7. Samsam Bakhtiari AM (2001) Gas-to-liquids: much smoke, little fire. Hydrocarbon Processing
8. Romanovw S (2001) Got gas. Hydrocarbon Processing
9. Campbell JM (1984) Gas conditioning and processing, vol. 1—the basic principles, 6th edn. Campbell Petroleum Series, Norman, OK
10. Campbell JM (1984) Gas conditioning and processing, vol. 2—the equipment modules, 6th edn. Campbell Petroleum Series, Norman, OK
11. Maddox RN, Lilly LL (1990) Gas conditioning and processing, vol. 3—computer applications for production/processing facilities, 2nd edn. Campbell Petroleum Series, Norman, OK
12. Segeler CG (ed) (1977) Gas Engineer's Handbook, 5th edn., 5th printing. Industrial Press, New York
13. (2004) GPSA Engineering Data Book, 12th edn. Gas Processors Suppliers Association, Tulsa, OK

14. Maddox RN (1985) Gas conditioning and processing, vol. 4—gas and liquid sweetening, 3rd edn. Campbell Petroleum Series, Norman, OK
15. Huntington RL (1950) Natural gas and natural gasoline. New York, McGraw-Hill
16. Jeffrey GA, McMullan RR (1967) The Clathrate Hydrates, Prog Inorg Chem 8:43
17. Ripmeester JA, Tse JS, Ratcliffe CI, Powell BM (1987) A new clathrate hydrate structure. Nature 325:135
18. Dillon W (1992) Gas (methane) hydrates—a new frontier. U.S. Geological Survey Fact Sheets
19. Collett T, Kuuskraa V (1998) Hydrates contain vast store of world gas resources. Oil Gas 96:90–95
20. Kuo JCW, Ketlar AD (1987) Evaluation of direct methane conversion to higher hydrocarbons and oxygenates. Final Report DOE Contract DE–AC22–86PC90009
21. Humphreys GC, Dunster M (1979) Engineering the World's Largest Methanol Plants. Design '79 Conference, Birmingham
22. Olah GA (1987) Electrophilic methane conversion. Acc Chem Res 20:422–428
23. Gas Processors Suppliers Association Engineering Data Book, 12th edn. Electronic FPS Version
24. Keller GE, Bhasin MM (1982) Catalytic conversion of methane by oxidative coupling to $C2^+$ hydrocarbons. J Catal 73:9
25. Pan D, Ji S, Wang W, Li C (2010) Oxidative coupling of methane in a dual bed reactor. J Nat Gas Chem V19, Issue 6, pp 600–604
26. Li Y, Dickenson AG, Moser WR, Ma YH, Balachandran U (2000) Oxygen-permeable dense membrane reactor for oxidative coupling of methane. J Membr Sci V170:27–34
27. Rasko J, Peirera P, Somorjai GA, Heinemann H (1991) Catal Lett 9
28. Yaghobi N, Ghorelshy MHR (2008) Comparative study on kinetic modeling of oxidative coupling of methane in laboratory and bench scales. Chem Prod Proc Model V3(1):16
29. Chemische Werke Huels AG (1975) Process for the Manufacture of Mixtures of Formaldehyde and Methanol by Partial Oxidation of Methane. GB Patent 1.410,709 (Oct. 22, 1975)
30. Fox JM, Chen TP, Degen BD (1988) Direct Methane Conversion Process Evaluation. Final Report, U.S. Department of Energy Contract DE-AC22-87PC79814, August 1988
31. Andrussow, U.S. Patent 1,934,839
32. Idol JD (1959) U.S. Patent 2,904,580, Sept. 15, 1959
33. Encyclopedia of chemical technology (2002) Kirk–Othmer, 6th edn, vol 7. pp 756–775
34. Gold T (2001) The deep hot biosphere. New York, Copernicus
35. Crockett LO, Widgery RC (1955) Natural Gasoline Producers' Stake in Chemical Manufacturing. Proceedings 34th Annual Convention, Gas Processors' Association
36. (2004) Gas Processors Suppliers Association Engineering Data Book, 12th edn. The GPSA offices are at 6526 East 60th Street, Tulsa, OK 74145
37. Gas Processors Association Standard 2145. Same address as GPSA
38. Energy Information Administration (2000) U.S. Crude Oil, Natural Gas and Natural Gas Liquids Reserves. Annual Report
39. Wender I, Klier K (1989) Review of indirect liquefaction. In: Schindler HD (ed) Chapter 5, coal liquefaction—a research needs assessment. DOE Contract DE–AC01-87ER30110
40. Prensky S (1995) A review of gas hydrates and formation evaluation of hydrate-bearing reservoirs. Society of Professional Well Log Analysts 36th Annual Logging Symposium Transactions, paper GGG
41. Trofimuk AA, Cheraski NV, Tsaryov VP (1977) The role of continental glaciation and hydrate formation on petroleum occurrence. In: Meyer RF (ed) The future supply of nature-made petroleum and gas. Pergamon, New York, pp 919–926

The Nuclear Industry

21

Tom Congedo, Edward Lahoda, Regis Matzie, and Keith Task

Introduction

The objective of the nuclear industry is to produce energy in the forms of heat from either fission reactions or radioactive decay and radiation from radioactive decay or by accelerator methods. For fission heat applications, the nuclear fuel has a very high specific energy content that currently has two principal uses, for military explosives and for electricity generation. As higher temperature reactors become more widely available, the high temperature heat (>900°C) will also be useful for making chemicals such as hydrogen. For radiation applications, the emissions from radioactive decay of unstable nuclides are employed in research, medicine, and industry for diagnostic purposes and for chemical reaction initiation. Radioactive decay heat is also employed to generate electricity from thermoelectric generators for low-power applications in space or remote terrestrial locations. Radiation produced from accelerator-based sources is used for geologic investigation (e.g., identifying oil deposits), materials modification, and contrast imaging of dense media (e.g., security inspections in commercial shipping). Fuel from the first atomic pile is shown in Fig. 21.1.

This nuclear technology is based on both the nuclear and the chemical properties of the atom. At the beginning of the twentieth century fewer than 90 chemical elements were known and there was only a dawning awareness of isotopes. Today, largely because of the nuclear industry, thousands of isotopes (or nuclides, depending on the properties of interest) have been identified. Brief definitions of several chemical and nuclear terms are given in Table 21.1.

As with other technology, nuclear technology involves a combination of science and art. However, it is unique because of the development of the atomic bomb that contributed to the ending of World War II. Many people view nuclear technology from the point of view of nuclear weapons and more recently nuclear accidents such as those at Chernobyl and Three Mile Island. This leads to the view that nuclear technology is only useful for explosive applications and that it is only with great care that it can be safely used. In reality, it is difficult to produce nuclear explosions and safe use of nuclear energy is really only a matter of following common-sense rules of behavior no more mysterious than those involved with handling dangerous chemicals or electricity.

This chapter first considers the following subjects.

- The status and outlook of the nuclear industry.
- Nuclear safety.
- The role of nuclear energy in meeting the world's energy needs.
- The nuclear processes on which the nuclear industry is based (Fig. 21.2a).
- The chemical process technologies involved in the fuel cycle for nuclear electric power and chemical process heat generation (Fig. 21.2b).

This is followed by discussions of radioactive waste management and transportation of nuclear materials, and consideration of various applications, such as nuclear power reactors used for the generation of electric power, the use of radioisotopes, and other military and civilian uses of nuclear materials (Fig. 21.2c). The processing of uranium ore, the enrichment of material for use as nuclear fuels, the production of electricity using nuclear fuel and the subsequent handling, storage, and in some cases reprocessing of the radioactively contaminated waste constitute some of the more sophisticated and challenging areas of fundamental and applied chemistry and engineering as well as requiring huge investments on the order of tens of billions of dollars for each stage in the process.

A reference section is provided for those who wish to have more detailed information.

T. Congedo • E. Lahoda (✉) • R. Matzie • K. Task
Westinghouse Electric LLC, Hartford, TX, USA
e-mail: lahodaej@westinghouse.com

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_21, © Springer Science+Business Media New York 2012

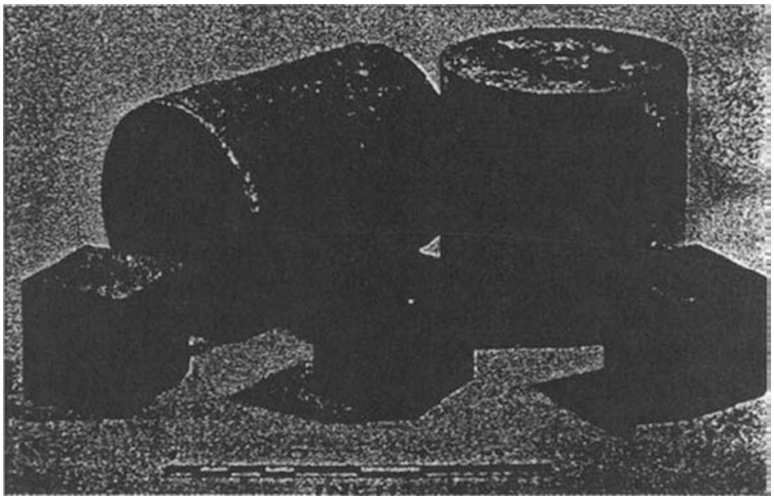


Fig. 21.1 Uranium metal fuel from the first atomic pile. This reactor contained 40 t of uranium oxide along with 6.2 t of uranium metal (ORNL News, 1-01-076)

Table 21.1 Definitions of atoms, chemical elements, isotopes, nuclides, and isomers

Atoms: Elementary particles of matter composed of protons, neutrons, and electrons. In a neutral atom the total number of orbital electrons is equal to the number of protons in the nucleus
Chemical elements: Atoms with unique properties related to their orbital electrons
Fissile: Capable of undergoing nuclear fission initiated by a slow neutron
Fissionable: Requiring a neutron of kinetic energy above a threshold value, to initiate nuclear fission
Isotopes: Atoms of the same chemical element with different masses related to a different number of neutrons in the nucleus
Nuclides: Atoms with unique properties related to the neutrons and protons in the nucleus of the atom
Isomers: Nuclides with the same numbers of protons and neutron but in different energy states
Radioactive decay: The process by which unstable nuclei become more stable
Radioactive half-life: The time during which the decay rate of a radioactive nuclide decreases by a factor of 2

Status and Outlook

Nuclear energy provides products that play vital societal roles:

- As a primary energy source for the production of electricity to meet the world's energy needs without polluting the atmosphere [1]. Currently, it is the only large-scale electrical production process that does not produce greenhouse gases (e.g., CO_2) or other pollutants such as SO_2, NO_x, mercury, or particulates.
- As stable and radioactive nuclides, which contribute significantly to research, medicine, and industry (see Table 21.2 and Fig. 21.2.).
- As radioactive decay heat sources that in conjunction with thermoelectric conversion devices provide low levels (kilowatts) of power for very long times in deep space probes and for remote terrestrial applications such as weather sensors.
- As nuclear explosives and in nuclear-powered submarines and ships, which have contributed to maintaining world peace since World War II, but have since declined in importance.
- Test reactors that are used to generate isotopes, do chemical analyses, and study materials.

Nuclear technology continues to be developed for military applications including nuclear explosives, submarines, and ships. Brief attention is given herein to explosives, primarily in terms of their historical role as a potential means for excavating harbors and canals and in the enhancement of gas and petroleum deposits.

It should also be noted that since the end of the Cold War, official military stockpiles of nuclear (and other) weapons have been diminished, by international agreement. At the same time, acquisition or fabrication of nuclear devices by nongovernmental terrorist groups or individual states has become a matter of increased concern.

Since the discovery of nuclear fission in the late 1930s, this technology has been developed to supply 14% of the world's electricity from 435 nuclear reactors located in 30 countries (2012) [2]. This is an international program regulated by the United Nations International Atomic Energy Agency (IAEA) which, as of 2012, includes 155 member countries. In 2012 there 104 nuclear units in the United States, including 69 pressurized water reactors (PWRs) and 35 boiling water reactors, which together generated 790 billion kWh of electricity, some 20% of the country's total electricity generation [2]. This represented a 217% increase in output since 1980 (see Fig. 21.3).

Fig. 21.2 The nuclear industry

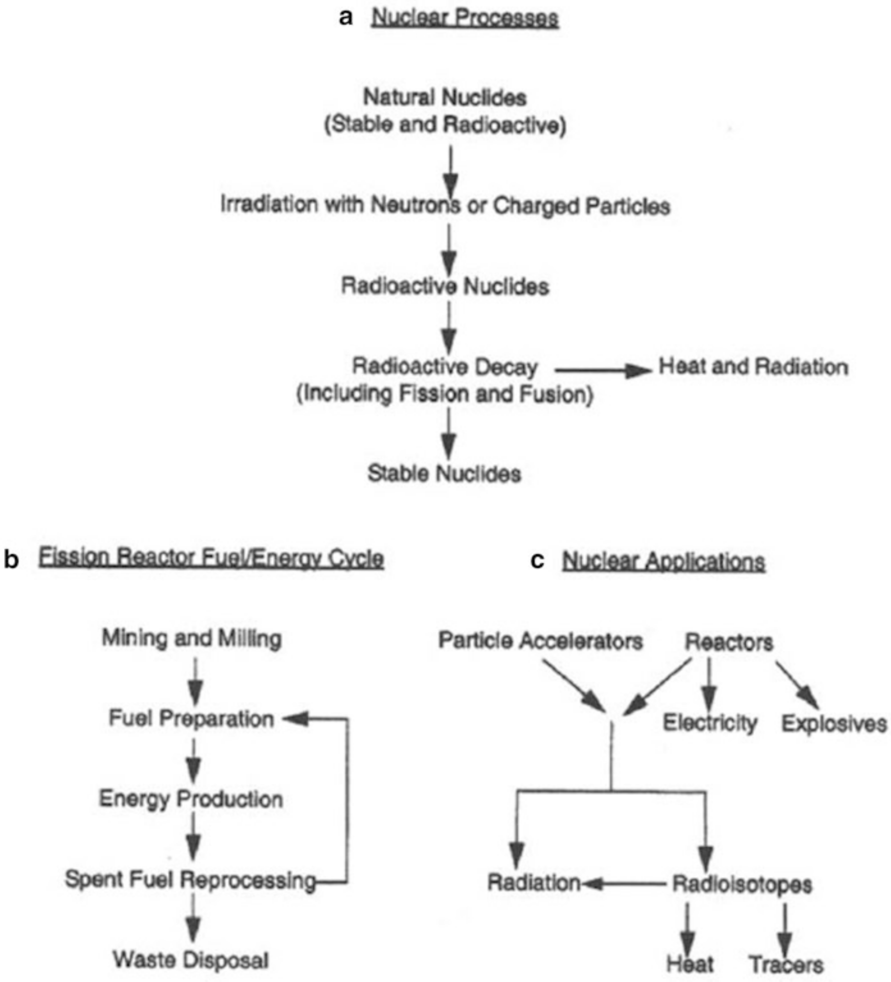


Table 21.2 Nuclear Industry Products (2012)

Civilian	Approximate number
Worldwide electric generating plants	435–433
Ships (icebreakers and transports)	~10
Research reactors	240
Radioisotope applications (tracers, radiation sources, thermal sources)	Millions
Radiation modification of materials (solid-state devices, treated polymers, etc.)	Small (mostly proprietary)
Military	
Weapons	20,000[a]
Submarines, ships (aircraft carriers, cruisers, transports)	~140

[a] Although it is a possibility that up to 10,000 more nuclear warheads may be awaiting dismantling or are in reserve in Russia [3]

The safety record of the nuclear industry has continued to be generally very good, despite the 1987 accident with a nuclear reactor at Chernobyl, in Ukraine, which killed a number of workers, caused the abandonment of adjacent communities and farmland, and resulted in radioactive fallout in neighboring countries. It should be noted that these were the first civilian nuclear fatalities since the beginning of the nuclear power industry [4]. It is generally agreed that this unfortunate occurrence was caused by serious design flaws, and that power reactors in use elsewhere are not subject to similar occurrences. Indeed, a somewhat similar yet also very different accident occurred in the United States in 1979 (Three Mile Island); but because of the reactor's inherent physics feedback mechanisms and its engineered safety features (required in the United States and recommended by the IAEA), there were no injuries and no significant radiation exposure, either to workers or the public [5]. More recently (2011) a tsunami triggered by a major earthquake

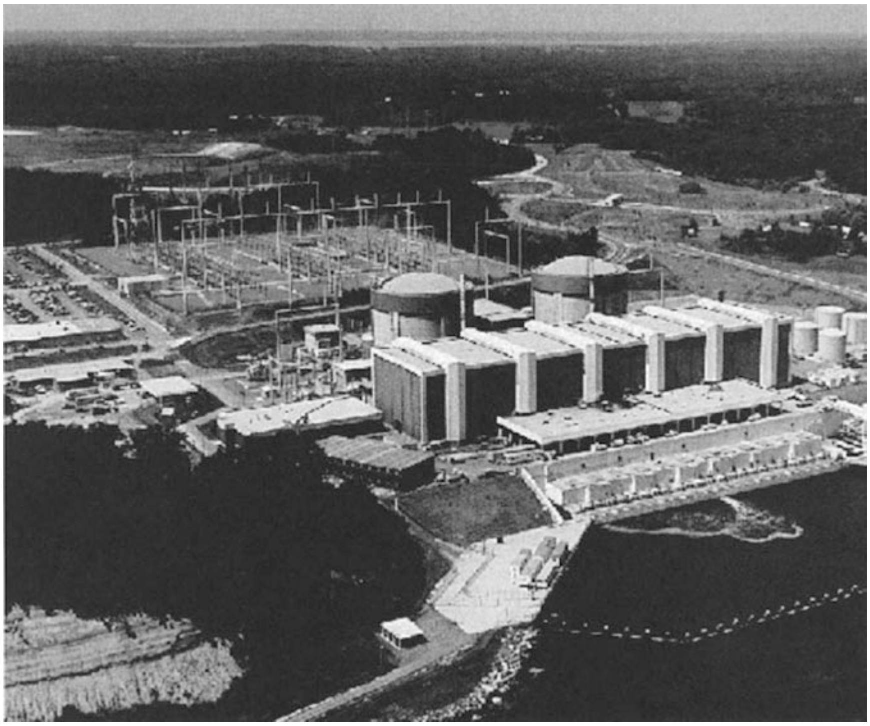


Fig. 21.3 The Calvert Cliffs Plant in Lusby, MD, Baltimore Gas and Electric. Capacity: 1,600 MW(e) from two reactors (Courtesy of EEI-Electric Perspective)

off the coast of Japan disabled the emergency gererators that powered the longterm cooling systems at the Fukushima Nuclear Poner Plant in Japan. Three reactors and some of the spent fuel pools suffered meltdowns and released significant amounts of radioactive material which contaminated large areas of the surrounding country side. Two workers died from non-radiation covles and no willian casualties occured.

Worldwide, aggressive public concern about nuclear safety has been delaying the development of waste disposal facilities for high-level wastes. Current intermediate-level waste facilities in the United States are nearly full, and there has been great difficulty in siting new facilities. It now appears that for many years the interim solution for the storage of spent fuel elements will be above-ground storage facilities at existing reactor sites with appropriate safeguards and security. On the energy side, increased attention to conservation and the ready availability of natural gas in the United States that can be used in low-cost combined cycle electric generating plants reduced the pressure for continued nuclear power development in the later years of the twentieth and the beginning of the twenty-first century. Supplies of natural gas continue to be plentiful and it accounted for about 23% of the energy used in the US in 2010. Due to new shake gas supplies that continue to come on line, prices will likely continue to be low for the foreseeable future, and the percentage of energy from natural gas will continue to increase, mainly from displacing coal use. Because of worldwide concerns regarding climate change and other environmental factors, there is increasing emphasis on reducing the use of all fossil fuels. Nuclear is the only non-fossil technology which is capable of being a significant and economically acceptable electrical power generating option.

Light water reactors (LWRs) and HWRs have been successfully used for electric power generation throughout the world with good operating and safety records. In recent years significant advances in nuclear reactor technology have been realized. A new generation of light and HWRs has been designed and licensed to further reduce the already very low probability of loss-of-coolant accidents such as occurred at Three Mile Island and station blackouts as occured at Fukushima and Chernobyl. In addition, the enrichment level of new fuel for LWRs has been increased, and the average fuel life has thereby been extended from 33,000 to as high as 50,000 MW days/mt, and the typical reactor cycle between shutdowns has been increased from 12 to 18 months. As a result, online times (i.e., capacity factors) have increased to the low 90 percents, providing low-cost as well as reliable electrical power generation.

All the components of the nuclear fission power system are fully operational except for ultimate waste disposal. However, spent fuel is not reprocessed in the United States because there is currently an adequate supply of natural uranium and enrichment services available domestically and from other countries at a lower cost than that of the recovered fissionable material from spent fuel. Also, the United States unilaterally declared a moratorium on reprocessing in the early 1980s in an attempt to reduce the spread of nuclear weapons. Current economics do not favor a return to reprocessing and fuel recycling in the United States it for the foreseeable future.

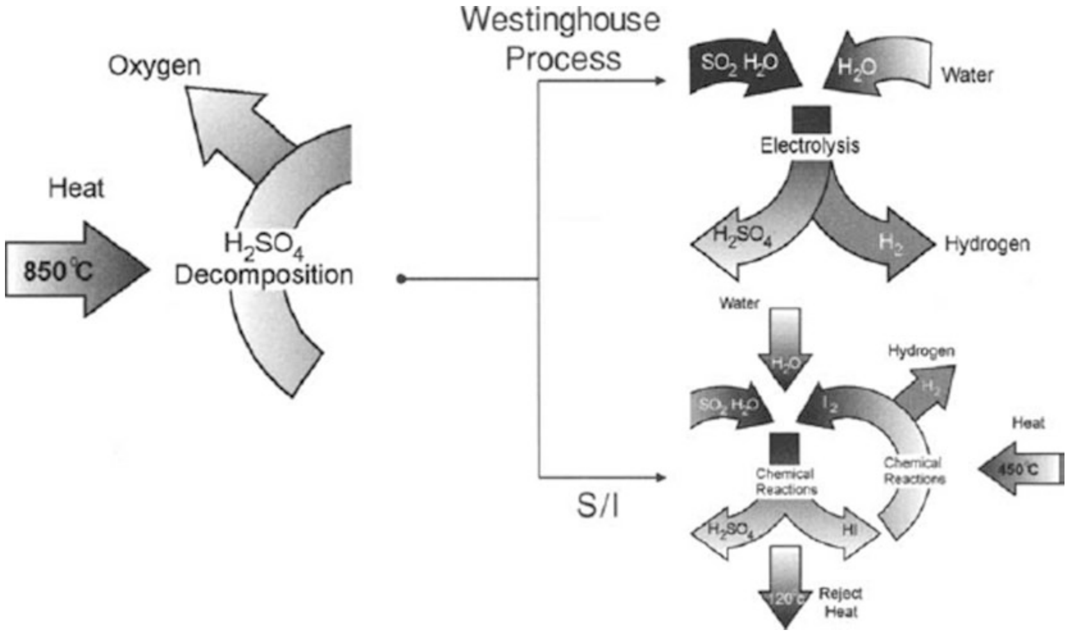


Fig. 21.4 High-temperature processes for generating hydrogen (Nuclear Hydrogen R&D Plan. Final Draft, U.S. Department of Energy, March 2004)

The nuclear industry makes available about 3,000 nuclides, including both the stable and the radioactive nuclides. Approximately 50 radioactive nuclides, along with some stable nuclides that have been isotopically enriched, are essential in research, medical, and industrial applications. Many of these are now produced commercially, but several still are dependent on government facilities. Some, for economic reasons, come from other countries. Radiation processing for sterilization of disposable medical supplies is an important operation using cobalt-60 from Canada. Electron accelerators have replaced radioactive nuclides as radiation sources for polymerizing plastic coatings on wire and paper. Perhaps the greatest disappointment experienced by the industry has been the public opposition to irradiation of food to extend its shelf life. The principal current application of this type is by the Russians, who are using electron accelerators for the deinfestation of wheat, and the use in the United States for treating herbs and spices.

In the continuing research and development of advanced nuclear energy, the liquid metal reactor (LMR), molten salt reactor (MSR), high-temperature gas-cooled reactor (HTGR), and fusion are the major activities. It would now appear that these alternatives will likely increase the cost of electrical energy in comparison with the present LWR/HWR technology based on low-cost uranium ore. HTGRs have been of interest since 2003 when the United States declared its intent to replace the use of oil-based fuels for transportation with hydrogen. Several hydrogen production processes (e.g., the Westinghouse Process and the Sulfur Iodine Process; see Fig. 21.4) have been investigated that utilize high temperature (850–900°C) heat from a reactor to provide a major portion of the energy for making hydrogen. These temperatures are only compatible with HTGRs. The advantage of using these high-temperature processes is the higher overall energy efficiency that is obtainable (see Fig. 21.5). For instance, the Westinghouse Process can achieve overall efficiencies of 33% or greater.

Plant-scale liquid metal (sodium) reactors (2008) have been in operation in France and Russia to establish the engineering technology and to evaluate their role in increasing the efficiency of uranium utilization. France, without a significant primary energy supply, wants to minimize imports of uranium. According to the USEIA web site, the United States imported over 40% of its uranium. In the United States, engineering test LMRs operated for many years and made significant contributions to LMR technology.

Whereas LMRs were first studied to reduce natural uranium requirements to approximately 20 mt/GW year electrical (GWye), development is now centered on reducing plutonium in the waste. However, at this time all such LMRs have been shut down and are being decommissioned. Programs are also underway to use conventional uranium dioxide fuel containing "blended down" highly enriched uranium (HEU) or mixed oxide (MOX) fuel containing oxides of uranium and plutonium in commercial nuclear power reactors in order to dispose of nuclear weapons materials. Other theoretical reactor concepts are also being investigated for disposal of actinides.

It is difficult to assess the progress of fusion research, a significant international research program. Russia, the European Community, China, Japan, and the United States are the principal participants. It is now clear that fusion will be a high-

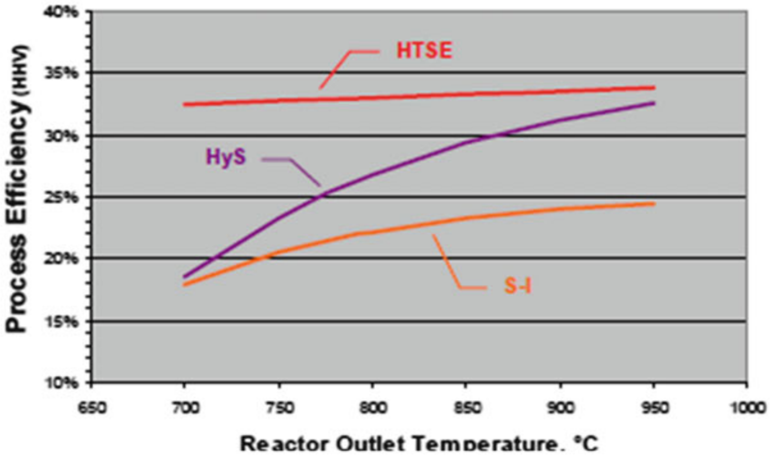


Fig. 21.5 Efficiencies of various H_2 production processes (Charles Bolthrunis et al. (March, 2009) NGNP hydrogen plant alternatives study, NGNP HPS SHAW-HPA, pp. 5–28)

cost source of energy and not economically justified within the foreseeable future. However, it is essential for the major world nations to cooperate responsibly in its development because of the severe technical challenges and high cost of research and development of this technology. The basic research has yet to achieve controlled ignition of the fusion reaction along with energy parity (energy input less than output).

At this time (2012), the outlook for the nuclear industry is difficult to project. France and Korea, countries without domestic energy resources, are continuing to move ahead to establish efficient nuclear systems for electric generation; and although they have a concerned public, their fears regarding waste disposal and proliferation are not overriding their need for electricity. China has embarked on an aggressive nuclear build program, but its projected electricity demand is growing so rapidly that this aggressive program will only result in about 5% of electricity from nuclear in the next 15 years. In Japan and Germany, countries which currently produce a significant amount of electrical power from nuclear, there is movement to eliminate the Nuclear power option.

Recently, a number of countries which had no previous reactors or new construction for a long period of time, have embarked on building new nuclear plants. Included in this list are countries new to nuclear power such as the UAE, as well as established nuclear users such as the US, Canada, UK, and Finland.

Nuclear Safety

Safety first has always been and continues to be the basic policy of the nuclear industry. This includes reactor safety by design as well as activities to discourage the proliferation of nuclear weapons and to prevent sabotage of nuclear facilities. This policy has been successful; the chance of death from a nuclear accident is over a million times less than death from ordinary human activities, and over a thousand times less than death from natural events (see Table 21.3). Safety remains the most challenging responsibility of the nuclear industry.

It is the public's strong perception that all nuclear activities are more dangerous than other accepted risks. Much of this public concern results from the atomic bombs and the government's secret program that produced the bombs. However, even with the early large-scale nuclear operations, there were few nuclear fatalities or life-shortening injuries. After 40 or more years of operation, these facilities were due to be retired; however, there is a strong movement now to extend the life of most operating reactors for another 20 years or more. Meanwhile, considerable operating and safety experience has been gained that may very significantly reduce public concern in the future.

The nuclear safety program originally was based on both established laws regulating all industrial safety and early knowledge of radiation health effects related to radium and X-ray exposure. Since then the safety regulations have been greatly enhanced, and they are present in the Code of Federal Regulations, Title 10, Chapter 1, titled "Nuclear Regulatory Commission" [6]. These regulations strictly control the management, engineering design, and operations of all nuclear activities. In addition the IAEA already detained previouly has established regulations as a reference for all national programs. Each country also has its own regulatory agency. There are two areas of concern in nuclear safety: radiation exposure of the public and of workers resulting from normal and accident conditions; and danger to world peace and order from nuclear weapons proliferation and sabotage.

Table 21.3 Average risk of fatality by various causes in the United States

Accident type	Annual total number*	Individual annual risk**
Total Transportation Accidents	45,832	1 in 4,429
Poisoning	24,313	1 in 8,349
Falls	22,736	1 in 8,928
Homicides	17,520	1 in 11,586
Fires	3,276	1 in 61,965
Drowning	3,237	1 in 62,712
Accidential Discharge of Firearms	721	1 in 281,553
Average Lightning Deaths***	58	1 in 350,000
Average Tornado Deaths***	57	1 in 3,561,403
Average Huricane Deaths***	48	1 in 4,229,166
Nuclear reactors****	–	1 in 357,615,8940

*2007 Statistics from the Centers for Disease Control http://www.cdc.gov/nchs/fastats/injury.htm

National Vital Statistical Reports, Volume 58, Number 1 August 19, 2009, Deaths: Preliminary Data for 2007

Table 2. Deaths, death rates, and age-adjusted death rates for 113 selected causes: UnitedStates, final 2006 and preliminary 2007

**US 2007 Population from Population Reference Bureau http://www.prb.org/pdf07/07WPDS_Eng.pdf

***National Weather Service – Lightning Safety http://www.lightningsafety.noaa.gov/overview.htm

****Calculation based on 100 LWRs operating under United States Nuclear Reugulatory Commission Supervision

Table 21.4 Annual estimated average effective dose equivalent received by a member of the population of the United States (From HP Society University of Michigan site, 1/18/05 [10])

	Average annual effective dose equivalent	
Source	(μSv)	(mrem)
Inhaled (radon and decay products)	2,000	200
Other internally deposited radionuclides	390	39
Terrestrial radiation	280	28
Cosmic radiation	270	27
Cosmogenic radioactivity	10	1
Rounded total from natural sources	3,000	300
Rounded total from artificial sources	600	60
Total	3,600	360

Radiation, like air, has a ubiquitous presence in the human environment. The amount of background radiation that humans are exposed to ranges from 100 to 300 mrem/year depending on their location in the world, and an additional 60 mrem/year comes from other normal sources (medical, consumer products, etc.). The principal natural radiation sources are cosmic rays from outer space, radon from geologic sources, and potassium-40 within the human body (see Table 21.4). The radiation exposure to the public resulting from normal nuclear operations is limited by federal regulation of engineering design and management controls. There is also an overriding requirement that limits the radiation exposure to "as low as reasonably achievable (ALARA)."

The annual radiation exposure of nuclear workers is limited to 5 rem, and the limit for public exposure (not radiation workers) from licensed activities is 100 mrem/year although people living near a nuclear power station are on average only exposed to <1 mrem/year [8]. The government also puts limits on fatality risk levels due to a nuclear accident:

> The risk of an immediate fatality to an average individual in the vicinity of a nuclear power plant that might result from reactor accidents should not exceed 0.1% of the sum of the immediate fatality risks that result from other accidents to which the US population is generally exposed, and the risk of cancer fatalities to the population near a nuclear power plant should not exceed 0.1% of the sum of cancer fatality risks from all other causes [9].

The levels of radiation exposure were first established on the basis of historical data, and since then they have been periodically evaluated by the International Commission on the basis of continuing experience.

In 1975 the United States Energy Research and Development Agency, the forerunner of the United States Department of Energy, sponsored the "Reactor Safety Study," the most thorough hazard analysis ever performed [10] for any technology. It concluded that on the basis of past licensing review practices, the risk from existing light water type power reactors was orders of magnitude less than other commonly accepted hazards. It further concluded that human error would be the expected cause of any accidents that might occur. This study also found that the cost of nuclear power reactor accidents would be several orders of magnitude less than the cost of other natural and human-caused events to which the public is accustomed (see Fig. 21.6).

In the 1950s the Nuclear Safeguards Program was established by the United Nations to administratively reduce the risk of nuclear proliferation. The objective of the Safeguards Program is to prevent the diversion of fissionable material, primarily plutonium, from the reprocessing of spent fuel. A force of onsite inspectors is employed by IAEA in this activity. This program is administered by the IAEA and is generally considered to be successful. Countries join this program voluntarily and do not include their weapons production facilities. India, Pakistan, and North Korea did not choose to join and developed their own nuclear weapons. Two nonmember countries have in the past or are developing nuclear weapons including South Africa and Israel.

In 1980 a study of proliferation control, titled "International Nuclear Fuel Cycle Evaluation," was carried out for President Jimmy Carter of the United States. The United States had called for this study, in which 50 countries

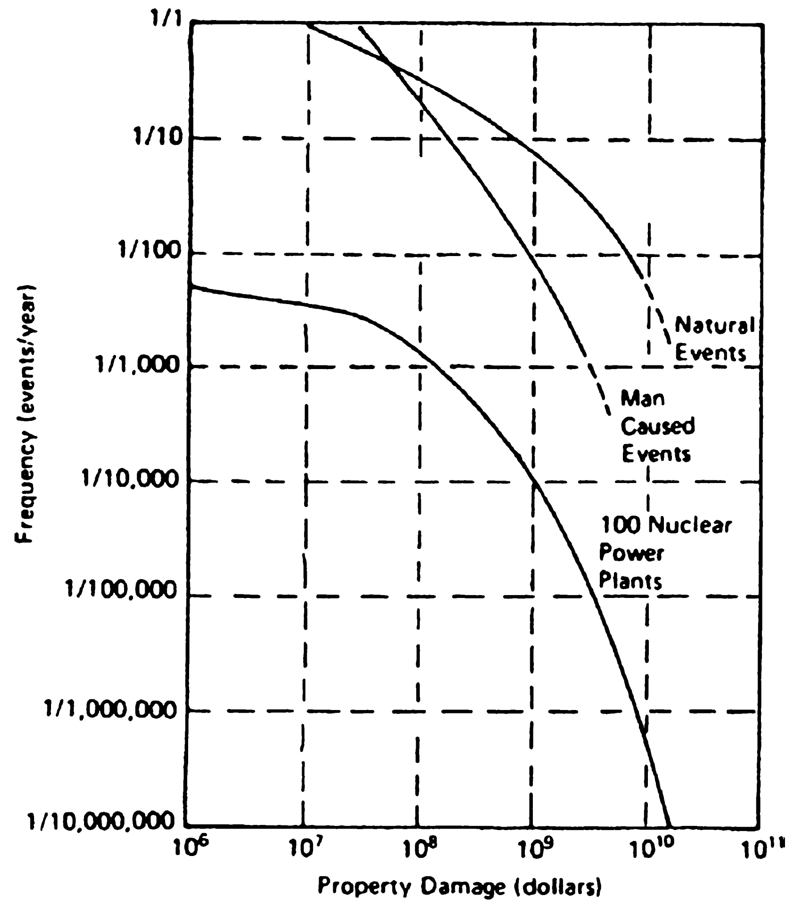


Fig. 21.6 Frequency of property damage due to natural and human-caused events (*Notes*: (1) Property damage due to auto accidents not included. (2) Approximate uncertainties for nuclear events are estimated to be represented by factors of 1/5 and 2 on consequence magnitudes and by factors of 1/5 and 5 on probabilities. (3) For natural and human-caused occurrences the uncertainty in probability of the largest recorded consequence magnitude is estimated to be represented by factors of 1/20 and 5. Smaller magnitudes have less uncertainty)

participated and which required several years to complete. The study concluded that only administrative controls such as those already in place would be effective, and these could only delay really determined nations from diverting nuclear fuel from power reactors to military weapons. By 2003, North Korea became the first example of the subversion of test/commercial nuclear materials for use in weapons.

The Earth's Energy Supply and Use

The earth is provided with two sources of energy, the "capital" resources deposited within the earth during its formation and the "income" resource continuously beamed onto the earth from the sun. Both are essential to human existence.

In the 1950s, global energy analyses were performed by King Hubbard for the US Geologic Survey and by Palmer Putnam for the Atomic Energy Commission. Hubbard's work reported the income energy from the sun to be equivalent to 178,000 million MW (Fig. 21.7). At noon on a cloudless day a square meter of the earth's surface facing the sun receives approximately 1 kW. The income resource also includes a very small contribution, less than 0.02%, from the gravitational forces of the sun, moon, and earth and from the thermal and nuclear sources within the earth. Part of the solar energy appears in the form of water and wind power, wood, alcohol, garbage, cow dung, solar heaters, and photoelectric generators.

Putnam's study, titled "Energy in the Future," was primarily concerned with the capital energy resources. These analyses are continued today by the Energy Information Administration in the US Department of Energy (see Fig. 21.8). The capital resources are the fossil and nuclear fuels present in the crust of the earth. The fossil fuels were created by the interaction of the sun with the terrestrial flora and fauna. This was a very low-yield process that over the 5 billion years the earth has existed captured about 2 days of the solar energy in the form of coal, oil, gas, and other combustibles. The recoverable quantities are difficult to estimate [11–14].

The nuclear fuels were created in the cosmic event that created the universe and were deposited in the earth as it took form. There are two families of nuclear fuels, those for fission (uranium) and those for fusion (protium [^1_1H], deuterium, helium-3, and lithium). Only uranium fission has been developed as a commercial source of nuclear energy.

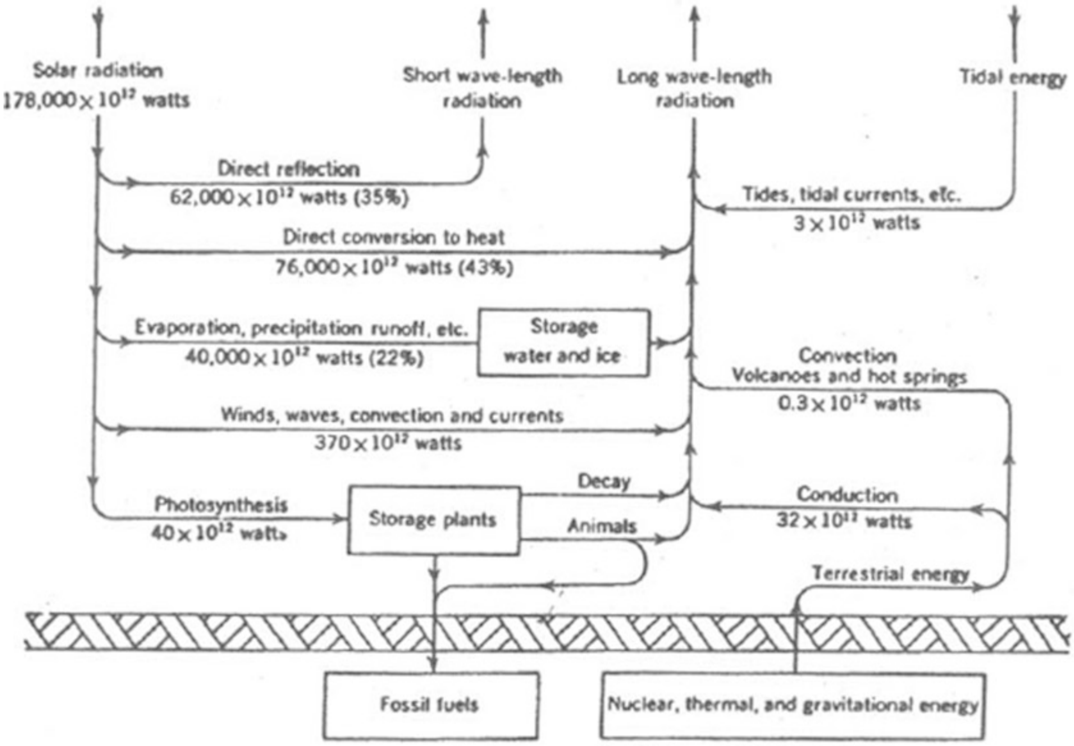
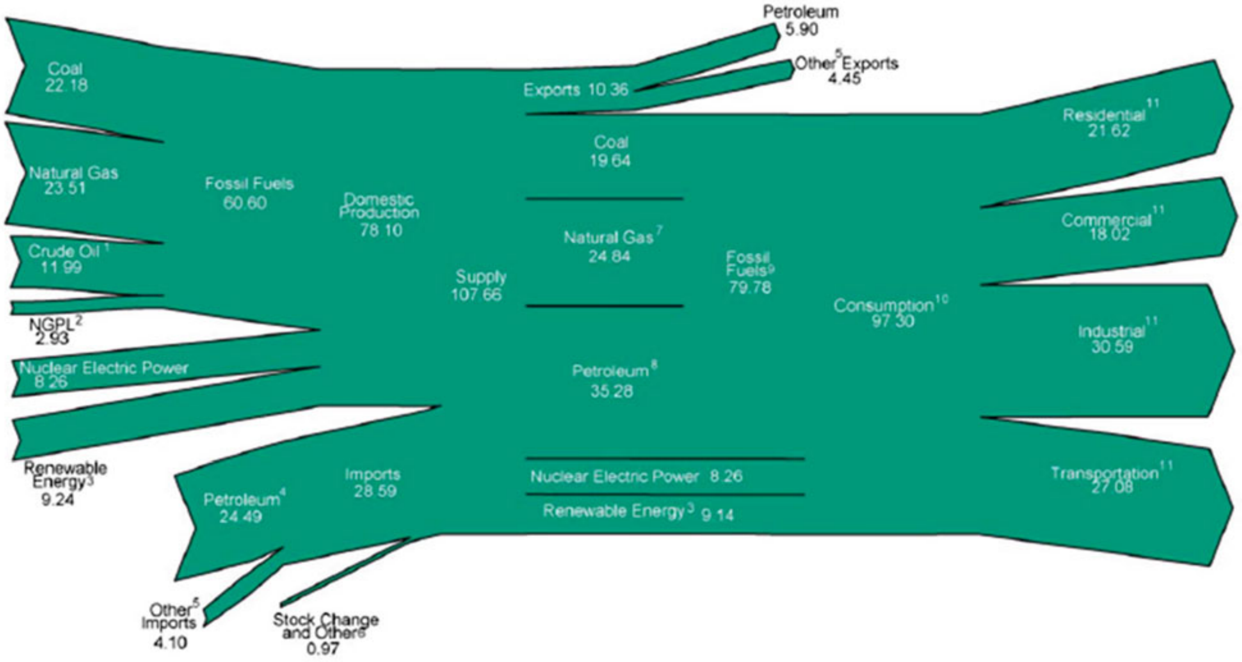


Fig. 21.7 World energy flow sheet (Hubbert, U.S. Geological Survey)



---

[1] Includes lease condensate.
[2] Natural gas plant liquids.
[3] Conventional hydroelectric power, biomass, geothermal, solar/photovoltaic, and wind.
[4] Crude oil and petroleum products. Includes imports into the Strategic Petroleum Reserve.
[5] Natural gas, coal, coal coke, biofuels, and electricity.
[6] Adjustments, losses, and unaccounted for.
[7] Natural gas only; excludes supplemental gaseous fuels.
[8] Petroleum products, including natural gas plant liquids, and crude oil burned as fuel.
[9] Includes 0.01 quadrillion Btu of coal coke net imports.
[10] Includes 0.13 quadrillion Btu of electricity net imports.
[11] Total energy consumption, which is the sum of primary energy consumption, electricity retail sales, and electrical system energy losses. Losses are allocated to the end-use sectors in proportion to each sector's share of total electricity retail sales. See Note, "Electrical Systems Energy Losses," at end of Section 2.

Notes: • Data are preliminary. • Values are derived from source data prior to rounding for publication. • Totals may not equal sum of components due to independent rounding.

Sources: Tables 1.1, 1.2, 1.3, 1.4, and 2.1a.

Fig. 21.8 U.S. energy flow, 2011 (quadrillion Btu) (Annual Energy Review, U.S. Department of Energy 2011)

Although fusion has been developed as a military weapon, the hydrogen bomb, it is premature to include the fusion fuels in the world's inventory of capital energy. The technology for controlled fusion is not available, nor is development of a controlled fusion process expected in the next decade. When available it would increase the capital supply to a level greater than that from all other sources combined [15].

In 1990, the National Academy of Science completed an energy study that concluded that commercial fusion power may be required by the year 2050. By that time 40% of the oil reserves will have been consumed, and current natural gas reserves can supply only about half of the projected needs. Since then, the discovery of additional deposits and improved production methods have met the need for natural gas and decreased prices.

Little use was made of the earth's capital energy until the nineteenth century, and now 5–8% of that resource may have been consumed. Yet even though most of the fossil fuels remain, there are increasing problems with their extraction and distribution and the pollution that is incidental to their use. The demand for energy is increasing as the world continues to industrialize. In 2002, the annual per capita consumption in the United States was 340 million Btus vs.190 million in Russia, 47.7 million in Brazil, and 13.5 million in India. The increasing demand will stem from population growth and increases in the standard of living. It will be accompanied by increased energy use for the extraction of marginal resources and the treatment of wastes before their release into the earth's environment.

Petroleum is and will remain the major source of mobile energy for the next century. However, petroleum production has passed its peak in the United States, and in several decades it may peak in the rest of the world. However, in 2001, approximately 45.8 trillion kWh of energy were consumed worldwide using oil, and projections indicate that this number will increase to approximately 71–74 trillion kWh in 2025, indicating a ~58.3% increase in oil consumption. Also in 2001, the US energy supply from petroleum products was 11.2 trillion kWh, and it is projected that this number will increase to 16.1 trillion kWh, indicating a 43.8% increase [8].

Coal can supply US needs for the next several hundred years, but there is increasing concern regarding pollution from both coal and petroleum. Unfortunately for the developed nations, using pollution-free energy sources such as nuclear and solar will do comparatively little to reduce worldwide atmospheric pollution. The sources of such pollution are worldwide. About 6 billion tons of CO_2 were generated in 1985, and this pollution source is expected to increase to 30 billion t/year by 2060. In the 1980s only 10% of the 6 billion t/year came from the United States. In 2002, this increased to about 5.7 billion tons (Annual Energy Review 2003, DOE) or about 23% of the world's total. In 2002, about 210 million tons of avoided carbon emission was achieved by nuclear (140 million) and renewables (mainly hydro) [16].

The benefits of nuclear power obtainable with the present technology depend on the availability of uranium resources. For example, using the present light water reactor technology, each 1,000-MW reactor requires a few thousand metric tons of uranium during its 60-year operation. Because of limited uranium ore reserves in the United States and the lower cost of uranium from other countries, in the late 1980s about 50% of ore consumed in the United States was imported. This number increased to 64% in 2003. However, US reserves should increase significantly when exploration again becomes profitable. In addition, the energy extracted from the uranium can be significantly increased by reprocessing the spent fuel. If the recovered mixed uranium and plutonium were used in a HRW, there might be as much as a 68% increase in energy generated. A further increase by a factor of 10 might be achieved going to a liquid metal cooled reactor. However, such advanced technology will result in higher electricity cost. In view of the current low cost for natural uranium, storage rather than reprocessing of spent reactor fuel creates a beneficial fission fuel reserve.

Nuclear Processes

The science of the nuclear processes has a long history. Democritus in the fifth century BC started the search for the elementary particle of nature he called the atom. John Dalton, about 1803, related atomic masses to hydrogen, and Mendeleev in 1869 correlated these masses with the atoms´ chemical properties and hypothesized missing elements. Many thought this ended the adventure, but in 1895 Roentgen discovered X-rays, and a year later Becquerel discovered natural radioactivity. A short time thereafter, in 1902, Rutherford and Soddy proposed that radioactivity was related to atomic changes, and in 1905 Einstein equated mass to energy. In 1932 Curie and Joliot discovered artificial radioactivity, and in 1938 Meitner and Frisch recognized that uranium fissioned when bombarded with neutrons. In 1942 Enrico Fermi and Glen Seaborg led development of the first controlled neutron fission reactor to produce plutonium-239 [17].

The nuclear processes of most interest to the nuclear industry are radioactive decay and the transmutation of nuclides. Whereas chemical processes relate to the interactions of orbital electrons of the atom, nuclear processes relate to interactions of neutrons, charged particles, and nuclides with the neutrons and protons in the nucleus of the atom. As noted above, several thousand nuclides and isomers are now known with only 287 of these occurring naturally. More continue to be found. As Mendeleev

Table 21.5 Radioactive decay process examples[a]

	Nuclide			
Decay process	Parent	Daughter[b]	Half-life[a]	Energy (MeV)
Alpha (α, 4/2He)	Pu-238	U-234	89.6y	5.5
Beta (β^-, e)	Cu-64[c]	Zn-64(s)	13 h	0.57
Positron (β+, e)	Cu-64[c]	Ni-64(s)	13 h	0.65
Orbital electron capture (EC, e)	Cu-64[c]	Ni-64(s)	13 h	1.68
Internal transition (IT)	Tc-99 m	Tc-99	6 h	0.14
Fission	U-235 Cf-252	FP[d], n FP, n	Prompt 2.6 years	~200 ~200
Fusion	D/T	He-4(s), n	Prompt	17

[a] Thorium-234 also is radioactive, and the decay process continues through ten more radioactive nuclides before reaching stable lead-210. Decay half-life calculated by $\tau_{1/2} = ((\ell n2(t - t_o))/((\ell n(N/N_o))$ where t = time and N = disintegration/s

[b] Daughters always include gamma rays

[c] Cu-64 has three decay processes (β^- 38%, β^+ 19%, and E, C 43%)

[d] Fission products

(s)Stable nuclides; others are radioactive

invented the chart of the chemical elements, Emilio Segre invented the chart of the nuclides to give order to the nuclear properties and processes [18].

Radioactive Decay

Radioactive decay occurs when an unstable atomic nucleus emits particles and/or energy to achieve a stable state. The nucleus contains neutrons and protons along with the energy that binds them together. Many reactions are involved in radioactive decay (see Table 21.5). These reactions are characterized by the type, energy, and rate of radiation emitted. A different nuclide, which may also be unstable, results from the decay process. This nuclide then will decay with its own unique radiation. Naturally occurring uranium-238 passes through 12 nuclides before reaching a stable end product, lead-210. The types of radiation include:

- Electrons, called beta particles when negatively charged and positrons when positively charged
- Helium ions, called alpha particles
- Electromagnetic energy, called gamma or X-rays, which accompanies almost all other emissions
- Neutrons
- Nuclide ions from fission and fusion

The energy of these emissions covers a wide range of values but is typically 190 million electron volts (MeV) for fission, 17 MeV for fusion, 5 MeV for alphas, 1 MeV for gammas, and 0.5 MeV for betas. The rate of radioactive decay is expressed through the half-life, the time required for the number of atoms of a specific nuclide to decrease by a factor of 2. The half-lives range from less than microseconds to greater than a trillion years.

As radiation is not detected by the human senses, special detection and measurement methods are necessary. The silver halides in photographic film are sensitive to radiation as well as to the electromagnetic energy of visible light; in fact, it was photographic film detection that led to the accidental discovery of radiation. This technique continues to be used, particularly for the study of cosmic irradiation. A closely related method is the use of crystals (sodium iodide and germanium silicide) that emit light when subjected to radiation. This light is measured with photoelectric cells that can be tuned to measure the strength of the radiation. Another popular method measures the discharge of electric condensers in the form of gas-filled ion chambers. The radiation passing through the chamber ionizes the gas and discharges the condenser. Ion chambers are particularly useful for soft alpha and beta radiation measurement. The Geiger counter is an ion chamber with its spontaneous discharge hooked to an audio speaker, whose resulting clicking provides a background noise. All these detection methods have special uses but at the same time a wide range of applications. For each application they are calibrated with radioactive standards that are traceable to national and international reference standards.

Fission

Fission is a relatively simple process. The process is explained in detail very well by Lamarsh [20]. Neutrons at room temperature (termed *thermal*) are captured by a fissile material producing an unstable nuclide which promptly fissions, yielding energy and fission products. (Note that certain nuclei, termed *fissionable*, are caused to fission only by the impingement of an energetic neutron.) The fission products generally include more than two neutrons, and under certain circumstances a fission chain reaction can be sustained. Under special conditions where more than one neutron per fission reaction produces a subsequent fission in the fissile/fissionable material in an uncontrolled manner, a violent explosion can result. However, considerable technical effort is required to hold the material together long enough to obtain a high-yield explosion. In a controlled

Table 21.6 The fission process

Reference equation			
84 % Fission products + ~200 MeV			
<Image/>			
Fission cross-sections[a] vs. *neutron energy*			
Energy (keV)	*Thermal*[b]		
U-235	585		
Pu-239	750		
U-233	531		
Thermal neutron-capture cross section			
Fissionable nuclides (includes fission)		*Moderating materials*	
U-235	684	H_2O	0.33
Pu-239	1,021	D_2O	0.0006
U-233	577	Helium	0.007
Fertile nuclides[c]		Carbon	0.004
U-238	2.7	Sodium	0.01[d]
Th-232	*7.4*	*Neutron poisons*	
Fuel cladding		Boron	760
Zircaloy	0.2	Cadmium	3,300
Stainless steel	3	Gadolinium	46,000

[a] Cross-section in barns (10^{-24} cm^2)
[b] Thermal energy at 20 °C, 2,200 m/s, 0.0253 eV
[c] Capture neutrons to produce fissile nuclides
[d] At 100–1,000 keV

nuclear fission process, such as in a nuclear power reactor, only one of the neutrons is captured by other fissile/fissionable nuclei. The remaining neutrons are used to produce more fissile/fissionable material, are absorbed in other components, or leak from the reactor.

Uranium-235. Natural uranium contains 0.711% uranium-235. The nuclear industry is still completely dependent on this single nuclide, which is the only naturally occurring one that can be effectively fissioned with thermal neutrons. A controlled nuclear fission chain reaction was first achieved on December 2, 1942 under the west stands of the football stadium at the University of Chicago. The Chicago event was achieved in an atomic pile, a stacked array of natural uranium bodies in channels of a large graphite block. The graphite was present to reduce the energy of the neutrons to a value where they are most readily captured in U-235 (see Table 21.6.) Approximately 85% of the 200 MeV released in a fission reaction is in the kinetic energy of the fission products and is instantly captured in the uranium in the form of thermal energy. The remaining energy is released in the form of radiation during decay of the radioactive fission products, e.g., strontium-90 and cesium-137. When this radiation is released, it too is largely converted into thermal energy.

The fission neutrons at birth have energies of approximately 1–2 MeV. In a thermal reactor the neutron energy is rapidly reduced through collisions with light nuclei to thermal (~0.02–1 eV), to promote more efficient capture. Besides the nuclear fuel, there are many other materials in the reactor core also competing for the neutrons, including: the moderator (the material used to slow down or thermalize the neutrons), fertile nuclides that produce additional fissile material (discussed in a later section), neutron poisons present in control rods, the coolant, fuel element cladding, and other structural materials.

Neutron poisons, which are nuclides with very high cross-sections for the capture of neutrons but which do not produce any neutrons, are an essential part of the nuclear reactor core. By their addition and withdrawal, the nuclear chain reaction is controlled with a multiplication factor of 1, that is, one fission per fission in the previous generation. To achieve high fuel burnup (utilization) in a pressurized LWR, boron (a poison) is present in the coolant (and in some cases on the fuel pellets) during the early part of each operating cycle. This serves to remove the excess neutrons (and therefore the excess reactivity of the fuel); subsequently, the boron is removed as the fuel is consumed and poisons from the fission products accumulate. This is the primary means of controlling the neutron multiplication throughout the operating cycle. Cadmium and/or boron is placed in control rods for fine tuning such as load-follow operation and reactor shutdown. Gadolinium and boron, which have very high cross-sections for neutron absorption, find use in emergency shutdown of HRWs and as a volumetrically dispersed component within the fuel pellets of some LWRs. Table 21.6 shows cross-sections at thermal energy for the principal fissile materials.

Plutonium-239. Plutonium-239 represents a fortuitous phenomenon. Whereas U-235 is the only significant fissile nuclide in nature, its major isotope, U-238, captures a neutron to produce another fissile nuclide, plutonium-239. A substantial amount of the energy produced during the life of uranium fuel is produced by the conversion of U-238 to Pu-239 which subsequently fissions. This process provides the basis for the nuclear breeding cycle.

Uranium-233. A second fissionable isotope, uranium-233, can be produced from naturally occurring thorium. It does not present an economically attractive option at present because of its dependence on highly enriched U-235 or Pu isotopes to bring the thorium cycle into operation and the large R&D expenditures required to develop the technology (for more information see ref [21]). In addition, U-232 is also generated along with the U-233 whose decay products produce high energy gammas. Fabrication of U-233 based fuel would therefore require extensive amounts of shielding making this fuel significantly less economical than U-235 based fuel.

Fusion

At present the only application for fusion is in thermonuclear military explosives, where the necessary temperature and

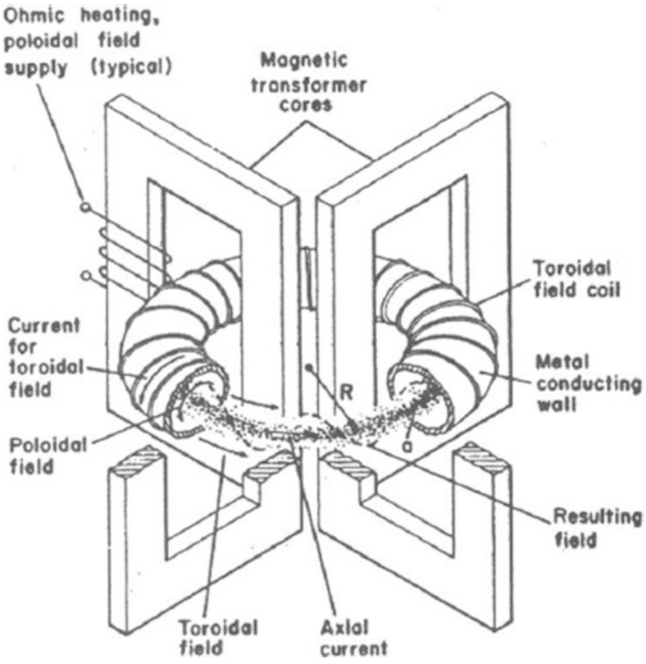


Fig. 21.9 Tokamak fusion reactor (Courtesy of U.S. Department of Energy)

pressure are achieved with a plutonium fission initiator for a very short time. The search for the science and technology needed to make fusion a viable source of energy for commercial electric power generation is the basis for one of the world's most important and challenging R&D efforts. The development of a controlled nuclear fusion reactor could provide a virtually unlimited source of energy. Like fission, fusion eliminates the atmospheric pollution associated with the use of fossil fuels. Also, it could significantly reduce the waste disposal problem of energy-producing processes [21, 22].

In the Tokamak fusion reactor depicted in Fig. 21.9, electric current to the poloidal coils on the primary magnetic transformer generates the axial current in the secondary plasma composed of deuterium and tritium ions. These ions are heated to ignition temperature and then the reaction becomes self-sustaining. The toroidal field coil suspends the plasma away from the metal conducting walls. Contact with the wall would both cool the plasma below ignition temperature and contaminate the plasma with heavy ions. The relevant reactions are given below.

Deuterium-Tritium Fusion

1. ${}^{2}_{1}\mathrm{H} + {}^{3}_{1}\mathrm{H} \rightarrow {}^{4}_{2}\mathrm{He} + {}^{1}_{0}n + 17.6\,\mathrm{MeV}$

It should be noted that this reaction is also used to provide rugged, durable commercial sources of 14 MeV neutrons for oil-well logging and scientific research. Similarly, deuterium–deuterium fusion provides sources of 2.5 MeV neutrons.

Tritium production is accomplished in a fusion reactor blanket (or a fission reactor, as is now done to produce tritium for the weapons program).

2. ${}^{6}_{3}\mathrm{Li} + {}^{1}_{0}n \rightarrow {}^{3}_{1}\mathrm{H} + {}^{4}_{2}\mathrm{He}$
3. ${}^{7}_{3}\mathrm{Li} + {}^{1}_{0}n \rightarrow {}^{3}_{1}\mathrm{H} + {}^{4}_{2}\mathrm{He} + {}^{1}_{0}n$

Protium lithium fusion would produce charged particles (90% of the energy in helium ions) for direct conversion to electricity, but higher temperatures and pressure would need to be achieved.

4. ${}^{1}_{1}\mathrm{H} + {}^{6}_{3}\mathrm{Li} \rightarrow {}^{3}_{2}\mathrm{He} + {}^{4}_{2}\mathrm{He} + 4\ \mathrm{MeV}$
5. ${}^{1}_{1}\mathrm{H} + {}^{7}_{3}\mathrm{Li} \rightarrow {}^{4}_{2}\mathrm{He} + {}^{4}_{2}\mathrm{He} + 17\ \mathrm{MeV}$

Of the several fusion reactions, deuterium-tritium fusion is the most feasible, as it has the lowest ignition temperature, 40 million °K (see Reaction 1). Deuterium comprises 0.15% of naturally occurring hydrogen, whereas tritium is produced by neutron fission of lithium-6 that is irradiated in a blanket surrounding a nuclear reactor core or by the absorption of a neutron by deuterium, for example, in a HRW. Nuclide separation is required to produce the deuterium and possibly the lithium. The tritium would be produced in a surrounding lithium blanket of the fusion reactor to sustain the process (see Reactions 2 and 3). The tritium for hydrogen bombs is produced in nuclear fission reactors using these processes. The alternative fusion reactions use protium and lithium and yield charged helium ions (see Reactions 4 and 5). These would contain 90% of the fusion energy and could be directly converted to electricity, but higher temperatures and pressures would be required.

There are now two approaches to developing a fusion reactor: magnetic containment and inertial confinement. The major effort, magnetic containment, uses the Tokamak concept to compress and heat the reactants in a plasma isolated in a vacuum away from the walls of the reactor. The fuel is injected as pellets and heated with electric ohmic energy to strive for the ignition temperature. Research along this line started in the early 1950s, and in 1968 the Russians discovered the Tokamak principles. (Tokamak is the Russian acronym for "toroidal chamber with magnetic coil.") Massive amounts of electrical energy are required to heat and compress the plasma. Ohmic energy now is supplemented with induction, microwave, and neutral beam techniques. The time of heating has been extended from microseconds to over a minute using these additional methods.

There are two immediate objectives in fusion research, first to achieve ignition self-sustaining burning and then to achieve parity (produce more energy than is required to maintain steady-state operation). In the 1990s the International Energy Agency (IEA) member countries spent approximately \$8.9 billion on fusion R&D. Today (2012), the majority of fusion research is carried out by the European Union (EU), the United States, Russia, China, India, South Korea and Japan. The largest Tokamaks that have been recently studied are the Tokamak Fusion Test Reactor (TFTR) located at Princeton, New Jersey, United States, operated from 1982 to 1997, and the Joint European Torus (JET) located in the United Kingdom (still in operation). JET was first to demonstrate breakeven (output power = input power) in 1997, although ignition had not been achieved as of 2012. However, it was decided by the Soviet Union in the mid 1980s that a next generation Tokamak was needed. In 1992 the Soviet Union, Europe, Japan, and the United States agreed to collectively design an International Thermonuclear Experimental Reactor (ITER). Although the United States pulled out of the collaborative group in the late 1990s, it rejoined in 2003. The cost of the ITER is thought to be about \$12 billion over the 30 year life of the project and its construction would take 10 years. The EU is expected to cover 45% of the project's total cost, with the other six partners—China, India, Japan, South Korea, Russia and the United States—contributing about 9% each [23]. Inertial confinement fusion is an alternative technique that uses lasers to generatic high intensity x-rays to compress a bead containing deutarium and tritium. This technique is being developed at Lowrence Liver more National Laboratory in the US, among others in Japan and the UN. When ignition and parity are finally achieved and understood, spending will need to be substantially increased to establish commercial feasibility. It is now realized that electricity from fusion will be more costly than that from fossil and fission processes. It is only as these resources are depleted that fusion will become economically feasible.

Nuclide Production

Nuclides are produced by capturing a nuclear particle in the nucleus of the target atom, which thus is transmuted to a different atom, one that in most cases is unstable and decays, as discussed in an earlier section [24].

The target atom may be any stable or radioactive atom. In a process termed irradiation, it is bombarded, leading to production of many of the nuclear particles discussed in the previous section. Most often the incident particle is a neutron produced in a nuclear reactor, but frequently it is a charged particle raised to the required energy in an electromagnetic accelerator. The charged particles often are protons but may be deuterons, helium nuclei (He-4 or He-3), or heavier ions [25, 26].

The rate of nuclide production is dependent on the number and the energy of particles bombarding the target, the cross-section of the target and the half-life of the product (see Table 21.7). The decay of these products can be a source of significant amounts of energy, as already discussed.

Table 21.7 Nuclide production process

$A = N\sigma\phi(1 - e^{-\lambda t})$	
Where	
$(1 - e^{-\lambda}t)$	= Saturation factor: for small values of λ, it equals λt; for $t = \tau_{1/2}$, it equals 0.5, the optimum irradiation time
A	= Disintegrations per second; 1 Ci = 3.7 E10 day/s
N	= Number of atoms; Avogrado's number = 6.023E23 atoms/g atomic wt
Σ	= Sorption cross-section; 1 b = 10^{-24} cm^2
Φ	= Irradiation flux, particles/s/cm^2; for electromagnetic accelerators reported in microamperes
Λ	= Decay constant, natural logarithm 2/half-life
T	= Irradiation time, seconds
$\tau_{1/2}$	= Nuclide half-life
Examples	
1. Co-60 production (Co-59 target at 1E14 s/cm^2, $\sigma = 37$ b)	
Irradiation time (years)	*Curies Co-60/g target*[a]
1.1	130
2.2	225
5.5	370
2. Pu-239 production (U-238 target at 1E14 s/cm^2, $\sigma = 2.7$ b)	
Grams Pu-239/metric ton target	
1	4,800
2	9,600
3	14,400

[a] These quantities ignore depletion of the products through decay and conversion to other products

Fission Products

The fission process produces radioactive as well as stable nuclides with mass numbers ranging from 72 to 167 and with two broad peaks in the regions of 95 and 138. The masses are identified rather than the specific nuclides because in fission many short-lived nuclides are produced that quickly decay by beta emission to a long-lived or stable member of the mass chain [18, 19].

The yields of selected mass chains that result from fission of U-235 and Pu-239 are shown in Table 21.8. In addition to the fission products with masses roughly half that of U-235, neutrons, tritium, helium, and beryllium are products of ternary fission. There are significant differences in some of the yields from U-235 and Pu-239.

Although the fission products could be recovered as byproducts from the waste from spent nuclear reactor fuel, special-purpose neutron irradiation of HEU (isotopically separated uranium-235) followed by chemical separation is

Table 21.8 Major fission products

Mass	Long-lived nuclides	State	Half-life	Percent yield[a]	
				U-235	Pu-239
72	Zn		46.5 h	0.001	0.001
79	Se	Meta stable	3.92 min	0.044	0.044
		Ground	6.5 E4–6.5 E5 years		
85	Kr	Meta stable	4.48 h	1.3	0.58
		Ground	10.76 years		
90[b]	Sr		28.78 years	5.8	2
99	Mo		2.7476 days	6.1	6.2
102	Pd(s)		1.02% (a/o) abundance	7.7[c]	28.1
103	Rh(s)		100% (a/o) abundance	3	7
124	Xe(s)		0.096% (a/o) abundance	23.7	27.2
131	I		8.0207 days	2.9	3.9
133	I		21 h	6.7	7.0
133	Xe	Meta stable	2.19 days		
		Ground	5.243 days		
137	Cs		30.07 year	6.7	6.6
140	Ba, La		1.678 days	6.2	5.4
167	Dy		6.2 min	0.001	0.001

[a] Because two fission products are emitted for every binary fission, the cumulative yield for the mass chains is 200%

[b] Mass-90 beta decay chain (half-life): Br(2 s), Kr(32 s), Rb(4 m), Sr(29 y), Y(2.6 day), Zn(stable)

[c] This is equivalent to 1.2 kg of palladium/mt of spent fuel burned to 32,000 MWd. One megawatt-day of thermal energy from fission is approximately equivalent to 1 g of fission products

(s) Stable products of the mass chains. Multiple chains contribute to the high yields for palladium and xenon

the normal production method. The major products, molybdenum-99 and iodine-131 with fission yields of 6.1 and 6.7%, respectively, have important medical applications. Mo-99, with a half-life of 2.7 days, is particularly interesting because its short-lived decay product, technetium-99 m with a half-life of 6 h, is the product actually used. To make this short-lived product available, the Mo-99 is sorbed in a silica gel column, and the Tc-99 m is eluted as needed.

During the 1980s, large-scale processing of the defense wastes at Hanford extracted and encapsulated 100 MCi of strontium-90 and cesium-137 to reduce heat generation in the waste storage tanks. Krypton-85 has been recovered from the airborne effluents at the Idaho processing plant.

In addition, development studies have addressed the recovery of stable rhodium, xenon, technetium, and palladium because of their limited availability in nature. For palladium, this would amount to about 40 kg/year from a 1,000 MWe reactor. However, a small amount of palladium-107 (half-life 6E6y) is present that could complicate its use. This is also a problem for technetium, which is 100% technetium-99 with a half-life of 2E5y.

Iodine-135 provides an example of a fission product chain that must specifically be considered in reactor operation. Iodine-135 with a 6.6 h half-life decays to 9.1 h xenon-135, which has a high cross-section for thermal neutrons. When a reactor is shut down, the xenon-135 grows in from its I-135 parent, because it is not being eliminated by neutron irradiation. For high-flux reactors this could lead to a "dead-time" of tens of hours during which the reactor cannot be restarted. These effects are controlled through the adjustment of fixed and dissolved neutron poisons during power changing maneuvers.

Neutron Transmutation Products

The production of neutron transmutation products involves the neutron irradiation of target materials followed by chemical separation. Although this promises high purity with high specific activity, the isotopic composition of the target along with other competing reactions frequently yields a spectrum of isotopes in addition to the desired nuclide.

Plutonium-239 and tritium for use as military explosives are the two major transmutation products. The nuclear process for Pu-239 production is the same as for energy generation, but there are some differences: (a) metallic natural uranium clad with aluminum facilitates later dissolution for plutonium recovery, and the reactor operates at a relatively low temperature because of the aluminum cladding and better heat transfer (due to the metallic natural uranium); (b) the irradiation cycle is limited to a few months to minimize the Pu-239 conversion to Pu-240 and Pu-241; and (c) a carbon or a heavy water moderator is used to increase the neutron efficiency.

More complex is the production of Pu-238 (used for isotopic heat sources) and californium-252 (used in research as a source of fission neutrons). For Pu-238 there are two irradiation cycles: the first starts with natural uranium but maximizes the neptunium-237 yield; then the Np-237 is separated, converted into aluminum-clad oxide targets, and irradiated to Pu-238 in a second cycle. The total process requires double neutron capture.

Californium-252 production was especially challenging, as it involved the sequential capture of 14 neutrons along with the intermediate separation and fabrication of two intermediate targets (americium and curium isotopes) when starting with Pu-239 [27]. This production campaign lasted 10 years, produced about 10 g of Cf-252, and then was terminated. The product was evaluated as a neutron source but had insufficient value to justify continuing production.

Neutron Activation Products

The neutron activation production cycle requires only neutron irradiation without chemical separation. The target and the product are the same chemical element but have different nuclide compositions. The specific radioactivity of the product is a function of the nuclide composition of the target, the neutron flux environment, the irradiation time, and the half-life of the product nuclide, along with the nuclear cross-sections of target and product nuclides.

The major neutron activation product is cobalt-60 with a reasonably long half-life (5.3 years). It is produced by capture of a thermal neutron by cobalt-59, the only naturally occurring isotope of that element. Cobalt-60 emits high-energy beta particles and gamma rays appropriate for radiographic, irradiation, and isotopic power applications. It is an ideal product because naturally occurring cobalt is a relatively stable metal that is uniquely monoisotopic cobalt-59. Furthermore, the target nuclear cross-section is good, and the product cross-section is lower than the target's by a factor of 10. Another factor in cobalt-60s long list of advantages is that the Canadian CANDU family of nuclear power reactors are ideal for production of Co-60.

Other important activation products include molybdenum-99 and iridium-192. However, the Mo-99 is better obtained from fission, as has been previously discussed. The 74-day Ir-192 has a gamma with less energy than Co-60 and is used for radiography of less dense materials.

Charged Particle Transmutation Products

Many radionuclides can be produced in cyclotrons, thus avoiding the use of more costly nuclear reactors. Many research hospitals now have cyclotrons to provide short-lived radionuclides of carbon, nitrogen, oxygen, and fluorine. The longer-lived products are produced commercially or in government laboratories [26, 27, 29]. A list of major isotopes and their uses is shown in Table 21.9.

Table 21.9 Major isotopes and their uses

Isotope	Use
D	D_2O moderator for HWR Fuel in fusion reactions
Li-6	Source for tritium Fuel in fusion reactions
Li-7	Water-cooled reactors use lithium-7 hydroxide as a water conditioner Possible coolant in LMFBR
F-18	Used in nuclear medicine, especially in labeling D-glucose
N-13	Tests done show that when used in ammonia, helps to determine global myocardial blood flow
C-14	Used in carbon dating
N-15 C-13 D-17 O-18	Used as tracer in living beings Using in NMR
B-10	Neutron absorber in nuclear reactors Used in the medical field in neutron-capture therapy
B-11	Can be used in tritium production and in NMR
Na-24	Used to locate pipe leaks
Mg-27	Used to locate pipe leaks
K-42	Can measure exchanged potassium in blood
Cr-51	Used to label red blood cells
Fe-57	Used as a spin isotope in chemistry
Fe-59	Used in blood studies Can be used to determine friction in machinery when used with steel
Co-60	Used in cancer treatment
Ga-67	Used to locate tumors
Kr-81	Used in lung ventilation studies
Tc-99	Used as a tracer in locating brain tumors
I-131	Used as a tracer when studying the thyroid gland
Yb-169	Used during brain scans
U-235	Fuel for most nuclear reactors
Pu-239	Used in nuclear weapons, fast breeder reactors, and MOX fuel reactors
Am-241	Used in smoke detectors

These radionuclides are produced by irradiating targets with beams of hydrogen ions (protons), but frequently deuterium ions (deuterons) are used (see Table 21.10). Some products require beams of helium-4 and helium-3 ions. The typical process involves the capture of the proton with the prompt emission of a neutron. This is called a p,n reaction. However, in other cases there may be protons, alpha particles, or up to five neutrons emitted. The resulting products decay generally by positively charged electron (positron) emission, but also decay by capture of an orbital electron.

To produce the beam of high-energy charged particles, a stream of hydrogen or helium is electrically ionized and accelerated in an oscillating magnetic field that holds the particles in a spiral path within the cyclotron. At the desired

Table 21.10 Nuclide production by charged particles

Typical reaction
${}^{14}_{7}\mathrm{N} + {}^{1}_{1}\mathrm{H} \rightarrow {}^{15}_{8}\mathrm{O} \rightarrow {}^{4}_{2}\mathrm{He} + {}^{11}_{6}\mathrm{C} \overset{20\mathrm{m}}{\rightarrow} \beta^{+} \rightarrow {}^{11}_{5}\mathrm{B}$ (**stable**) This reaction normally is reported as N-14(p,α)C-11. Natural nitrogen gas is bombarded in a cyclotron with a 30-mCi current of 4–13 MeV protons (p). O-15 is produced, but an alpha particle is promptly ejected, producing C-11, the desired product. The yield for a 20-min bombardment is 0.2 Ci
${}^{58}_{28}\mathrm{Ni} + {}^{1}_{1}\mathrm{H} \rightarrow {}^{59}_{29}\mathrm{Cu} - 2{}^{1}_{1}\mathrm{H} + {}^{57}_{27}\mathrm{Co} \underset{270\mathrm{d}}{\overset{\mathrm{EC}}{\rightarrow}} {}^{57}_{26}\mathrm{Fe}$ (stable) Ni-58[p,2p]Co-57. Enriched Ni-58 (enriched in the Oak Ridge Calutron) is plated on a copper heat exchanger, bombarded for 5 h with 1,000 mA of 20 MeV protons. Cu-59 is produced but two protons are promptly ejected producing Co-57, the desired product. The yield for a 14 h bombardment is approximately 28 mCi/mA hour
Other (p,n) processes: Li-7 to Be-7; Fe-56 to Co-56; Cu-65 to Zn-65; Ag-109 to Cd-109
Other processes and products: N-14(d,n)O-15; O-16(p, α) N-13; Ne-20(d, α)F-18

β^{+} represents a positron
EC represents electron capture

energy the particle beam is magnetically deflected to the target. The beam energy can vary from 4 to 800 MeV but is usually less than 30 MeV. The beam current ranges from 3 to 2,500 μA. The target is designed to accommodate the ten to several hundred watts of thermal energy deposited by the beam. Gas and liquid targets are used for the short-lived products to facilitate processing for use. For the long-lived products that require longer irradiation and higher energy flux (up to a kW/cm^2), the target material may be plated on a water-cooled heat exchanger. The target materials in many cases are separated stable nuclides.

The Department of Energy's laboratories have accelerators that provide up to 800 MeV proton beams. Other machines support high-energy physics research but occasionally are used to produce special nuclides such as xenon-127.

Reactor Materials Processing

Reactor materials processing is concerned with preparation of the special materials used in nuclear reactors. This includes the nuclear reactor fuels along with the reactor vessels, plumbing, heat exchangers, coolants, and moderators. All the materials that go into the nuclear reactors are deliberately and thoroughly controlled by standards established by the nuclear industry with the participation of the Nuclear Regulatory Commission and administered by the American National Standards Institute. These materials include the uranium oxide pellets, zirconium cladding, and zirconium or steel fittings of the fuel assemblies, the corrosion–and radiation-resistant steel used for the reactor vessels, and the various corrosion–and erosion-resistant steels for piping, valves, pumps, and heat exchangers. In addition, there are the coolants (water, sodium, and helium), the moderators (water, deuterium, and graphite), and the neutron sorbers (boron, erbium, cadmium, hafnium, and gadolinium). Each of these materials requires intensive processing to achieve safe and reliable operation of the nuclear reactors.

Isotope Enrichment

Isotope enrichment is a major activity of the nuclear industry. The naturally occurring chemical elements are composed of isotopes with varying nuclear properties. By enriching the chemical element in the nuclide with the desirable nuclear property, the performance of the nuclear processes is enhanced. In addition, the sensitivity of isotopic analysis permits the use of enriched chemical elements in place of radioisotopes as tracers in research, medicine, and industry.

The separation of chemical isotopes is based on small differences in their physical and chemical properties. For the lower-mass isotopes, chemical exchange, distillation, and electrolysis have been used. For the higher mass isotopes, techniques based on mass have been used, including jet nozzles, gaseous diffusion, centrifugation, thermal diffusion, and ion activation [29, 30]. A newer method uses lasers that produce coherent light tuned to the specific wavelength of a vibration bond related to the desired isotope in an atom or molecule. This technique is still under development but promises much higher per stage separation factors resulting in lower energy consumption and capital costs. The main issue with this technology is the availability of low capital and operating cost lasers with sufficient wavelength selectivity.

The most universal of these techniques uses the Calutron, an electromagnetic accelerator employing the first arc of a cyclotron (Fig. 21.10). In 1945 it provided the enriched U-235 that was used in an early atomic bomb. Since then it has been used to enrich small quantities of most of the stable isotopes. The separation is achieved by ionizing the feed material and then accelerating the ions through the field of a bending magnet. The differences in mass causes the heavier elements to pass through a larger arc with the ions trapped in carbon pockets located 180° from the feed point. The accelerating voltage and magnetic field are controlled to optimize performance. Separation factors of 10–100 are obtained, but the yields are only 1–10%.

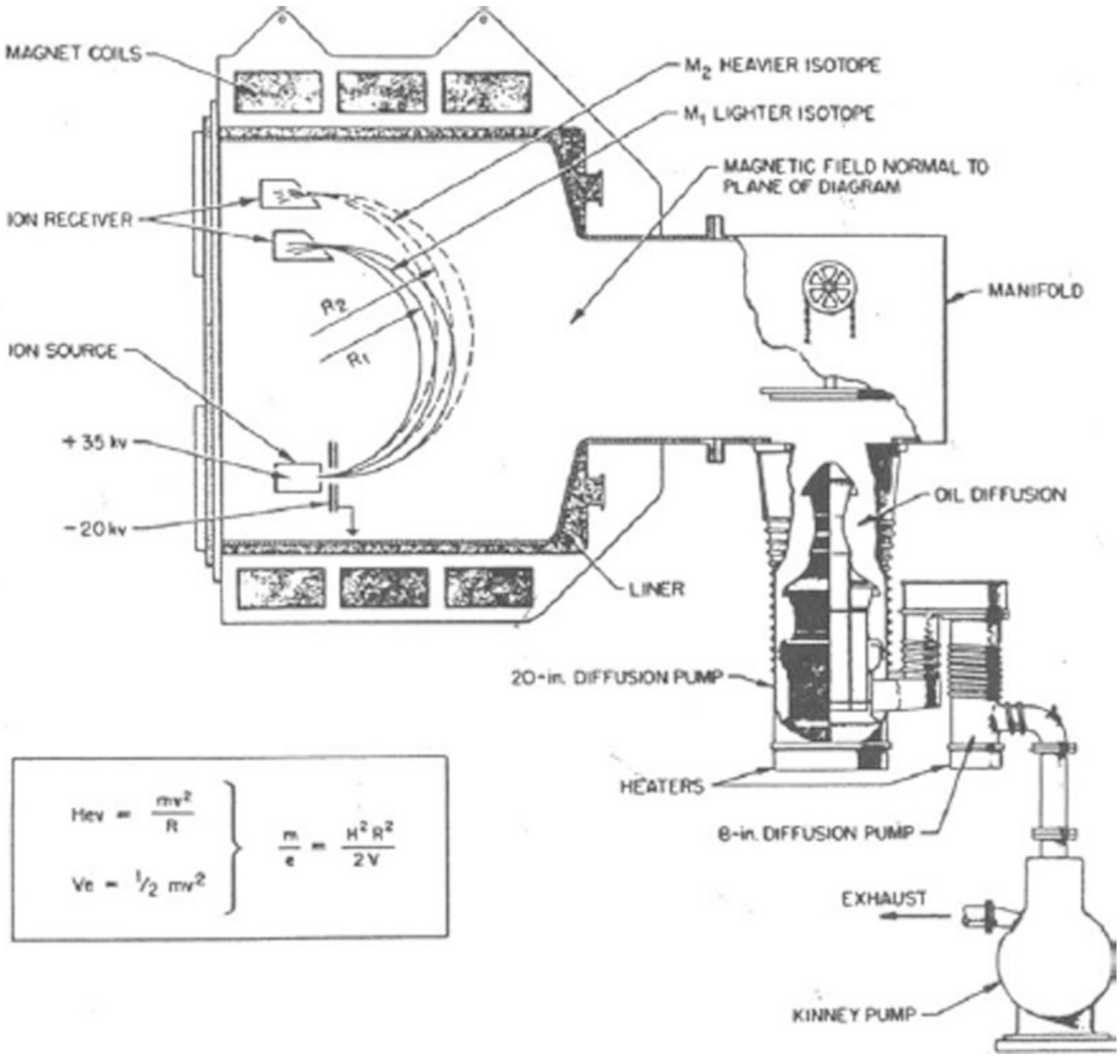


Fig. 21.10 Oak Ridge Calutron for separation of isotopes. The first method to achieve large-scale separation of U-235 and today the source of research quantities of most stable and a few radioactive nuclides (Courtesy of USDOE)

Large-scale production has been applied to enrich U-235, deuterium, boron 10, and lithium-6. The original gaseous diffusion technology has now been replaced by high-speed centrifuges as the principal technology for U-235 and is discussed further as part of the fuel cycle. Dual-temperature chemical exchange processes are used to enrich hydrogen and lithium isotopes. Distillation is used for boron 10.

Several methods can be used to obtain high-purity deuterium for use as a coolant in HRWs. Natural hydrogen contains 0.16% deuterium and is enriched to 10–20% deuterium in pairs of 200 plate columns [31]. The chemical exchange system is hydrogen sulfide gas and water. One column operates at 27°C and the second at 220°C, with the enriched product removed from the bottom of the cold column that flows to the second set of columns. From the second set it goes to distillation and then to electrolysis to yield 99.8% deuterium. The specific enrichments at the crossovers between processes are controlled by energy considerations. Other methods in use are the chemical exchange reaction between liquid ammonia and hydrogen and nitrogen gas using potassium amide as a catalyst. At least two plants are in commercial operation: one in Argentina and one in India.

The chemical exchange system employed for lithium-6 enrichment is lithium amalgam and aqueous lithium hydroxide. It also employs paired dual-temperature columns.

Cryogenic distillation is used to enrich carbon-13, nitrogen-15, oxygen-17, and oxygen-18. Typical of these processes is carbon monoxide distillation, which has a C-13/C-12 separation factor of 1.008 between the vapor and the liquid. The initial 20-m packed column is tapered from 2.5 cm at the boiler to 10 cm at the reflux condenser. This brings the carbon-13 from 1.1 to 12%. In a second column it is brought to 93%.

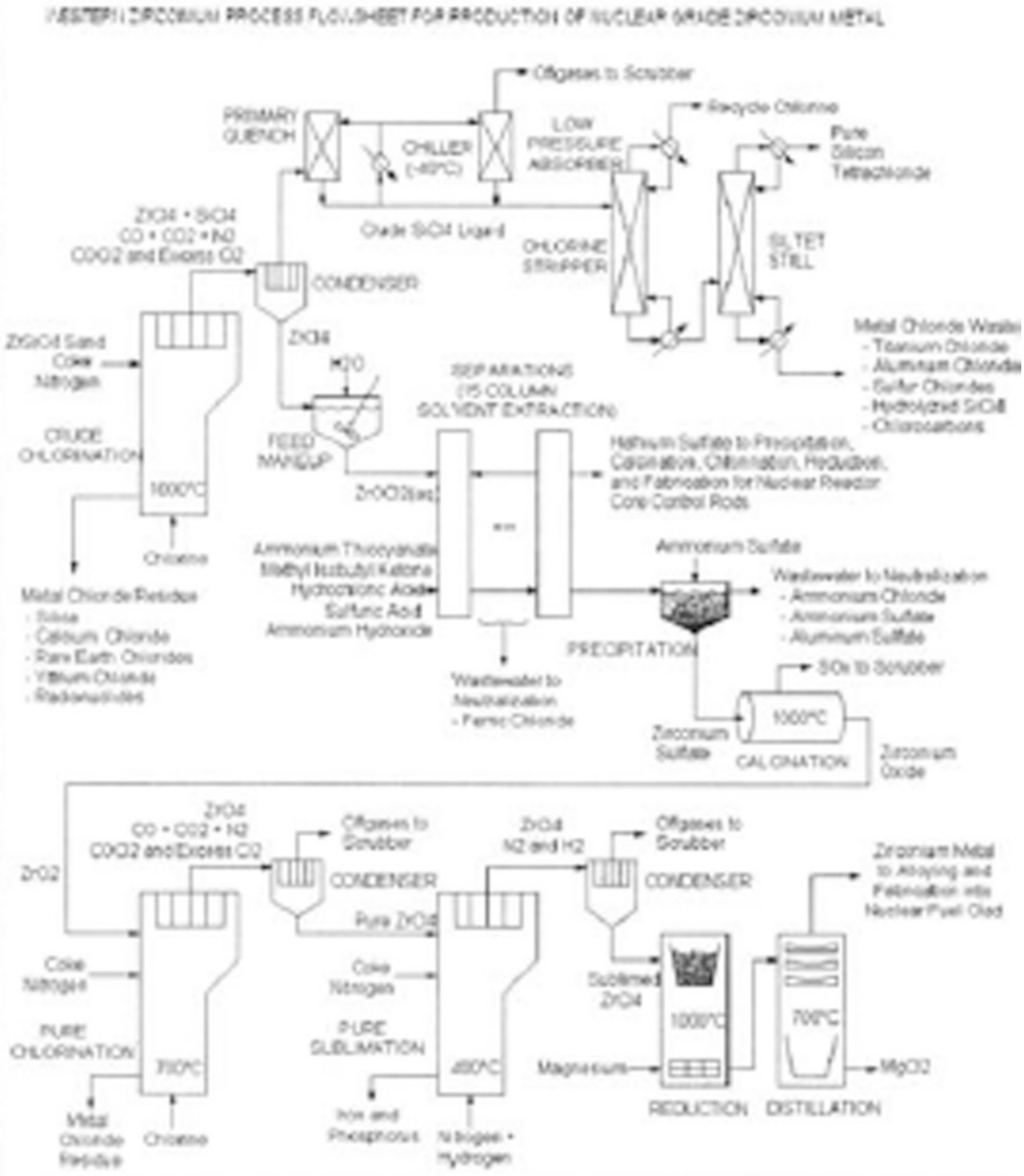


Fig. 21.11 Schematic of Western zirconium plant in Ogden, UT

Zirconium Production

Nuclear-grade zirconium alloys are used for the manufacture of nuclear fuel rod tubes and structural components. Processing begins with zircon sand ($ZrSiO_4$) obtained from Australia, India, or a variety of locations (see Fig. 21.11). Zircon sand is co-milled with coke, and injected into a large fluidized bed reactor (crude chlorination). The reactor is heated to approximately 1,000°C by induction heating of the graphite liner, and the bed is fluidized with chlorine gas. Zircon is converted to zirconium tetrachloride by the following reactions.

$$\begin{aligned} ZrSiO_4 + 4\,Cl_2 + 4\,C &\rightarrow ZrCl4_4 \\ &+ SiCl_4 + 4\,CO; \\ ZrSiO_4 + 4Cl_2 + 2C &\rightarrow ZrCl4_4 \\ &+ SiCl_4 + 2CO_2. \end{aligned}$$

The process gases pass through two stages of condensation, separating low-volatility $ZrCl_4$ product from the remaining offgases. Metal chloride impurities that escape from the chlorinator also condense with the zirconium product. The offgases pass through several stages of cryogenic quenching with liquid $SiCl_4$, recovering $SiCl_4$, Cl_2, and $COCl_2$. Distillation operations then separate silicon tetrachloride from chlorine and phosgene. The latter two species are recycled back to the crude chlorinator, and $SiCl_4$ is purified further by activated charcoal absorption, and sold as a valuable by-product. Remaining offgases are scrubbed with aqueous NaOH to remove traces of chlorine and phosgene, and are then vented; salt solutions resulting from scrubbing operations are sent to evaporation ponds.

The zirconium tetrachloride product must then be purified before reduction to metal. In particular, hafnium must be removed to less than 100 ppm Hf:Zr because of the high neutron absorption cross-section it exhibits, and phosphorus and aluminum must be removed to even lower specifications due to their deleterious metallurgical impact on the final zirconium alloys. The tetrachloride product is first dissolved in water under carefully controlled conditions to produce an acidic $ZrOCl_2$ solution. This solution is complexed with ammonium thiocyanate, and contacted with methylisobutyl ketone (MIBK) solvent in a series of solvent extraction columns. Advantage is taken of the relative solubilities of Zr, Hf, and Fe thiocyanate complexes to accomplish a high degree of separation of hafnium and iron from the zirconium.

Purified zirconyl chloride solution is then reacted sequentially with surfuric acid and ammonium hydroxide to precipitate a complex zirconium oxysulfate. The precipitate is washed, filtered, and stripped to remove traces of MIBK, and then calcined to drive off sulfur and convert the product to ZrO_2. The precipitation process leaves behind most of the aluminum and phosphorus. The hafnium stream leaving solvent extraction is treated similarly, producing a HfO_2 by-product.

Zirconium dioxide is then carbochlorinated a second time (pure chlorination) and converted once again to $ZrCl_4$. This operation is very similar to crude chlorination, but takes place at lower temperature due to more favorable thermodynamics. The purified $ZrCl_4$ product is then reduced to zirconium metal by direct reaction with molten magnesium metal, according to

$$ZrCl_4 + 2\ Mg \rightarrow Zr + 2\ MgCl_2.$$

This reaction (known as Kroll reduction) is carried out in sealed vessels at approximately 1,000°C, producing a mixture of $MgCl_2$ and a porous zirconium metal morphology known as sponge. Because of the large density difference, the molten $MgCl_2$ floats above the sponge; most of it is recovered by mechanical separation, and sold as a high-purity byproduct.

The purified sponge is then crushed, combined with alloying elements, pressed into large cylindrical ingots, and electron-beam-welded into a solid cylinder. This cylinder is then remelted by vacuum arc melting to obtain the proper metallurgical properties. All of these operations must be done with great care to avoid introduction of oxides and nitrides into the final metal product, both of which will result in embrittlement. From here, the ingot is heated, forged, and formed into either flat plate or heavy walled tubing known as Trex (Tube reduced extrusion). The Trex is then extruded and pilgered into the final nuclear fuel to form, and the plate is formed into other structural internal components. A variety of different zirconium alloys are produced, having specific applications for different reactor designs.

The Uranium Fuel Cycle

The uranium fuel cycle for the uranium–plutonium system is a multicomponent system of chemical process operations that begins with mining uranium ore from the earth as the starting material and ends with the radioactive waste products in above-ground or belowground terminal repositories. Some radioactivity is released in the air and water discharged to the environment at concentrations and quantities below those specified by federal regulations. The topics to be considered in this section are mining, milling the ore, uranium fuel preparation, and reprocessing [32, 33]. The following section discusses radioactive waste management.

A similar set of processes has been partially developed for the thorium–uranium system but is not discussed here because it is not expected to be employed in the next several decades. The important feature of the thorium cycle is that it could be used to achieve breeding (to produce more fissionable material than is consumed) in thermal reactors, but nuclear as well as chemical factors have frustrated this development (for more information, see ref. [21]). Thorium based cycles may in the future povide a means of burning very long half-life actinides such as Pu, cm and Am to make a high level Nuclear waste that decays to lower radiatoxicity levels in hundreds rather than thousand on tens of thousands of years.

It should be noted that breeders would not reduce the demand for uranium ore for many decades because several LWR and/or HWR converters (which produce fissionable material, but less than consumption) are required during the run-in of a breeder cycle to equilibrium. The doubling time of a breeder (the time required for the breeder to produce sufficient fissionable material to start up a second breeder reactor) might be a significant part of its operating life. The tritium for the fusion cycle will be made in nuclear

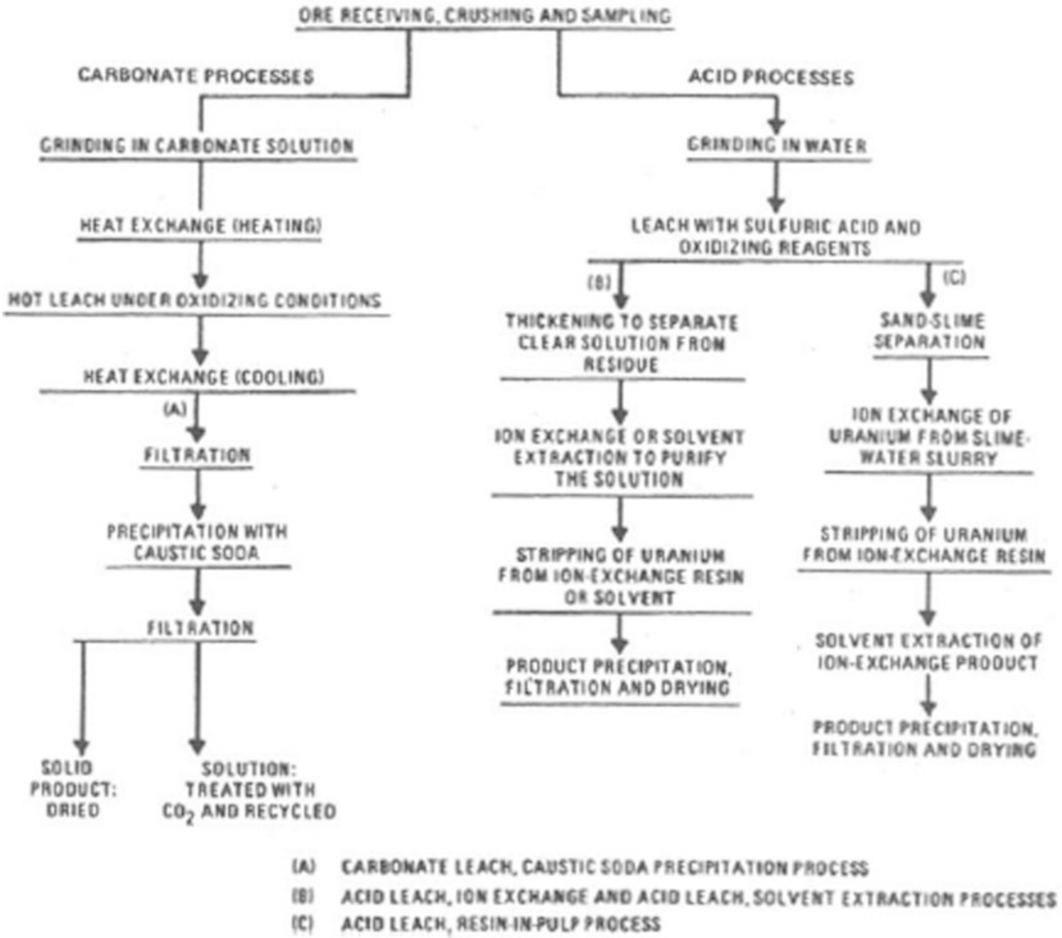


Fig. 21.12 Uranium ore mill processes: (**a**) carbonate leach, caustic precipitation process; (**b**) acid leach, ion exchange and acid leach, solvent extraction processes; (**c**) acid leach, resin-in-pulp process

reactors, as it now is for nuclear weapons. The nuclear industry will always be dependent on a continuing supply of uranium from ore.

Mining

Uranium mines are primarily the open pit type, but there is significant production from deep mines as well as from solution mining. Sometimes uranium is produced as a by-product of mining operations for vanadium, phosphate, copper and gold.

Where uranium ore deposits occur in permeable aquifers with low-permeability geologic formations above and below them, the uranium is extracted by circulating a carbonate leaching solution. The leach solutions are sulfuric acid, ammonium carbonate, or sodium carbonate–bicarbonate along with air, peroxide, or sodium chlorate as the oxidizing agent. Holes are drilled to the bottom of the ore body on a grid pattern, 50–200 ft apart. The holes are cased with screened sections through the ore body. Varying combinations of injection and withdrawal of the leaching agents are applied to recover the uranium. The uranium is recovered from the leaching agent by ion exchange. The leachate is adjusted and recycled.

Milling

In milling, the uranium is leached from the ore, separated from major contaminants, and converted to yellow cake [or sodium diuranate $(Na_2U_2O_7)^-$]. The uranium ores milled in the United States generally contain 0.05–0.2% U_3O_8, with an average of 0.1%. To save transport costs, the mills are located near the mines. There are two major processes, one based on a carbonate leachant and the other on sulfuric acid (see Fig. 21.12.).

Carbonate process. In this process the ores are leached with hot sodium carbonate for 24 h, and sparged with air to

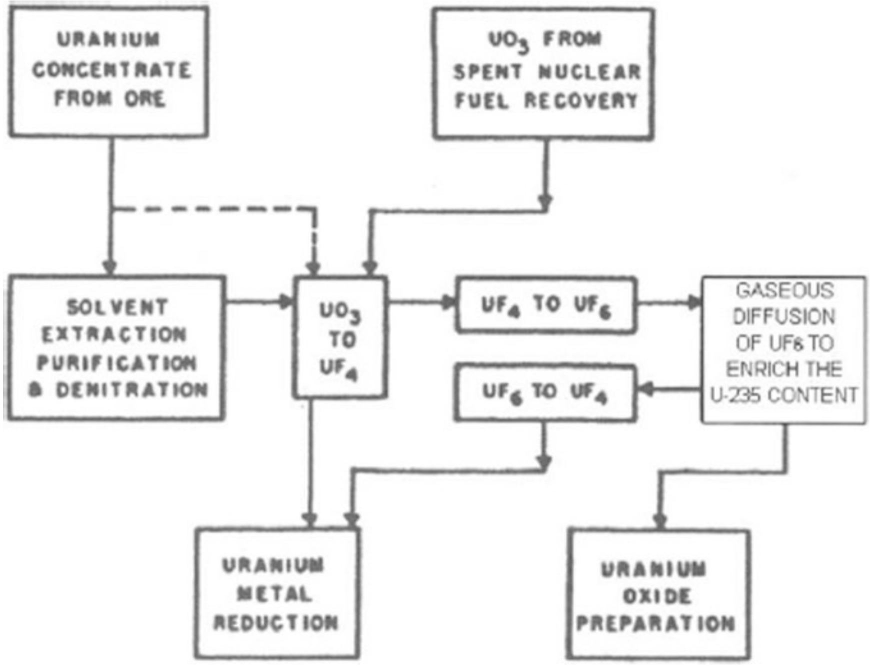


Fig. 21.13 Uranium feed materials flow sheet

provide oxidation. The leachate is cooled in countercurrent heat exchangers, heating the carbonate solution for the next batch. The carbonate leachate is filtered on rotary drums, and the uranium is precipitated with sodium hydroxide and filtered. The filtrate is converted back to carbonate by sparging with carbon dioxide, usually from a boiler flue gas, and recycled. The dried precipitate in the form of sodium diuranate is about 90% U_3O_8. Carbonate leaching is fairly specific for uranium, and the product is fairly low in contaminants.

Sulfuric acid process. In the sulfuric acid process, the pulverized rock is leached for about 8 h with sodium chlorate or manganese dioxide added for oxidation. The leachate passes through a series of thickeners countercurrent to the flow of the wash water. The liquid (leachate) is then separated from the solids using filtration. The acid leachate is not very specific for uranium and therefore goes through ion exchange or solvent extraction for purification. The solvent is a hydrocarbon-diluted 2-diethyl-hexyl phosphoric acid applicable to the recovery of both vanadium and uranium when both are present in the ore. A variation is the RIP (resin-in-pulp) process. After thickening, to avoid the filtration step, the slimecontaining leachate goes into tanks where screen baskets of anion exchange resins, mechanically sloshed up and down, absorb the uranium. After loading, the uranium is stripped from the resin with an acidified chloride or nitrate solution. These product streams are precipitated with sodium hydroxide and the resulting sodium diuranate is dried. An additional cycle of ion exchange may be required to achieve the desired product quality.

Fuel Preparation

There are two general processing routes for making feed for the nuclear fuel manufacturing plants. The primary route is through solvent extraction for purification, denitration to UO_3, and reduction with hydrogen to UO_2, followed by either hydrofluorination to UF_4 and reduction to metal, or fluorination to UF_6 followed by isotopic separation and reduction to UO_2. The other route starts with higher grade yellow cake, skips solvent extraction, and goes directly to hydrofluorination, with distillation of the final hexafluoride added for purification.

Uranium fuel preparation takes the UF_6 and is converted to either (a) aluminum-clad uranium metal for the weapons plutonium production reactors or (b) to Zirconium-clad UO_2 for electricity production in the light and heavy water power reactor (see Fig. 21.13).

Solvent extraction. The yellow cake is dissolved in nitric acid and extracted from this aqueous phase by 5% tributyl phosphate (TBP) in a hydrocarbon diluent. The diluent reduces the density and viscosity of the TBP, enhancing the aqueous/solvent phase separation. The extraction is very specific for uranium, with separation factors of 10^3 to 10^5. The product thus obtained is an aqueous uranyl nitrate solution (Fig. 21.14).

Denitration. The uranyl nitrate solution from solvent extraction is converted to UO_3 by evaporating the solution to a final boiling point of 120–140°C, followed by calcination at 620°C. The product characteristics are dependent on the type of calciner (pot, trough, or fluid bed) and significantly affect the subsequent steps (see Fig. 21.15).

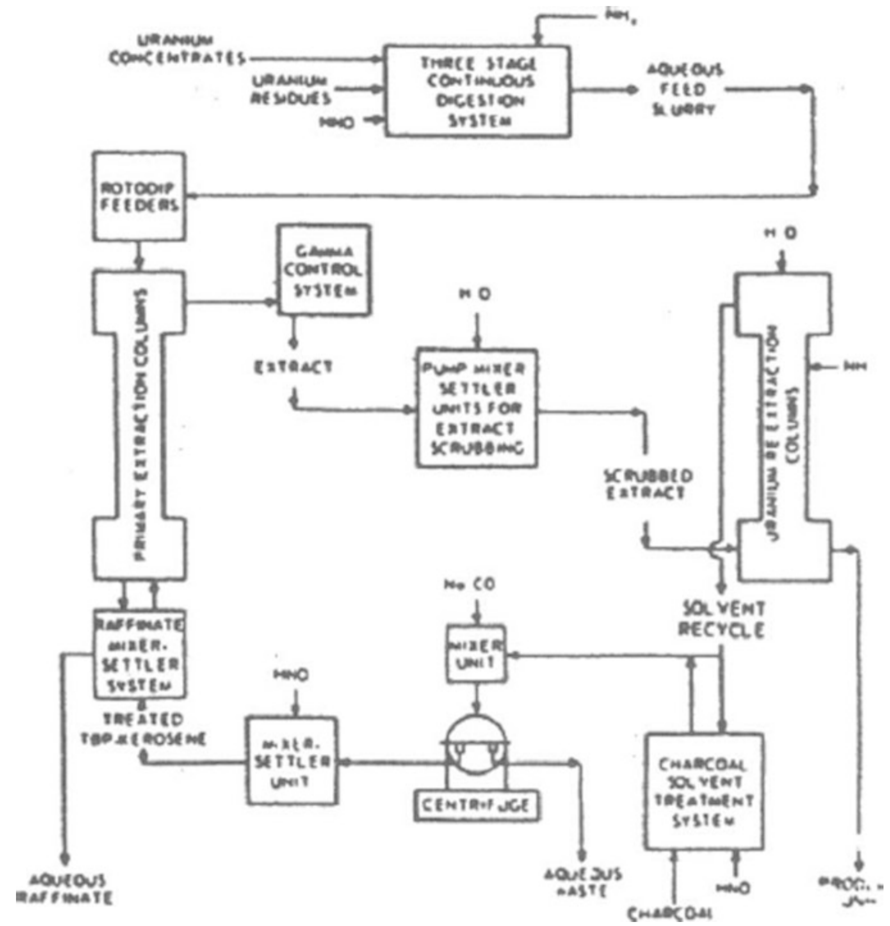


Fig. 21.14 Uranium solvent extraction process for purification of ore concentrate and scrap; slurry feed eliminates clarification cost and losses (Courtesy of USDOE)

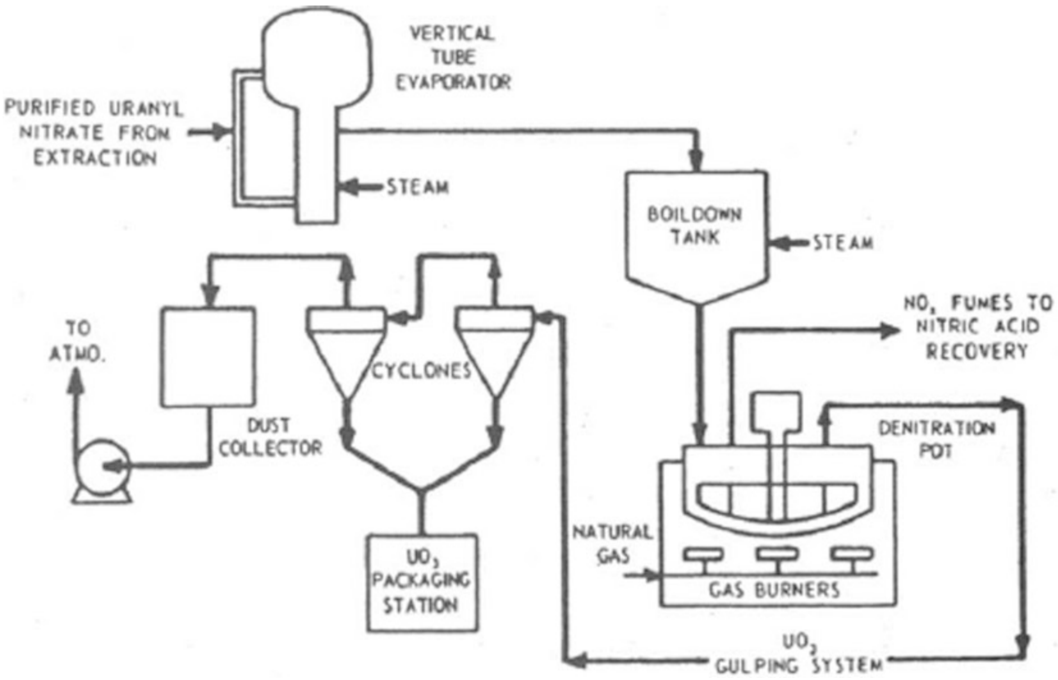


Fig. 21.15 Uranium denigration process prepares UO_3 for conversion to metal and UF_6. A continuous fluid-bed process has been developed (Courtesy of USDOE)

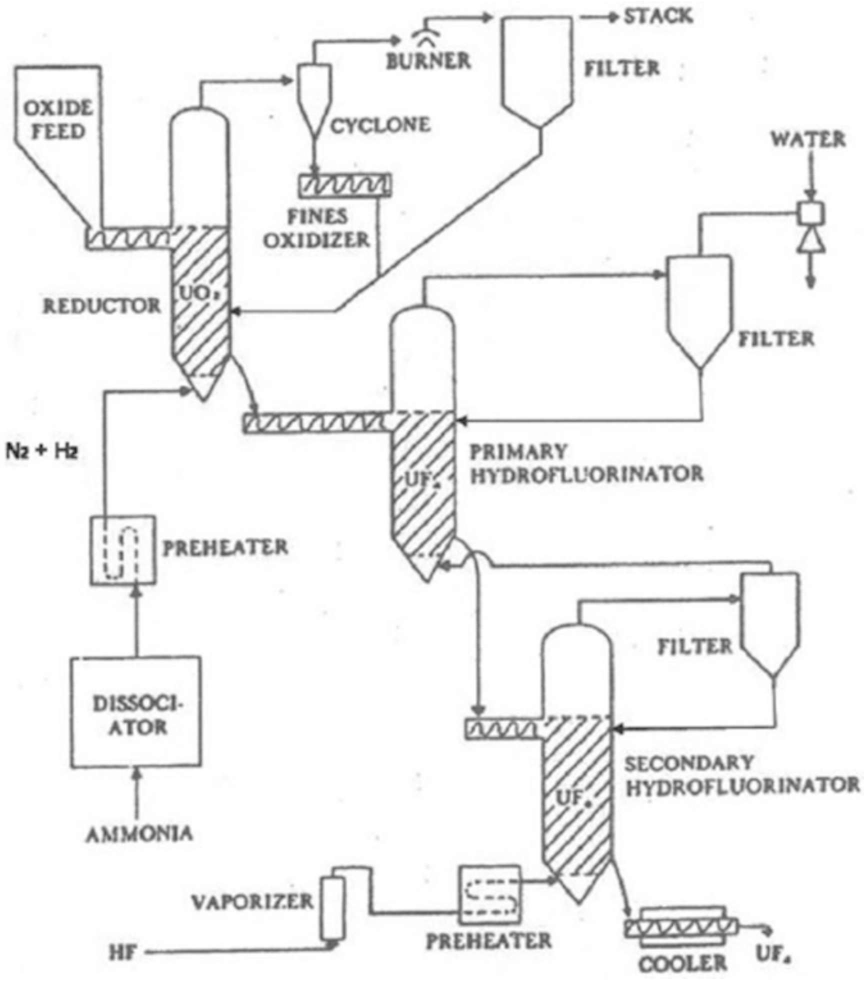


Fig. 21.16 Fluid-bed system for UO_2 conversion to UF_4 (Metropolis process)

Conversion to uranium hexafluoride. The calcined UO_3 is reduced to UO_2 with hydrogen, then converted first to UF_4 with hydrogen fluoride and finally to UF_6 using fluorine. These steps usually are carried out in fluidized bed reactors, but the conversion to UF_4 sometimes is done in a stirred trough reactor [34].

The fluid-bed process at the Honeywell Metropolis Plant uses a series of three sets of fluidized beds (see Fig. 21.16). The first bed is fluidized with hydrogen and reduces the UO_3 to granular UO_2. Then it is hydrofluorinated to a granular UF_4 in a two-stage fluid bed to achieve efficient consumption of the hydrofluoric acid. Finally, in two parallel fluidized bed reactors, the granular UF_4 burns in a fluorine atmosphere to UF_6 vapor. The fluorine is fed in through the ash at the bottom of the reactor to maximize the uranium yield, and then exits through the UF_4 screw feeder to minimize the loss of the high-cost fluorine.

UF_6 purification. Distillation of the UF_6 is required for purification when the yellow cake does not go through solvent extraction. At Metropolis, the molybdenum, vanadium, and other impurities are removed in a pair of columns, one a 120-ft, 100-plate column operating at 200°F and 85 psia to remove the high volatiles and the second a 45-plate column operating at 240°F and 95 psia to remove the low volatiles. The UF_6 is condensed into 10-t (48b) cylinders and delivered to the enrichment plant.

Reduction to metal. Uranium metal is produced by bomb (high-temperature chemical reactor) reduction of UF_4 with magnesium metal. In this process, granular UF_4 is blended with magnesium metal pellets and tamped into a steel reactor lined with the reaction by-product, magnesium fluoride (see Fig. 21.17). After the container is capped, it is placed in a furnace where the temperature is raised to the ignition temperature at which the magnesium and UF_4 react spontaneously. The reaction mass reaches a temperature sufficiently high for the liquid uranium metal to form a puddle within the bomb. This is a very empirical technology, dependent on the control of many physical factors.

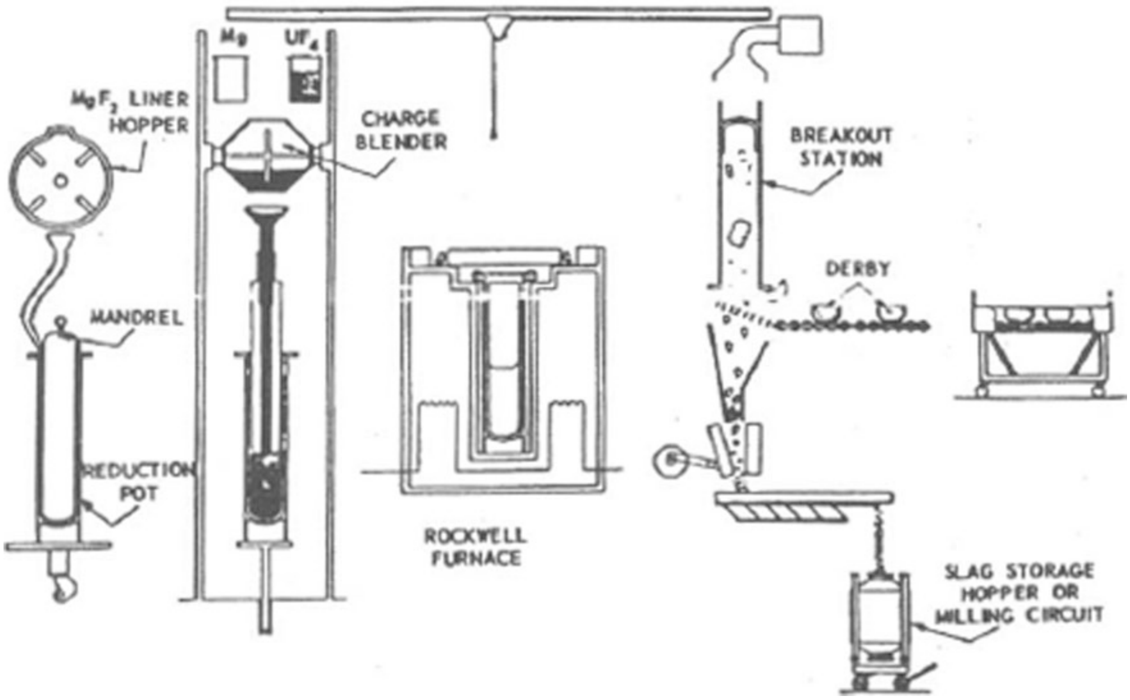


Fig. 21.17 Uranium metal reduction process; a similar process is used for plutonium metal production (Courtesy of USDOE)

Failure to maintain control of the process parameters will produce dispersed uranium shot instead of the desired regulus.

The resulting 350-lb uranium regulus, called the derby, is broken out, remelted in a vacuum furnace, and held at 1,454°C to volatilize and remove the impurities. It then is recast in graphite molds to produce the ingot. This is formed into 1- to 2-in.-diameter rods by extrusion and rolling-mill operations followed by machining and cladding with aluminum. Its primary use is for plutonium production; however, some of the depleted metal is used for shipping-cask shields, military projectiles, and counterweights. The magnesium fluoride by-product is ground and screened to provide material for lining the metal reduction bombs. Excess MgF_2 is disposed of as a low-level waste.

Uranium enrichment. Enrichment of uranium-235, from 0.711% as present in natural uranium, is essential to the economical operation of LWRs where the fuel life is a function of the enrichment. With approximately 4.95% U-235 fuel, the pressurized LWRs produce an average of about 55,000 thermal megawatt-days of energy/metric ton of uranium during the 4.5 years the fuel is in the reactors. In the naval submarine and ship reactors using highly enriched U-235, the fuel life exceeds 10 years. During the original development, gaseous diffusion was selected over electromagnetic separation and thermal diffusion to separate uranium isotopes. Enrichment represents about 15–25% of the production cost (fuel plus operations and maintenance cost) of nuclear-generated electricity. Gas centrifugal separators are now used in newer enrichment plants. Laser technology is now in various stages of development that might reduce this cost by a factor of 4. Separation in the ultra-high-speed centrifuge depends on isotopic mass difference, as does the jet nozzle [35]. The last-named device is a pneumatic cyclone with a high-velocity gas flow induced by a high differential pressure.

Laser-induced separation is being developed, based on differential activation energies of the uranium hexafluorides. This latest enrichment technology uses laser irradiation of the hexafluoride vapor to selectively decompose the U-235 F_6 and precipitate U-235 F_5. This process promises to recover much of the U-235 that remains in the tails from gaseous diffusion plants. Although this process is not yet commercial (as of June 2012), it is very advanced in its development stage. The only process that is currently developing this technique is the SILEX process in Australia. GE is presently testing this process at their Wilmington, NC plant.

Gaseous diffusion. In the gaseous diffusion process, the UF_6 flows through a porous nickel membrane called the barrier.

The heavier U-238 F_6 flows more slowly than the U-235 F_6, and the theoretical separation factor for an equilibrium stage is:

$$\alpha = \sqrt{\frac{{}^{238}\mathrm{UF}_6(352)}{{}^{235}\mathrm{UF}_6(349)}} = 1.0043$$

The optimum conditions require elevated temperature and reduced pressure with a pressure differential across the barrier sufficient to transport half the UF_6 flowing through that stage. However, by varying this flow ratio, the stages are reduced in size as the enrichment of the U-235 is increased. This practice, called tapering, is very important because a large number of stages are required. In the largest stages, axial flow compressors driven by electric motors rated to 3,300 hp transport the UF_6, and 640,000 kg of this material is circulated to produce one separation work unit (SWU).

When it was operating at full capacity, about 150 mt of UF_6 was fed to the plant daily in the process described here. Tapering reduces the time required for the plant to come to equilibrium.

The Department of Energy's gaseous diffusion plant in 1980 had 10,812 stages, consumed 6,000 MW of electric power and 1,350 million gallons of water a day, and took months to come to equilibrium. The building housing the plant had a combined floor area of a square mile. Plans to further expand the plant using gas centrifuges were scrapped in the 1980s when the expected growth of nuclear power was not realized. In addition, other countries and companies, including Russia and URENCO (which services such countries as the United Kingdom, Holland, and Germany), now offer such services [36].

The capacity of processes employing diffusion and mass separation is reported on the basis of the work required, and is expressed as SWUs. (The SWU is a measure of the work required to separate uranium of a given U-235 content into two components, one having a higher and the other a lower U-235 content.) It takes 5.3 SWU to enrich natural uranium to 3.44% U-235 (see Table 21.11). The Department of Energy facility has the capacity of 11.3 million SWU/year, and there is a world capacity of about 60 million SWU/year

Table 21.11 Separation work unit requirements (SWU/kg product)[a]

Percent U-235	SWU
0.5	0.17[b]
0.6	0.11[b]
0.711	0.00
1.0	0.4
2.6	3.4
3.4	5.3
4.0	6.5
5.0	8.9
20.0	45.7
98.0	270

[a] Tails 0.2% U-235

[b] Added for depleted fuel. Reference cost $112/SWU/kg

Uranium hexafluoride conversion to oxide. There are several processes used commercially to make nuclear fuel. A diagram of the conversion and fuel manufacturing process is shown in Fig. 21.18. The oldest is called the ADU process. In this process, the UF_6 is hydrolyzed in water to a UO_2F_2 solution which is mixed with an aqueous ammonia solution to precipitate ammonium diuranate and then calcined in a hydrogen and steam atmosphere to UO_2. Another process is called the AUC process. In this process, UF_6 is injected into a solution of ammonium carbonate to form an uranyl tricarbonate precipitate. This precipitate is then calcined in a fluidized bed using steam and hydrogen to UO_2 powder. In both the ADU and AUC processes, a liquid waste of ammonium fluoride is produced that is treated using lime to produce a calcium fluoride waste and recycle ammonia. Disposal of the calcium fluoride is sometimes difficult depending on how contaminated with uranium it is. In order to get around producing a solid waste that is hard to decontaminate, the IDR process was developed. In this process, UF_6 and steam are combined in a nozzle to form UO_2F_2 powder. This UO_2F_2 powder then drops into a calciner where it is calcined to UO_2 in a hydrogen and steam atmosphere. The advantage of this process is that the waste stream is only HF which can be disposed of (or sold) much more readily than the wastes from either the AUC or ADU processes. This is because the HF can be easily decontaminated (separated from any residual uranium).

The UO_2 powder produced from these conversion processes is then mixed with U_3O_8 recycle material, die lubricants, and other materials and then granulated, pelletized, and fired in a hydrogen furnace to achieve high density (see Fig. 21.18.). Typical densities are >95.5% of the theoretical UO_2 density (10.96 g/cm^3). The pellets are then ground to size in wet centerless griders and inspected for chips, cracks, etc. Rejected pellets and grinder sludge are oxidized in air furnaces at about 400°C to U_3O_8 which is then recycled to the pelleting operation.

As with most industrial operations, uranium-bearing wastes as well as excess uranium scrap is produced during the course of normal operations. Because of the value of the

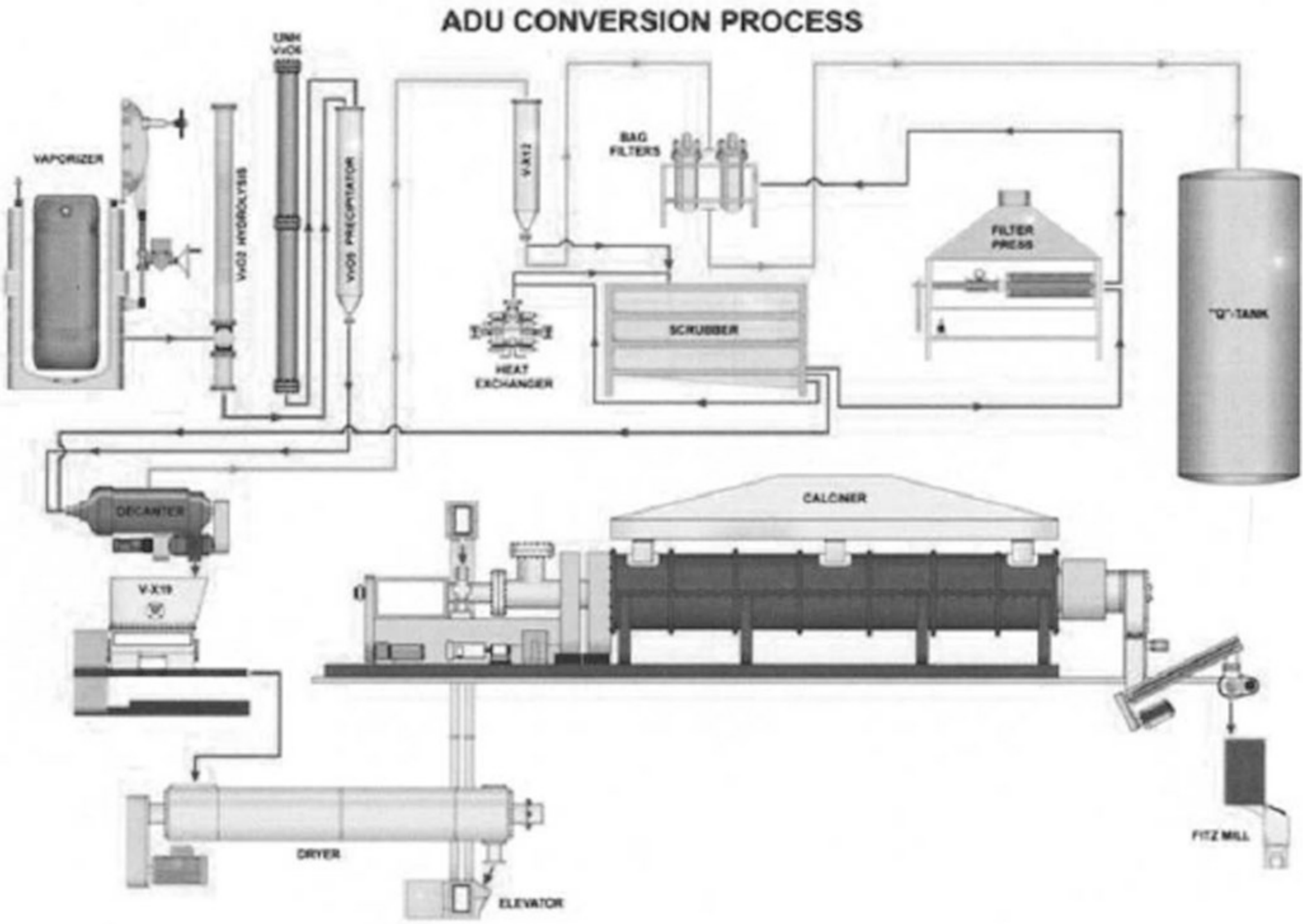


Fig. 21.18 UF_6 conversion to commercial UO_2 nuclear fuel (Courtesy of Westinghouse Electric, LLC, Nuclear Fuels Division)

uranium in this waste and scrap, it is processed to recover the uranium. The scrap and waste is treated with nitric acid to dissolve the uranium. This uranium nitrate solution is then treated using solvent extraction to recover the uranium from the other waste components. The raffinate from this waste is then neutralized with lime to precipitate any metals and then discharged. The resulting uranyl nitrate is processed in either an ADU or AUC process.

The oxide pellets are then ground to size and loaded into zirconium tubing to form fuel rods. There are approximately 80 miles of these rods in a nuclear power reactor. Another method, called the Cezus process, involves dissolving purified (Zr, HF)Cl_4 in $KAlCl_4$, which is then distilled to separate the zirconium from the hafnium. This method has the advantage of not requiring a second chlorination step. However, the compounds are highly corrosive and capital and maintenance cost are high.

Spent Fuel Reprocessing

At this time (2012), it is not know how the used Nucler fuel rods will be treated. An earlier plan of burial in Yucca Mountain (near Las Vegas, NV) was rejected. The spent fuel still contains a large percentage of fissionable material (see Table 21.12 for an example from a 32,000 MW days/t fuel) because the buildup of fission products has effectively poisoned the fuel by capturing too many neutrons. Another option instead of disposal is to reprocess the spent nuclear fuel. Most of the reprocessing to date has been performed for the weapons program. Currently, there are plants in England, France, Russia and India for commercial reprocessing. Japan is building a plant. Reprocessing provides more efficient use of natural uranium for the uranium cycle. Note that reprocessing does not reduce the activity of the fission products in the spent fuel but concentrates it into a high-level waste.

High-level waste reprocessing is the most hazardous operation in the nuclear industry. It is there that the largest quantities of fissionable and radioactive nuclides are handled in aqueous solution. These large-scale operations require both remote control and remote maintenance of the plant to protect the workers from radiation. In addition, the air and water effluents along with the solid refuse must be closely monitored to assure that the public is protected. Finally, the fissionable material requires strict accountability to ensure that it is not diverted to unauthorized uses.

As nuclear safety is a foremost public concern, it is well to note that the operation of the plutonium recovery plants since 1945 has met design specifications. There have been no significant health effects on either the workers or the public. The recovery plants were designed with large underground tanks for interim storage of the aqueous waste. This waste is then processed to an oxide, blended with glass-forming additives, and then melted into a glass mass (log) within a metal canister. The use of glass to encase the high activity waste has been chosen for defense waste in the United States and for defense and commercial waste in the rest of the world. Cement/grout or glass has been picked for stabilizing the mid- and low-activity waste that results from this processing. It is proposed that the canisters containing the high-level waste will then disposed of for the long term (hundreds or thousand to millions of years) deep underground in geologic formations. In the United States, the proposed geologic formation was at Yucca Mountain in Nevada. However, the status of this site as a repository is not known at this time due to political considerations. Other counteries such as the UK, France and Japan have also not yet identified deep geological reposition.

Table 21.12 PWR and BWR fuel burnup (late 1980s)

		Fuel composition (g/mt)		
Fuel burnup (Mwd/mt)		U-235	Pu-239	Pu-241
PWR–plutonium withdrawal				
Charge	–	32,500	0	0
Discharge	32,000	8,359	5,327	1,213
PWR–plutonium recycle				
Charge	–	6,869	19,849	4,082
Discharge	32,000	2,919	8,765	4,647
BWR–plutonium withdrawal				
Charge	–	25,000	0	0
Discharge	27,000	6,403	4,808	1,034

The main separation process in spent fuel reprocessing is the Purex (plutonium–uranium extraction) process. Fig. 21.19 shows a generalized flow diagram for nuclear fuel recovery via solvent extraction [37]. The solvent is TBP in a hydrocarbon diluent. The process was first used at the Ames Laboratory for uranium purification, then at Oak

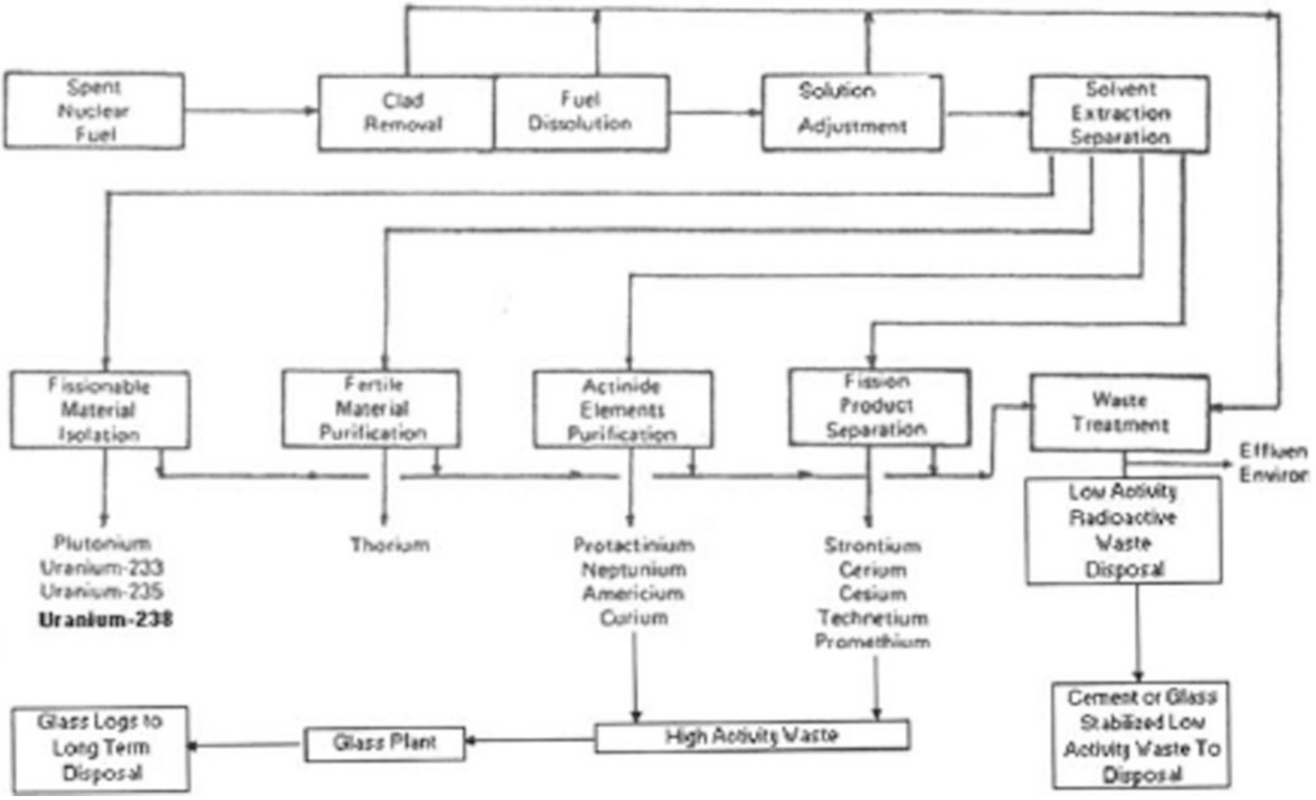


Fig. 21.19 General solvent extraction process for spent fuel recovery

Ridge National Laboratory for spent fuel. Although other processes were used in earlier days, the Purex process, with various modifications, is now used for many chemical separations in the uranium fuel cycle. It first was used on a large scale to recover uranium from the bismuth phosphate waste generated by the original plutonium recovery operation. This technology was shared with the world at the first Atoms for Peace Conference in Geneva in 1955.

A typical Purex process includes the following.

1. Aging the spent fuel to reduce by radioactive decay the 8-day iodine-131 that would be released to the atmospheric during dissolution
2. Shearing the Zircaloy-clad fuel rods into approximately 1 in. lengths so the UO_2 can be dissolved
3. Dissolving the UO_2 in nitric acid
4. Solvent extraction to separate the plutonium, uranium, and fission products
5. Anion exchange to clean up the plutonium
6. Sorption to clean up the uranium

Modifications to this process can be made to effect recovery of neptunium, americium, curium, californium, strontium, cesium, technetium, and other nuclides. The efficient production of specific transuranic products requires consideration of the irradiation cycle in the reactor and separation of intermediate products for further irradiation.

The facilities and equipment are designed for worker and public protection and for accident avoidance. Also, provision is made to allow remote replacement, decontamination, repair, and disposal. The instrumentation used in the process is designed to measure and report all the independent and dependent parameters essential to controlling and monitoring the operation. Management of the defined standard operating procedures, including quality control and quality assurance, is dictated by international standards and safety regulations.

Dissolution. The spent fuel is dissolved in nitric acid to prepare it for solvent extraction. This is a batch process that is made complex by the fuel cladding and radioactivity. A typical fuel assembly for a PWR contains about 250, 0.37-in. zirconium-clad rods 12-ft long containing in total about 1,100 lb of UO_2. The assembled rods are mechanically sheared into 1 to 2 in. lengths and dropped into a perforated basket. This is placed in the dissolver tank where the uranium oxide is dissolved as discussed above. The basket then is removed from the dissolver, and the zirconium hulls and associated hardware therein are dumped into containers, compressed, sealed, and sent to transuranic waste storage. Ultimately, they will go to a disposal facility.

Separation and purification. In the Purex process discussed here, the uranium, plutonium, and fission products are separated by solvent extraction into three different streams (Fig. 21.20). The plutonium stream goes through anion exchange (discussed later) to reduce traces of ruthenium, and the uranium stream goes through silica gel sorption to reduce traces of zirconium. The fission-product stream, which contains the fission products with the transplutonium products, is sent to waste treatment. About 99% of the uranium and plutonium is recovered in separate product streams, and decontamination from fission products by a factor greater than 10^7 is effected [38]. The spent fuel solution from the dissolver, adjusted to 2 M nitric acid, flows into the middle of the first column. From there it flows downward countercurrent to a 5% TBP/95% hydrocarbon solvent, which is introduced into the bottom section. The uranium and plutonium transfer to the solvent, leaving most of the fission products in the acid phase, which pass out the bottom of the column. The U/Pu solvent solution is scrubbed in the top half of the column with 3 M nitric acid to wash out additional fission products.

The valence of the plutonium in the U/Pu solvent stream is reduced from +6 to +4 with hydrazine to lower its extractability, and the stream flows to the middle of the second column, where the plutonium is stripped out of the solvent with mild acid. The uranium passes on in the solvent stream to the third column, where it is washed from the solvent with a weak acid. The fission product stream flowing from the bottom of the first column, containing less than 1% of the U and the Pu, is the high-level waste. It is evaporated to remove most of the acid, and neutralized with sodium hydroxide before going to earth-covered storage tanks in the United States. In Europe, the high-level waste steam is oxidized, mixed with glass formers, and disposed of in glass logs.

The uranium and plutonium streams flow separately through second sets of two columns for re-extraction and stripping for further separation from fission products, according to the scheme shown for Pu in the lower portion of Fig. 21.20. The acid waste streams from the second cycles may be recycled to the first to eliminate the second cycle losses.

Centrifugal contractors or pulsed columns are used for these solvent extraction operations in preference to longer packed columns because the latter would complicate the shielding problem. Longer columns increase the exposure time and therefore the decomposition of the extractant and solvent due to the radiation field. Mixer-settlers also can be used. The solvent is recycled after treatment to remove decomposition products caused by radiation and chemical effects. This treatment may include a combination of distillation with acid and caustic washes.

From the second cycle the plutonium goes through anion exchange for final purification (Fig. 21.21). The principal problem here is due to ruthenium, which is difficult to remove because of its many valence states. The uranium stream goes through silica sorption primarily to remove zirconium, which seems to be carried along as a colloid.

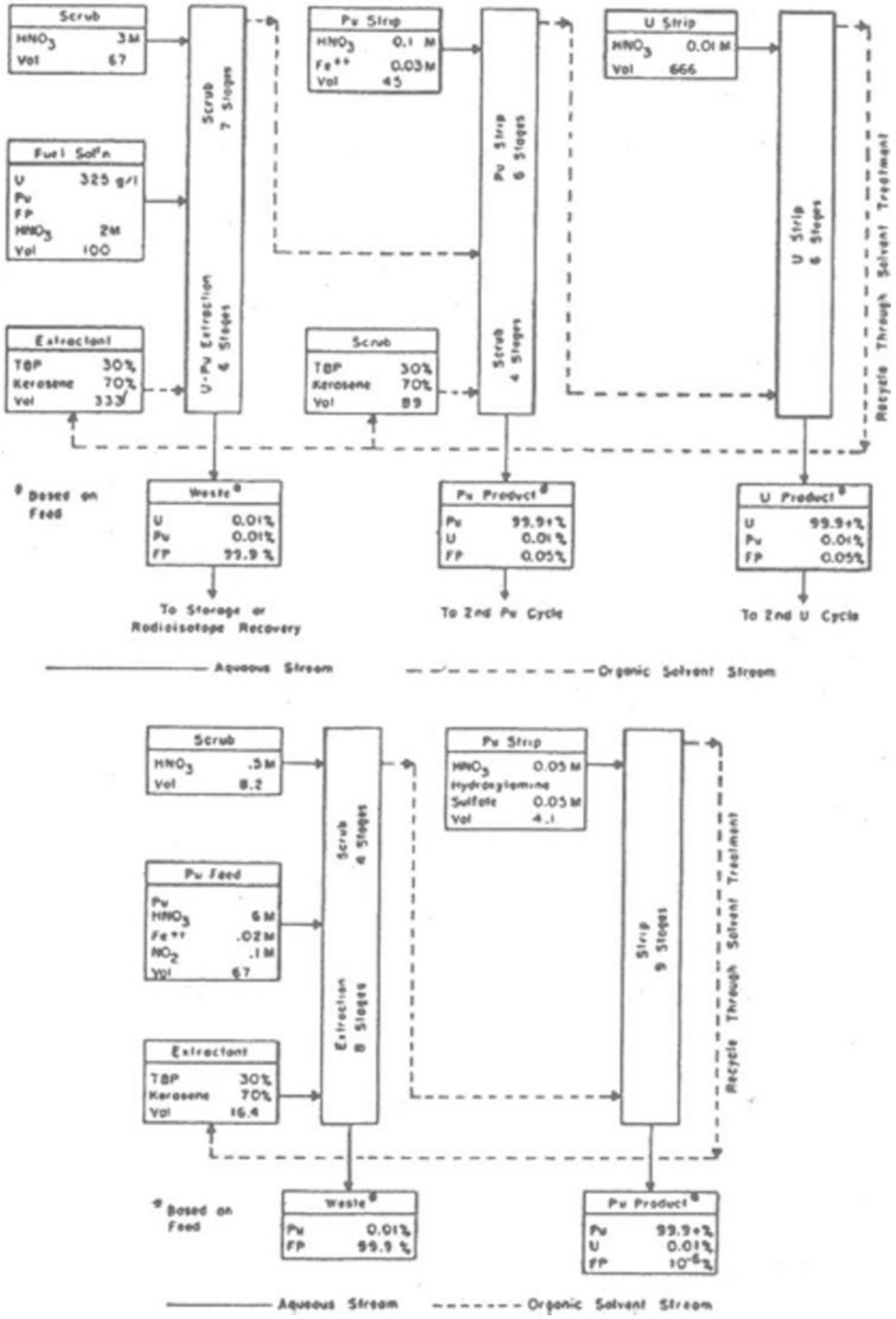


Fig. 21.20 Purex process showing (top) first solvent extraction cycle, and (bottom) second plutonium solvent extraction cycle

The process is installed in buildings called canyons, where there are rows of concrete shielded cells serviced by overhead gantry cranes for remote replacement and servicing of valves, pumps, piping, and other equipment. Penetrations through the cell walls are offset to prevent radiation from streaming through. Some equipment is located in cells equipped with windows, TV cameras, and manipulators. Cells for remote repair of the equipment may be similarly equipped.

It is occasionally necessary for personnel to enter some of these facilities. This requires considerable time for decontamination and placement of local shielding. The operations that will be performed are carefully planned and rehearsed in order to get the entry over with as swiftly as possible. A special group of employees, the health physicists, are responsible for radiation safety throughout the plant, and they provide close monitoring of these entries.

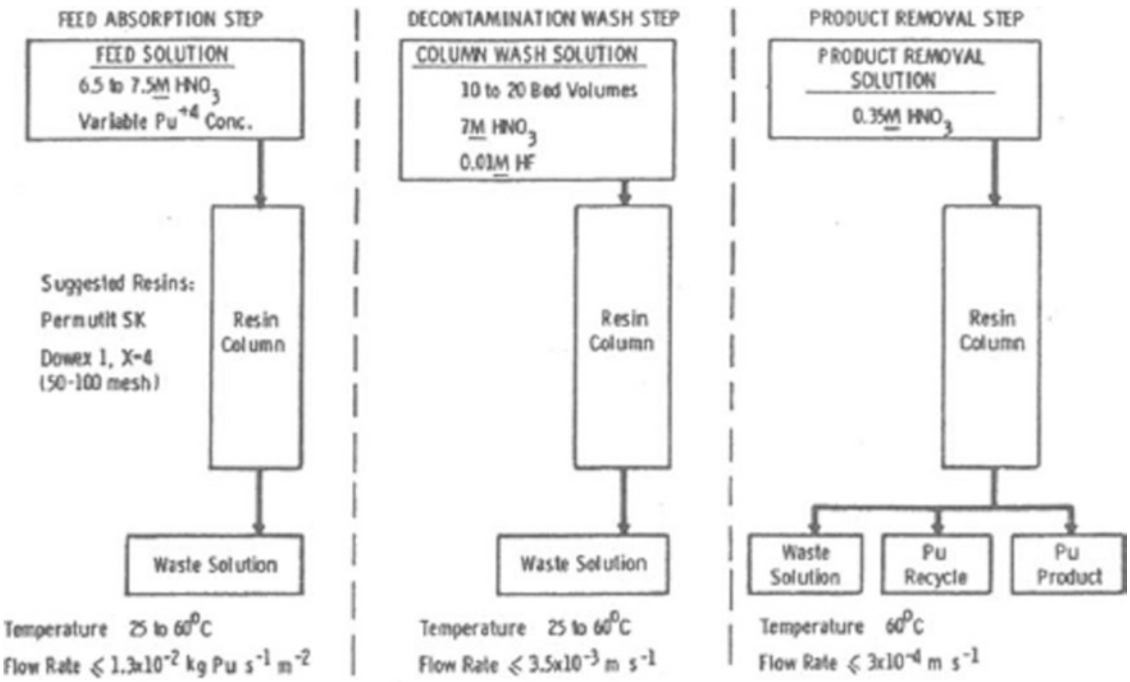


Fig. 21.21 Plutonium ion exchange flow sheet

Another unique consideration is the prevention of nuclear criticality, i.e., the start of a chain reaction leading to a self-sustaining nuclear reaction within the cells. In the dissolver and first cycle, criticality is prevented by the presence of the uranium-238, which absorbs neutrons. Later in the process, where the plutonium is separated from the uranium, criticality is prevented by proper design of the vessels and piping. This includes the cell floor and sumps, where materials would collect in case of leakage from the equipment. To prevent criticality, the vessels are limited in either diameter or thickness. Vessels and piping are placed in arrangements designed to avoid a critical array. In addition to design considerations, the risk of criticality can be lowered by limiting the amount of total fissile material in any given vessel, and when possible, by choosing a medium which is not an appropriate moderator for fission.

In one modification of the Purex process, the plutonium is not separated from the uranium. In this version, the first cycle has only two columns instead of three. In addition to reducing the criticality risk, this modification reduces the risk of unauthorized diversion of the plutonium.

Radioactive Waste Management

Radioactive waste management involves the treatment, storage, and disposal of liquid, airborne, and solid effluents from the nuclear industry's operations, along with those from other activities that employ the radioactive products. Its strategy involves four approaches: limit generation, delay and decay, concentrate and contain, and dilute and disperse. Combinations of all four of these usually are employed to manage each waste stream [39].

There are three types of nuclear wastes, based on their radionuclide characteristics:

- Uranium-contaminated waste, principally from mining, (the "Tails" (Non-Uranium containing minerals contain radium and radon as uranium decay products)) milling, and enrichment [40]. Under some circumstances uranium's 15-h radon decay product requires additional control. The uranium tails from enrichment generally are not included in waste tabulations because in time they may be incorporated as a fertile material in breeding.
- Plutonium-contaminated waste, principally from weapons materials processsing. There are two classes of plutonium wastes: the transuranic, containing mostly plutonium and other transuranics, and the high-level wastes that contain significant heat-generating products, strontium, and cesium.
- Other radionuclide-contaminated waste, generally called low-level waste, as generated in the reactor and fuel cycle, as well as in the radionuclide applications (see Tables 21.13 and 21.14.).

Limiting the generation of waste is the first and most important consideration in managing radioactive wastes. The Purex process was developed to eliminate the solids additions to the -high-level wastes in earlier technology. Improved housekeeping procedures have greatly reduced the low-level and transuranic wastes generated in the power plants, defense plants, and laboratories. Delay and decay is frequently an

Table 21.13 Radionuclides in a PWR spent fuel assembly[a]

Nuclide	Half-life (years)	mpc (Ci/L)[b]	Curies: years after discharge			
			10	100	1,000	10,000[c]
Zr Clad						
Fe 55	2.7	8E-7	180	10^{-8}	–	–
Co 60	5.3	5E-8	970	0.01	–	–
ZR 93	9E5	6E-8	0.10	0.10	0.10	0.10
UO_2 fuel—fission products						
Kr 85	10.4	3E-10	2,300	6.8	10^{-25}	–
Rh 106	2.2 h	–	270	10^{-25}	–	–
Pd 107	7E6	–	0.05	0.05	0.05	0.05
I 129	2E7	6E-11	0.01	0.01	0.01	0.01
Xe 133	0.014	3E-10	–	–	–	–
Cs/Ba 137 m	30	2E-8	74,000	9,300	10^{-5}	–
UO_2 fuel–transuranium products						
Ra 226	1,620	3E-11	10^{-7}	10^{-5}	10^{-3}	10^{-1}
Np 237	2E6	3E-9	7.6	7.6	7.0	3.1
Pu 238	89	5E-8	1,000	500	0.5	0.06
Pu 239	24,360	5E-8	142	142	139	108
Pu 240	6,580	5E-8	236	235	214	85
Pu 241	13	2E-7	36,000	500	0.01	0.003
Am 241	458	4E-9	770	1,700	410	0.005
Cm 244	17.6	7E-9	480	15	10^{-14}	–

[a] Characteristics: 33,000 MWd/tU burnup, 14.7 kg U^{235}, 446 kg U^{238}, 108 kg Zr, 10 kg steel and Inconel

[b] Normal human consumption in water would be equivalent to a radiation dose of 500 mrem/year, approximately 2–5 times the natural radiation background

[c] After 10,000 years the radioactive toxicity is approximately constant for several million years

Table 21.14 Estimated cumulative waste quantities and land requirements: United States[a]

	Volume (10^6 ft^3)			Land (acres)		
	1980	1990	2000	1980	1990	2000
Low level						
Defense	60	80	90	540	650	680
Commercial	15	45	83	150	300	420
TRU[b]						
Defense	0.05	0.05	0.05	20	100	200
Commercial	0.01	0.01	0.01	Nil	Nil	Nil
High level[d]						
Defense (sludge and calcine)[c]	–	–	0.004	0	0	900
Defense (LLW salt cake)	–	–	0.008	Nil	Nil	15
Commercial spent fuel[d]	0.014	0.7	1.4	20	50	2,200
Disposal facilities[d]						
Near-surface burial grounds[e]	–	–	–	690	950	1,100
Deep geologic vaults	–	–	–	Nil	Nil	2,700

[a] IRG Appendix D, assuming 148 GWe in 2000 and commercial fuel not processed

[b] TRU in interim surface storage until decision made regarding disposal

[c] Packaging of defense wastes started in early 1990s

[d] Spent fuel and defense high-level waste in surface storage until decision made regarding disposal

[e] Includes scrap from decommissioned facilities

important strategy because much of the radioactivity in nuclear reactors and accelerators is very short-lived, a few minutes to a few days. Concentrating and containing is the objective of the treatment activities for the longer-lived radioactivity. The concentrate and contain strategy generally involves converting the airborne, liquid, and solid waste to stable solids in corrosion-resistant containers for storage and for transport to and emplacement in carefully selected disposal sites.

Finally, both dilution and dispersion of large volumes of air and water effluents containing very low quantities of radioactivity generally are necessary. The concentration of radioactivity in these effluents is controlled by federal and international regulations, and such effluents are continuously monitored before release to the environment from the waste treatment and other activities. In a particular operation, the regulations may be reflected by a set of actions that are triggered by successively higher levels of radioactivity.

The first action, where correction by the plant operator is called for, occurs at a small fraction of the permissible level. A somewhat higher level of radioactivity requires both operator action and regulatory notification. The next higher level requires regulatory participation, and if the final level is reached, the operation must be shut down.

Liquid Waste Treatment

Liquid waste treatment is primarily concerned with conversion of an aqueous solution to a solid form suitable for storage and disposal. Evaporation is the most effective process and achieves decontamination factors for the evaporate of 100–100,000 per cycle.

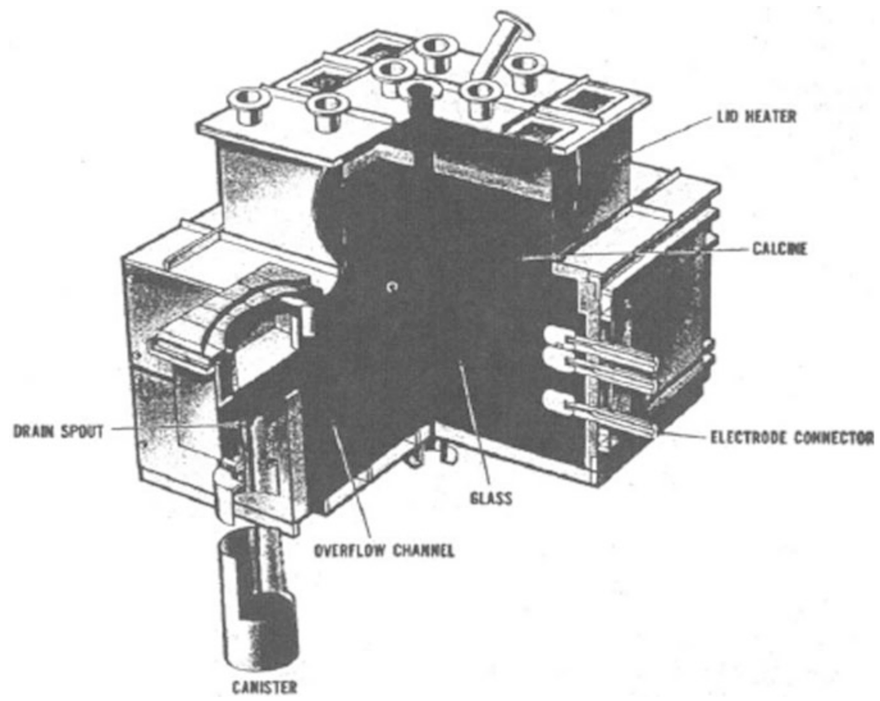


Fig. 21.22 Slurry-fed vitrification furnace converts aqueous high-level waste into a boro-aluminum silicate glass form (Courtesy of USDOE)

Ion exchange and scavenging precipitation can be employed at a lower cost than that of evaporation when decontamination factors of 5–100 are adequate.

As an example, in the US defense establishment most of the radioactivity from plutonium production is found in the liquid high-level waste from the first cycle of the Purex process. This liquid is neutralized with sodium hydroxide and stored in earth-shielded tanks. There a sludge settles out that contains most of the radioactivity. The residual liquor is partially evaporated to decrease the volume of the waste, and sodium nitrite crystallizes on top of the sludge. In a process to be used at the Savannah River plant, the sodium liquor fraction, containing most of the cesium fission product, is pumped from the tank and treated using solvent extraction. The concentrated cesium and strontium stream is then combined with the sludge that was slurried from the tank and washed to remove the salt liquor. The resulting mixture is then mixed with a pulverized glass frit and converted to a boro-aluminum silica glass in a vitrification furnace, cast in a canister, and sealed (see Fig. 21.22). This high-level waste form is then sent to a high-level repository. The sodium nitrate and the residual liquor will be combined with cement and other components to make a “Saltcrete” for long-term disposal in above-ground storage vaults.

Airborne Waste Treatment

The largest volume of radioactive waste is made up of the ventilation air from processes, plants, and laboratories. However, the quantity of radioactivity is relatively low. It generally occurs in a particulate form readily removed by dry filters. A special class of paper filters developed for this purpose (HEPA or high efficiency particulate air) removes 99.97% of 0.3-μm particles. These filters frequently are preceded by scrubbers, sorbers, and roughing filters to extend the life of the paper filter.

A notable exception to the above is the airborne effluent from spent nuclear fuel reprocessing and from tritium production. In this case, iodine-131 volatilizes from fuel reprocessing and must be removed by passing the airborne effluent through a silver catalyst bed, which provides a decontamination factor of about 100,000. Prior to reprocessing, the spent fuel is stored for a sufficient time for decay of the 8-day iodine-131, usually 10 half-lives, to make this treatment adequate. The processing effluent also carries increased levels of krypton-85, xenon and tritium, gases, and solid radioactive particulates. The particulates are removed by filtration, but the other substances generally can be released. In tritium production, cold traps and sorbers are used to remove tritium from the airborne effluents.

Solid Waste Treatment

The solid wastes from nuclear operations include the concentrates from liquid and airborne treatment along with paper, clothing, lab glassware, and scrap equipment. The transuranic and low-level wastes both receive the same but separate treatment. In plutonium operations the first step frequently is monitoring to separate the transuranic from the low-level wastes. Some of these are incinerated to reduce their volume, but this process produces additional airborne and liquid waste. Physical compaction frequently is more useful. Equipment items, including large reactor components, are decontaminated by using a series of acid, chelating, and caustic washes. Finally, the solids may be incorporated in asphalt or concrete to reduce their leachability and then are drummed. Most of these solid wastes go to low-level disposal.

From power reactors with no reprocessing, the high-level waste consists of assemblies of zirconium-clad spent fuel rods to be packaged in stainless steel canisters. If the spent fuel is reprocessed, then the high-level waste will be converted to a silicate glass form similar to that from defense operations. The uranium and plutonium content in the waste, and also possibly the strontium-90 and cesium-137, would be reduced by 90–99%. The nuclides are useful, and their removal reduces the waste disposal problem in several ways: it reduces the toxicity and heat load of the waste and also reduces the incentive for future generations to undertake the very costly recovery of what may become valuable energy-producing resources.

Storage of Spent Fuel

Spent fuel originally was originally supposed to be stored at each nuclear power reactor in water pools designed to accommodate a 5-year inventory on the assumption it would be reprocessed later. However, reprocessing did not develop, and there has been a delay in identifying the ultimate repository. The storage pools were reracked to increase their capacity, and now the older reactors also are installing dry storage capacity to accommodate the life of the reactor, approximately 60 or even 80 years, as power plant license extensions are granted [41].

The federal government was committed to begin accepting spent fuel in the late 1990s and was planning to provide away-from-reactor storage that would have included a fuel rod consolidation facility. The nuclear utilities have been assessed at 0.001/kWh electric that is expected to pay for disposal of their high-level spent fuel wastes. As is usual for all new nuclear operations, the site and the schedule for this facility was dependent on public acceptance. Because this public acceptance was not forthcoming, the political support was not available in the US Congress to make a decision on an ultimate final burial site for high-level waste. While many options have been considered (see Table 21.15), as of 2012, no site has been selected.

Table 21.15 Conceptual methods for radioactive waste disposal

Terrestrial	Shallow land burial Deep geological vaults Continental sites Island sites Caves Deep hole Sub seabed Ice cap Hydrofracturing In situ rock melting
Extraterrestrial	Solar orbit–space shuttle Moon crater–rocket/soft lander Solar escape–electric cannon
Transmutation	Fission reactors Fusion reactors Electromagnetic accelerators

It appears that the repository capacity for high-level wastes will be heat-limited to 1 kW/acre. This is equivalent to 1 t of fuel after storage for 10 years. This can be increased to about 2 tons od fuel if the uranium and the plutonium were removed, and considerably less than that if the strontium and the cesium also were removed. Although reprocessing would reduce the transuranic content by a factor of 10–50, this amount would be an insignificant fraction of the transuranic hazard in the waste.

The most restrictive of the current regulations require this waste to be contained for about 10,000 years. Intensive international R&D has been conducted, along with extensive geologic surveys by responsible scientists. A 1990 National Academy of Sciences study concluded that this requirement is beyond the possibility of proof. Extensive experimental work has been carried out, and perhaps the most interesting study is related to a group of natural nuclear reactors found by the French in Gabon, Africa. During the Neolithic time, where uranium was being deposited in a swamp, the developing ore body repeatedly achieved criticality. This occurred over a period of centuries at a series of interconnected sites. It appears that the long-lived products were closely held within the ore body. Unfortunately, the translation of this information to other geologic environments poses many uncertainties.

In 1989 a deep mined repository, The Waste Isolation Pilot Plant, was opened in a bedded salt formation near Carlsbad, New Mexico. The site has been receiving transuranic wastes since and has now disposed hundreds of tons.

Low-Level Waste Disposal

The solidified low-level radioactive wastes for which ultimate disposal must be provided fall into two broad categories: the uranium-containing wastes from milling and the front end of the fuel cycle, including enrichment; and the radionuclides from the nuclear reactors and spent fuel operations, together with the radionuclides from applications in research, medicine, and other industries [42].

The proposed technology for disposing of the second category of wastes consists of near-surface burial for low-level wastes after stabilization in cement.

Mill wastes. Agreement has been reached with the public regarding mill waste, and disposal of the accumulated backlog is nearly complete via monitored surface repositories. The uranium-containing wastes from milling are mounded and covered with earth. This earth cover prevents erosion and delays escape of the 14-h half-life radon gas, the gaseous decay product of uranium, long enough to significantly reduce released levels. They also reduce rainwater teaching and dispersion of the radium daughter decay product. These mill waste repositories are located near the mines and mills and are not a very different hazard from the original naturally occurring uranium deposits [43, 44]. The depleted uranium from the enrichment operations is stored in cylinders as uranium hexafluoride for possible future use in the uranium–plutonium breeding cycle. Other uranium-containing wastes from enrichment and fuel fabrication go to the low-level repositories.

Low-level wastes. Low-level wastes presently are being entombed in near-surface, earth-covered trenches. These trenches are designed to minimize the inflow of water and drain to a sump that is constantly monitored. Treatment facilities are maintained to process drainage that exceeds licensed levels of radioactivity for release in the groundwater.

One of the major repositories for commercial low-level waste is near Barnwell, South Carolina and is nearing its capacity. Other repositories around the country were closed when full or when required to be closed by state regulatory action. The major repository for very low-level or incidental nuclear waste is near Salt Lake City, UT and there is another located in Texas. There have been no health effects or releases of radioactivity exceeding licensed limits from these low-level repositories. The schedule for opening additional low-level repositories has been delayed for several years by the same type of public opposition encountered by proposals relating to treatment and disposal of municipal and other wastes.

Transportation of Nuclear Materials

The transport of nuclear materials reflects many years of experience based on millions of shipments. These shipments range from microcuries in liquid and gaseous forms in glass vials carefully packaged in sealed cans inside cardboard boxes to megacurie quantities contained in spent fuel packaged in hundred-ton shipping casks. They are shipped by both private and public carriers, by car, bus, truck, railroad, airplane, barge, and ship.

During these years there have been many incidents involving all forms of transportation, both commercial and military, and a few of these events resulted in some contamination of facilities. However, there have been no health effects to any individual transport worker or the general public.

Shipments of radioactive materials, originated by many different shippers throughout the world, are controlled by an established code of national and international regulations. Although most nuclear operations relate to a localized population, transport potentially affects large numbers of people. Many communities have attempted unsuccessfully to prohibit nuclear transport through their jurisdiction. Some, notably New York City, have succeeded.

Shipments of spent nuclear fuel constitute the largest quantity of nuclear materials being transported. These shipments involve 50-t truck casks with the capacity for one to several fuel assemblies. Rail and ship casks weighing about 100 t carry many assemblies. The design of these casks provides:

1. Radiation shielding to limit radiation to the workers and the public
2. Passive cooling for the heat-generating spent fuel, up to several kW per assembly
3. Assured containment in the event of the maximum accident

Current casks are designed to handle 5-year were aged spent fuel that travels by truck, rail, and ship. There are frequent ocean shipments from Japan to England and considerable rail shipment in Europe. Most of the spent fuel shipments in the United States are by truck. No cask failure has been experienced in any accident to date.

The transport of plutonium wastes to the transuranic and high-level repositories will represent a major increase in the quantities of radioactive materials shipped; therefore an extensive demographic study has been made to examine potential exposures of the public and to select routes to minimize such exposure [45, 46].

New casks are being developed to maximize the efficiency of the longer-aged spent fuel shipments and to reduce the exposure of the workers involved in loading and unloading the casks. There are expected to be two families of casks; those transporting the spent fuel from the reactors to a temporary storage facility, primarily truck casks; and those transporting the spent fuel from the storage facility to the repository, primarily rail casks. Another set of casks is being developed for the high-level defense waste glass at the Hanford and Savannah River facilities, and the commercial high-level waste at West Valley, NY.

The Nuclear Reactor

Development of the commercial nuclear power industry for electricity generation began in the early 1950s. The first commercial electric power was produced using a LWR at Shippingport, PA in 1957. By the 1970s, large numbers of power reactors had come online. In the United States all but one were LWRs, but in Canada the preference was for heavy water reactors (HWRs). There were many gas-cooled graphite reactors (GCRs) in England, France, and Russia and one in the United States. France and Russia had the first operating liquid metal fast breeder reactors (LMFBRs) operating, with several engineering test reactors in the United States. The characteristics of typical units are presented in Table 21.16. Nuclear reactors can be classified in several ways:

- *Burners* vs. *breeder*s: Burners depend on uranium from ore but may recycle the residual fissionable material from the spent fuel to reduce the ore requirement. Breeders produce more fissionable material than they consume by converting either uranium-238 into plutonium-239 or thorium-232 into uranium-233. The breeders depend on the plutonium-239 or uranium-233 from the burners to bring them into equilibrium, and this may take 30 years.
- *Choice of neutron moderator-light water (protium), heavy water (deuterium), liquid sodium or lead or molten salts, or solid (graphite)*: Except in the case of graphite, the moderator also serves as the coolant. In the case of a graphite moderator, gas (He or CO_2) or sometime water is used as a coolant.
- *Neutron energy (thermal* vs. *fast)*: The liquid metal/salt reactor operates with fast neutrons to breed using the uranium cycle. The water and graphite reactors operate with thermalized neutrons to more effectively burn the fissile material.
- *Pressurized* vs. *boiling LWRs*: The PWR transfers its energy from the fuel to an intermediate heat exchanger to generate the steam that drives the turbine. The boiling reactor generates steam within the reactor that goes directly to the turbine.

The LWR, CANDU, and graphite reactors are the major producers of electricity from nuclear energy. When compared on the basis of cost, reliability, and safety, they are competitive. The PWR is the most widely used LWR (267), but there are also many BWR's in use (84). In 2010 there were 51 CANDU's and other heavy water moderated reactors operating in the world, including Canada, where they were developed with several in other countries as well. In addition, there are 33 other types of reactors, including sodium cooled, and graphite-moderated (Nuclear News, 1th Annual Reference Issue, March 2012).

When PWRs and CANDUs are compared on the basis of uranium utilization, the CANDU requires 19% less uranium from ore on a per kWh basis (see Table 21.17). As there is now an abundance of low-cost uranium from ore and enrichment service, the electric utilities continue to favor the PWR

Table 21.16 Characteristics of nuclear power reactors

	PWR[a]	BWR[b]	CANDU[c]	LMFBR[d]	HTGR[e]	AGR[f]
Electric power (GWe)	1.1	1.06	0.74	0.23	0.33	0.63
Plant efficiency (%)	33	33	30	40	42	42
Fuel	UO_2	UO_2	UO_2	UO_2	UO_2ThO_2	UO_2
Clad	Zr	Zr	Zr	SS	C	Mg
Uranium loading (t)	99	155	130	–	95	–
U-235(5)	3.0	2.9	0.81	12[g]	40	0.75
Burnup (GWd/t)	30	28	7.3	100	110	–
SWU (t/GWey)	109	114	0	0,03	130	–
Fueling internal (years)	1.1	1	0.003	0.5	1	–
Fuel fraction	0.3	0.2	0.003	0.3	0.3	–
Power density (kW/kgU)	38	25	26	160	115	38
Coolant/Moderator	H_2O	H_2O	D_2O	Sodium	He/C	CO_2
Pressure (bars)	155	73	100	2	45	40
Temp (°C)	327	–	310	560	750	670
Flowrate (t/s)	21	15	12	3	0.4	–
Conversion ratio	0.6	0.6	0.6	1.2	0.8	0.6

[a] Zion 1 1973
[b] Peach Bottom 2 1974
[c] Bruce 1 1977
[d] Phenix 1973
[e] Fort St. Varain 1979
[f] Hinkley Point B1 1976 (advanced gas-cooled reactor)
[g] Includes Pu-239

and BWR types. Operating experience with graphite reactors in England, France, and Russia also has been generally satisfactory. However, the graphite moderator represents a large inventory of combustible material, which contributed significantly to the severity of the Chernobyl and windscale accidents. Further development of the high-temperature, gas-cooled, graphite-moderated reactor is being considered for the advanced Generation IV reactors to be deployed from 2015 and beyond. Due to its much higher operating temperature (~900°C), this type of reactor can be used to provide process heat for a variety of industries.

A new family of advanced nuclear reactors has been designed by the PWR, BWR, and CANDU suppliers, which are now (2012) being licensed and built in Finland (1), China (4) and the US (4). It is the industry's objective to use proven technology to provide reactors that are simpler and less costly to build than older designs. These reactors would have wider safety margins through the increased use of passive safety systems, relying more on reactor design and less on operator action and outside power and water resources to prevent accidents. They would be made less complex by drastically reducing the number of components necessary for operation. This design also would facilitate plant construction by requiring less concrete, fewer valves and pumps, and less piping. Most likely the future will see introduction of modular built systems that will be prefabricated at the factory. Safety considerations always will be paramount [47, 48].

Table 21.17 PWR and CANDU requirements for Uranium from ore

	mt/GWy-electric	
	PWR	CANDU
Once-through cycle	217[a]	175[b]
Pu/U recycle	157	129[c]

[a] 3.25 % U-235 fuel

[b] 0.71 % U-235 fuel (natural uranium)

[c] Pu/U recycled to CANDUs, from PWR once-through spent fuel (PWR/TANDEM)

Light Water Reactors

Pressurized water reactor. The PWR contains three coolant systems: the primary system, which removes heat from the reactor and partially controls nuclear criticality; the secondary system, which transfers the heat from the primary system via the steam generator to the turbine-driven electrical generator; and the service water system (the heat sink), which dumps the residual coolant energy from the turbine condenser to the environment. The service water is recirculated from a river, lake, ocean, or cooling tower. In the primary system (Fig. 21.23), dissolved boron is present to control

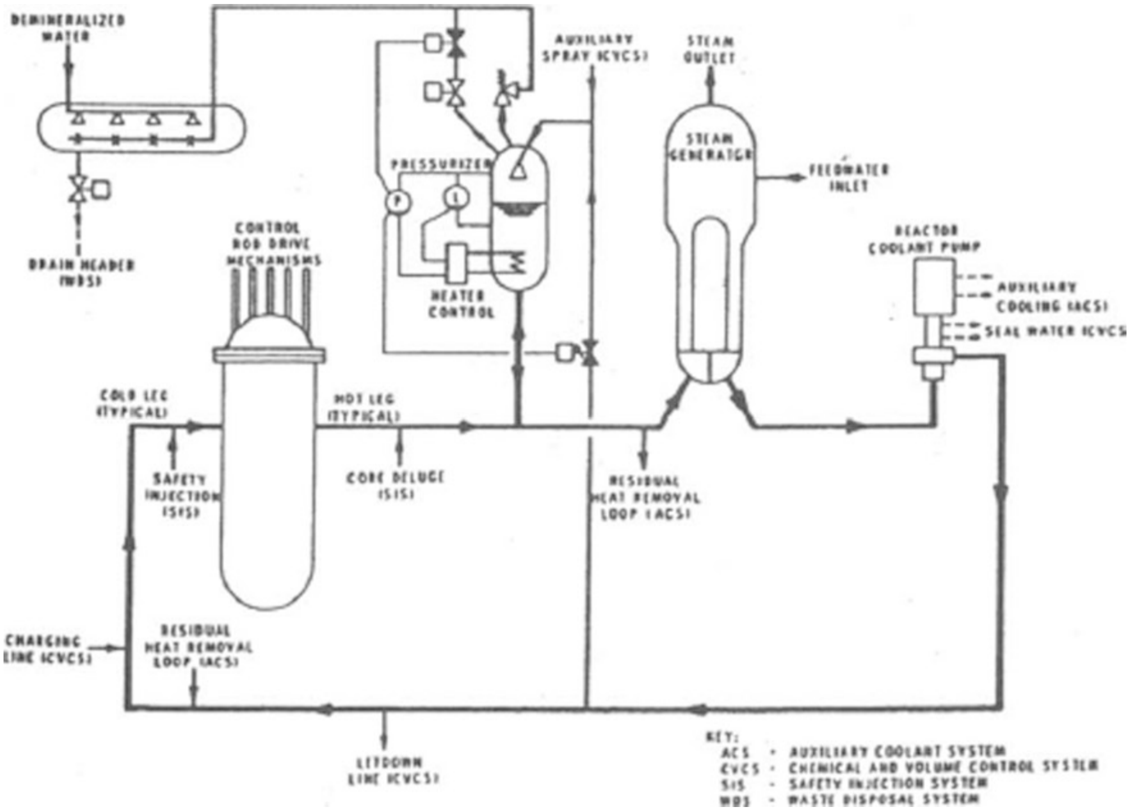


Fig. 21.23 Pressurized water reactor (PWR) for power generation (Courtesy of USDOE)

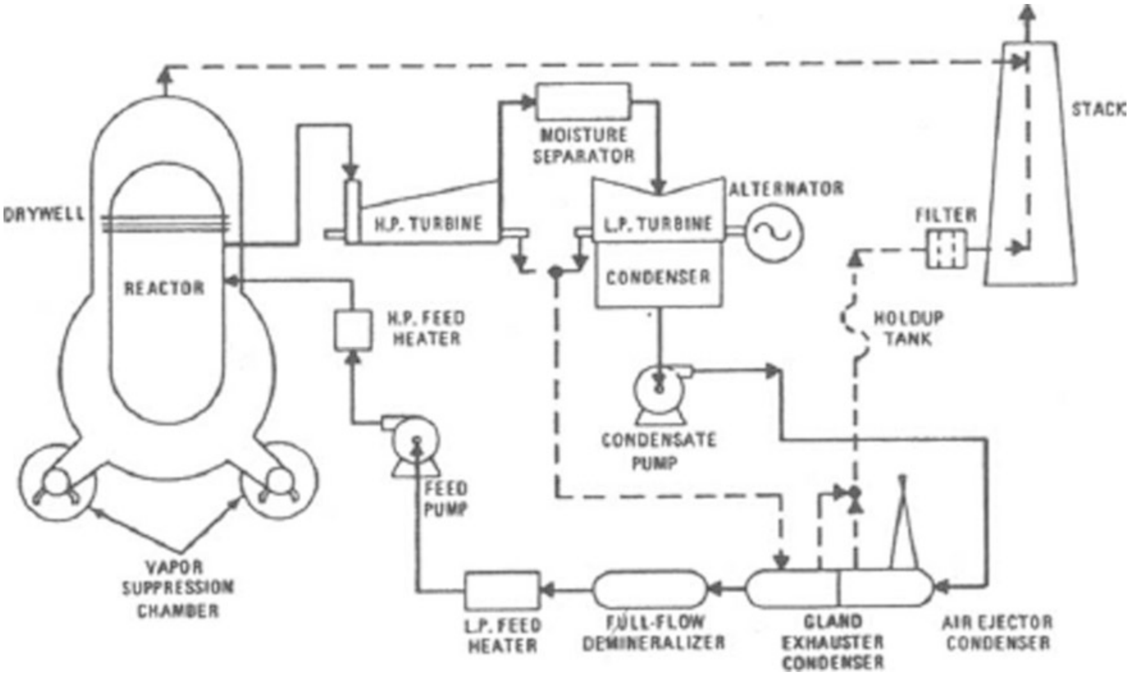


Fig. 21.24 Boiling water reactor (BWR) for power generation (Courtesy of USDOE)

nuclear criticality. Fixed-bed ion exchange units are used to maintain the water quality in both the primary and the secondary systems. The chemical and volume control system reduces boron concentration during the power cycle to compensate for fuel burnup. These operations are carried out continuously through bypass systems [49].A more complete view of the current technology is presented in Table 21.16.

Boiling water reactor. In the BWR, steam is generated in the reactor and goes directly to the turbine (Fig. 21.24). Although the BWR eliminates the intermediate heat exchanger, it places the steam turbine in a radioactive environment. In the primary system, there are full-flow ion exchange units treating about 30,000 gal of water/min, which serves two objectives: removal of radioactivity resulting from neutron-activated corrosion products and fuel cladding leaks; and corrosion protection of the reactor, fuel, and turbine from possible leakage of service water to the primary coolant from the turbine condenser.

In the following paragraphs, three important aspects of LWRs are discussed: fuel elements, the primary coolant system, and reactor containment. They each play an important role in providing multiple barriers that prevent radioactive fuel and fission products from being released to the environment.

Fuel elements. The fuel element is designed to provide the primary containment for the radioactive fuel and fission products over the three- to six-year operating life of the PWR and BWR fuel. In addition, without significant problems, it has provided the containment of the spent fuel after discharge from the reactors for more than 50 years under water in the storage pools at some of the reactors. At some of the older reactors it is now being packaged in stainless steel canisters and placed in dry storage vaults at the reactor site [50].

The quality of the reactor coolant depends on the integrity of about 80–110 miles of 0.0225 inch thich wall Zirconium alloy tubing operating at temperatures up to 350°C containing uranium oxide pellets. The PWR fuel assembly, which is about 14 ft long, is made up of up to about 264 fuel pins (Fig. 21.25). A typical PWR contains about 110 mt of fuel charged with up to 4.95% U-235. The economics of the fuel cycle require this degree of enrichment to achieve a reasonably long operating life and high energy yield from each assembly. Because these nuclear power stations represent a significant part of the total power system capacity of the United States (about 20% in 2012), frequent shutdowns are undesirable. With the increased experience and expectation of a once-through fuel cycle, the enrichment has been increased to raise the energy yield to an average of 50,000 MW days/metric ton and to extend the time between refueling shutdowns to 18 months or longer. This reduces the amount of replacement power needed during shutdowns, which generally comes from the older, more costly, fossil-fuel-powered plants or must be purchased from outside the system. Approximately one-third of the fuel is replaced in the core of the reactor during each refueling shutdown.

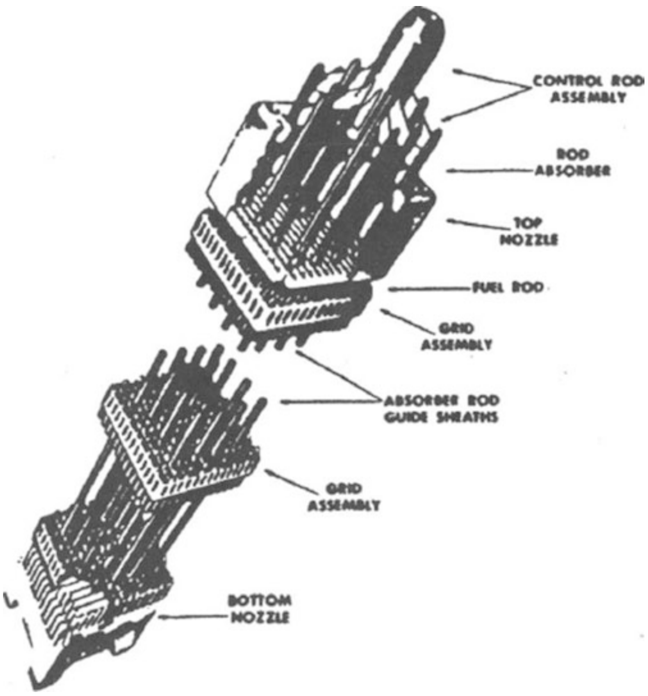


Fig. 21.25 Cutaway of PWR fuel element with the control rod cluster assembly. Element contains about 1,200 lb UO_2 in up to about 264 rods (Courtesy of USDOE)

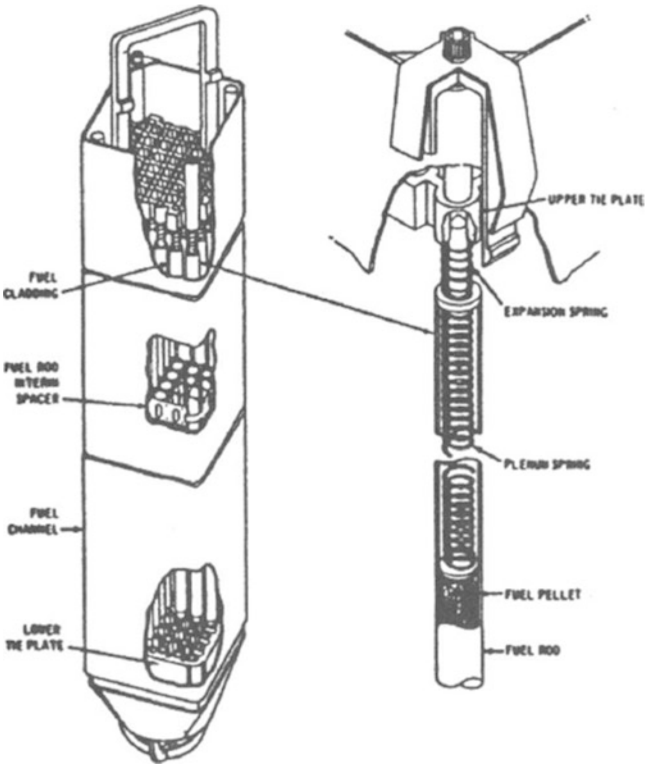


Fig. 21.26 Cutaway of BWR fuel element containing 488 lb UO_2 in 49 rods (Courtesy of USDOE)

In order to improve fuel cycle economics and protect the reactor vessel from radiation-induced embrittlement, partially burned fuel is often loaded in the core periphery. Appropriate placement of the new fuel and shuffling of other partially burned fuel ensures satisfactorily flat power distributions.

A 1,000 MWe BWR core contains about 165 t of UO_2, which is charged up to 4.95% U-235 (Fig. 21.26). Here the integrity of the 0.032 inch thick wall fuel cladding (Fig. 21.26) is even more important than in the PWR because the steam is generated in the reactor core and goes directly to the turbine. This has proved to be a feasible design and eliminates the PWR expense of replacing steam generators.

Two engineering system demonstrations were performed to reduce the uranium-from-ore requirements of LWRs: recycle of the plutonium and conversion to the thorium–uranium cycle to achieve thermal breeding. The demonstration phase of the plutonium recycle development was carried out in seven power reactors. Several LWRs originally were started up on the thorium–uranium cycle, and a light water breeder reactor core using the thorium cycle was tested in the Shippingport reactor. Although these tests did not resolve all the technology questions, no major problems were uncovered. In the United States, plans for fuel recycling and the use of the thorium cycle have been postponed because of the low cost of uranium from ore as well as the perceived danger of diversion of the fissionable materials.

Primary coolant system. Nuclear technology faces challenges on most of the frontiers of engineering, with design and fabrication of the primary coolant system being one of the greatest. The reactor vessel and associated piping and heat exchangers comprise the second barrier to prevent the release of radioactive materials from the fuel to the environment. The steel for the reactor vessels must meet metallurgical standards developed for them to withstand the temperature and pressure cycles as well as neutron irradiation, which in time reduces the ductility of the metal. Carbides and phosphides precipitated in the grain boundaries prevent slippage along the crystal planes. The design criteria for optimum performance require operating pressures of 1,050 psia for BWRs and 2,250 psia for PWRs. Typical dimensions for vessels serving 1,000 MWe stations range from 21 ft in diameter by 70 ft in height for BWRs, to 14 by 42 ft for PWRs, with vessel weights of 782 and 459 t, respectively. These vessels are shop-fabricated, and their transport to the reactor site tests the ingenuity of the carriers [51, 52].

As anticipated, there have been occasional equipment failures involving reactors, but the safety systems have been sufficiently redundant so that one or more have always worked. Even in the Three Mile Island accident in 1979, the safety systems worked as designed. Much of the damage resulted from operator actions to override the safety systems.

As concluded in the Reactor Safety Report, the limitations of the operator created and then seriously aggravated the Three Mile Island incident. Nevertheless, the features of the containment system prevented significant exposure to the plant personnel or any off-site individual, this despite failure of the barriers provided by the fuel cladding and the primary coolant system.

LWR tests-to-failure had been performed to evaluate accident scenarios involving loss-of-coolant accident (LOCA) events such as occurred in the Three Mile Island incident. The power burst tests in a 20 MW PWR have created fuel failures and defined the initiating conditions. The LOCA tests with a 50 MW PWR have demonstrated recovery from catastrophic major feedwater and steam line breaks without fuel damage.

Advanced reactors such as the Westinghouse AP600, AP1000, and General Electric ESBWR have been developed that produce the same amount of power but with many fewer components. These reactors will have increased reliability and greater safety due to the reduced component count that reduces both the cost and the number of points at which failures can occur. In addition, these newer reactor designs are designed to be passively safe. That is, active involvement by the reactor operators is not required to maintain the reactor in a safe state. These types of designs were developed to allow time for reasoned decisions to be made in case of an accident or major equipment failure while still keeping the reactor in a safe state.

CANDU Heavy Water Reactor

This Canadian reactor design originally utilized (1) natural uranium as the fuel, thus avoiding the necessity for uranium enrichment, and (2) heavy water as both coolant and moderator, to maximize neutron utilization. Current designs plan to use slightly enriched fuel and light water as a coolant. In addition to electrical power, this reactor can efficiently produce cobalt-60, which has commercial markets for product irradiation. Each reactor can produce 30–50 MCi annually, and in 1999, the catalog price was $1.20/Ci. The CANDU requires 18% less uranium from ore than the LWR once-through cycles need. The spent fuel contains 0.23% U-235, roughly equivalent to the tails from the gaseous diffusion enrichment operation. It also contains 0.27% fissionable plutonium. The average residence time of a fuel is about 1 year, with approximately 0.3% replaced each day. This requires refueling during operation. The CANDU fuel bundles contain 37 Zircaloy-clad UO slugs, a less complex design than that of LWR assemblies. Many alternative fuel cycles are being considered for the CANDU to further decrease the requirements of uranium from ore, but they will require enrichment or reprocessing [53].

Liquid Metal Fast Breeder Reactor

The salient features of the liquid metal fast breeder reactor include a fuel doubling time of 10–30 years, along with a high coolant temperature for more efficient energy conversion and low operating pressure through the use of liquid sodium or lead as the coolant. Operation of the reactor with fast neutrons increases the neutron multiplication factor for the mixed uranium–plutonium oxide fuel. The power density in the LMFBR is about five times greater than that in the LWR cores; so the reactor vessel is much smaller. This core is surrounded with radial and axial blankets of fertile material. These blanket regions, which capture the leakage neutrons, must be used to achieve breeding, which consists of the production of fissile material from fertile material. Finally, the use of the uranium–plutonium fuel cycle makes optimum use of the plutonium produced in the LWRs.

All major nuclear countries have participated in LMFBR demonstration plants, ranging from about 200–500 MWe. In the United States, the Experimental Breeder Reactor I (EBR I), the starting point of this program, went critical in 1951. LMFBRs were operated in the United States, England, France, Japan, and Russia. In France, the Phenix, a 590 MW/230 MWe reactor, was operated starting in 1973, and the Super Phenix, a 1,200 MWe reactor, first generated power in 1987 and was in commercial operation. The EBR I, Phenix and Super Phenix have all been shut down. In the former Soviet Union, the BN-350 MWe reactor produced power and steam heat for Shevchenko on the shore of the Caspian Sea from 1973 to 1999; and another reactor, the BN-600, has been operating since 1980. In the United States, the 400 MW Fast Flux Test Facility started operation in 1980 to test fuels and materials. The objectives of this program were satisfied in 1989, and it was shut down.

During the 1980s, the Argonne National Laboratory was developing a metal-fueled LMFBR in contrast to the oxide-fueled version described above. The principal advantage of the metal fuel was to be the ease of recycle using electrometallurgical technology. The program was called the Integral Fast Reactor with emphasis on high burnup instead of breeding, possibly 185,000 MW-days/metric ton vs. 50,000 for the present-day PWRs (Fig. 21.27) Early work with metal fuel showed that it expanded because of the presence of the xenon and ruptured the cladding. The fuel rods have been redesigned to accommodate this phenomenon. Finally, each reactor site would have its own reprocessing unit, and this would greatly reduce concern about proliferation. Cost estimates indicate that the cost of the electricity generated would exceed to that from LWR and HWR systems. Much work remains to be done to confirm all these attractive indications [54]. However, the high cost of reprocessing, the fast reactor and fuel manufacture as well as the safely and maintenace issues associated with

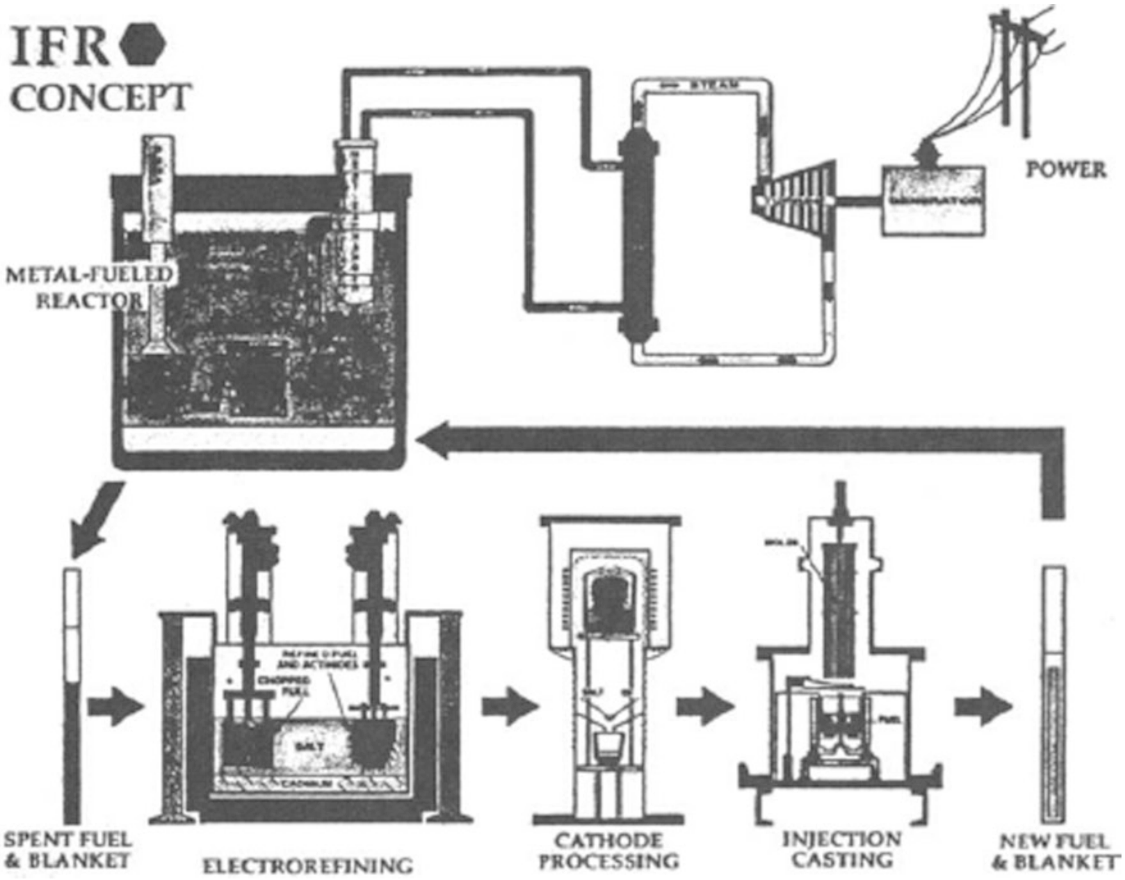


Fig. 21.27 Schematic of integral fast reactor and fuel cycle concept (IFR) (Courtesy of USDOE)

molten sodium metal make this technology commercially unacceptable for electrical production.

Other Nuclear Reactors

Many other nuclear reactors have been developed for research, engineering development, nuclide production, and mobile power. Most noteworthy are the nuclear propulsion systems for naval applications. In the United States, the PWR technology for electric power generation is an outgrowth of the nuclear naval development program. The first nuclear submarine, the *Nautilus*, was commissioned in 1954. Today there are more than 100 nuclear submarines, along with a number of aircraft carriers and other surface vessels. Other nuclear propulsion systems have been studied for arctic tractor trains, aircraft, and rockets for space propulsion, but these studies were terminated before completion owing to insufficient benefit or other funding priorities and environmental concerns. Recently, nuclear reactor development was started again to provide power on orbiting space stations in connection with the United States Strategic Defense Initiative program and for propulsion as part of the NASA manned and unmaned planetary explorer programs [55]. As part of this activity, a "space" reactor was purchased from the Russians who have been employing them for many years.

The remaining classes of nuclear reactors range from zero-power, subcritical neutron sources for university training to large-scale reactor systems for plutonium-239 production. Portable reactors have provided heat, power, and water to US bases in Alaska, Antarctica, and Panama. Private industry has operated various test reactors for reactor studies and radioisotope production.

Radiation Processing

Next to the generation of electric power, radiation processing is potentially the most important commercial application of nuclear energy. Radiation processes have been developed for treating food and medical supplies to inhibit growth of bacteria, viruses, fungi, and insects, and for polymerization of plastics and rubber. They take the place of thermal and chemical treatment to achieve these objectives.

Radiation also is an important tool in the medical treatment of cancer. However, what may be one of its most important applications, extending the storage life of foods, has been seriously delayed by public concern, in spite of extensive tests showing that irradiated foods meet all international safety standards. Currently, the United States irradiates spices for the main purpose of bug removal.

In early radiation processing work, cobalt-60 was the principal source of radiation, but now it has been displaced by machine-produced electrons wherever possible. Cobalt-60 is commercially available from the Canadian CANDU power reactors.

Commercial information regarding the application of machine-produced electrons is closely held, but the following are a few examples of their use:

1. Low-energy electrons, up to 0.5 MeV, are in general use to polymerize thin plastic sheeting and coatings. Fixed beams up to 2.5 m wide irradiate 0.4-mm-thick products with 10 kiloGray (kGy) at 1,600 m/h.
2. Electrons with energies up to 5 MeV are widely used in scanned beam machines. The typical scan width is 1–2 m when applied to the polymerization of thicker materials. These machines also are being used for irradiating grain and chicken feed. The Russians are treating 200 t/h of grain with a dose of 0.4 kGy, and the Israelis treat chicken feed with 3.8 kGy at 15 t/h. In both cases the flowing-bed depth is about 6 mm.
3. A 5-kW beam of 7 MeV electrons is delivering a 3 kGy dose to flat cakes of deep frozen cakes of chicken meat at a rate of 50 t/day; disposable medical supplies are being treated at the rate of 100 m^3/day [56].

Radioisotope Applications

The tremendous release of energy from nuclear reactions makes possible a unique family of applications for long-lived radioisotopes that are important to health, science, and industry. Whereas fission and fusion occur almost instantaneously, other radioactive decay processes occur in times ranging from a few minutes to thousands of years. The general areas of application may be grouped into irradiation, thermal energy generation, and tracer applications [57].

Radiation Sources

Radiation from radioactive nuclides is used to detect changes in density or other characteristics of materials, to promote chemical, physical, or biological changes, and to provide a source of thermal energy. Some radioactive materials find use in industry, research, and medicine as tracers for physical, chemical, and biological processes [58].

The very important radioisotope cobalt-60 has a 5-year half-life and emits 1.17 and 1.33 MeV gamma rays. Co-60 is made by irradiating naturally occurring Co-59 with neutrons in a reactor. It is the major radioisotope in a number of applications including the following.

1. Destruction of malignant tumors.
2. Sterilization of prepackaged medical supplies, particularly those that are heat-sensitive or can be contaminated by chemical agents.
3. Food irradiation to extend shelf life.
4. Radiographing metals to determine weld quality (Iridium-192, with a 0.3 MeV gamma ray, provides better definition than Co-60 in some cases but has a shorter half-life.).
5. Polymerization of coatings on electrical conductors and paper and monomers impregnated into wood and concrete (Machine radiation generally is preferred in applications involving thin materials.).
6. Density measurements for level controllers in silos and other vessels bearing solids.

There also are many applications for alpha and beta radiation sources, including:

1. Polonium-210 alpha activation of beryllium to yield neutrons that in turn are used to start up nuclear reactors.
2. Plutonium-238 activation of beryllium to produce neutrons for logging drill holes to detect hydrogenous materials in geologic formations.
3. Strontium-90 beta sources to measure the thickness of paper in paper manufacture and also to discharge static electricity by ionizing the air in high-speed printing operations.

Radioisotope Thermoelectric Generators

These devices (radioisotope thermoelectric generators, RTGs) provide reliable long-life sources of electrical energy. The thermal energy required to drive the thermoelectric element is provided by a long-lived radionuclide. The electricity is generated by passing the thermal energy through the thermoelectric element from the heat source at one end (the hot foot) to a heat sink at the other end (the cold foot). Germanium silicide is now the thermoelectric material of choice; because of its higher thermal stability, it can be driven with larger temperature differentials. Lead telluride is more efficient per degree of temperature difference but has lower thermal stability [59].

Plutonium-238 is the most important heat source; it is an alpha emitter whose radioactive decay energy is easily absorbed and converted into thermal energy. It also is easily shielded. In addition, its long half-life, 90 years, makes possible hundred-watt power sources with a design life exceeding 10 years. Although solar cells generally comprise the

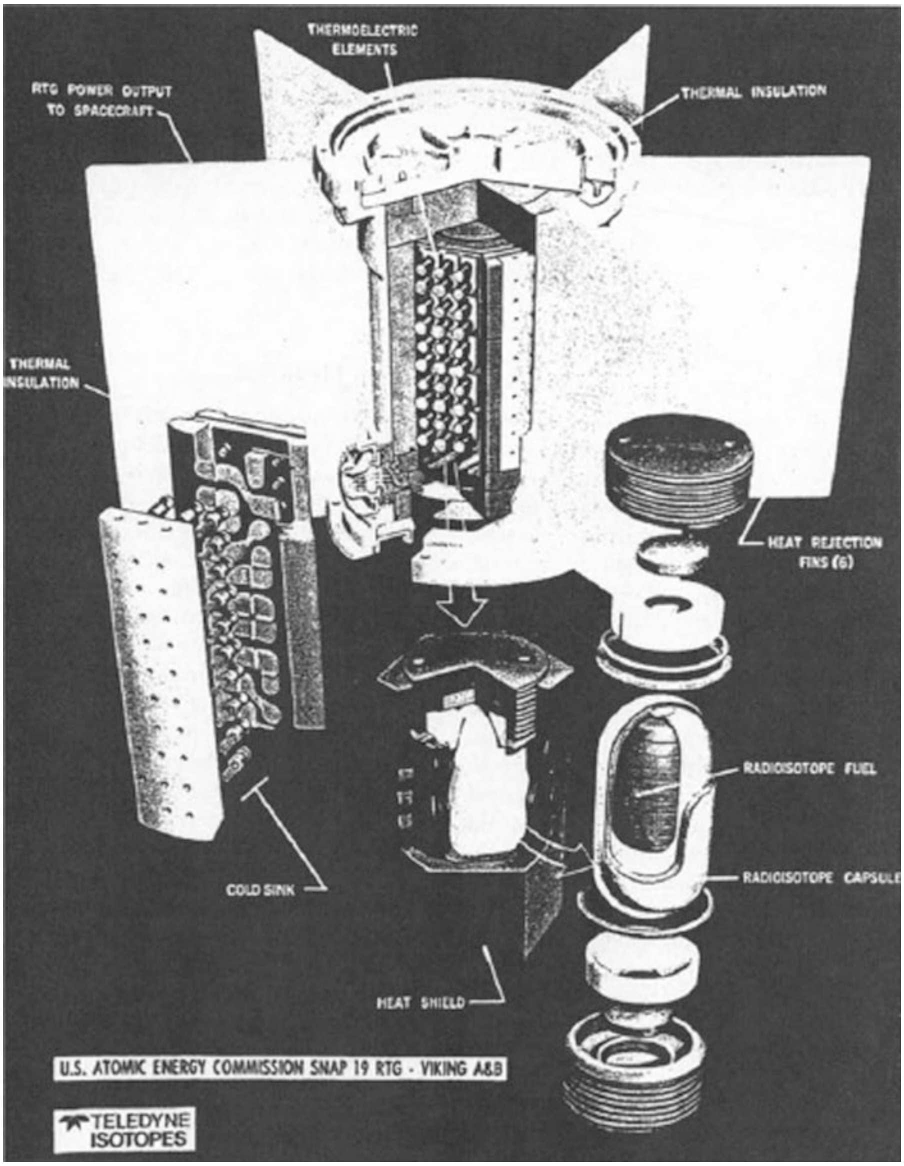


Fig. 21.28 The multi-hundred-watt electric generator. Three units on Voyager generated 475 W from the decay heat of plutonium-238 acting on thermoelectric couples (Courtesy of USDOE)

technology of choice to power extraterrestrial missions, a Pu-238 RTG was used for the Apollo moon landings. Such devices are used when the mission path is shielded from the sun, and they continue to be used for deep-space missions (see Fig. 21.28).

Terrestrial RTGs have also been developed and are commercially available. A beta emitter, 30-year strontium-90, has been the major radionuclide used for terrestrial applications. Although beta particles (negative electrons) are easily sorbed and converted to thermal energy, the sorption process is accompanied by the emission of X-rays that require shielding. This necessitates a heavy device, about 300 lb for a 5-W source. Because strontium-90 is a fission product, the supply now is limited to 100 MCi recovered in the 1960s from the Hanford defense waste. This is equivalent to about 10 kW of electrical energy, assuming a 5% conversion efficiency.

Tracers

The radiation emitted from decay of single atoms can be easily measured. Thus, small amounts of radioisotopes can be incorporated into systems to study, for example, chemical reactions and the flow of fluids through complex systems such as organs of the human body. They also can be used in more mundane applications such as locating leaks in buried piping. Many organic compounds tagged with carbon-14 in specific positions are commercially available.

Because C-14 with a half-life of 5,700 years occurs in nature, it can be used to determine the age of carbon-containing artifacts. The basis for such determinations is the loss of C-14 that was trapped in the artifact (animal, human, fired pottery, etc.). The related art and science of carbon dating are fairly complex and thus subject to significant uncertainties.

Table 21.18 Radionuclides used in nuclear medicine

For organ function and tumor visualization	
Technetium-99 m	Carbon-11
Iodine-131	Oxygen-15
Gallium-67	Nitrogen-13
Strontium-85	Fluorine-18
Indium-111	Iodine-123
Xenon-133	
For irradiation of tumors	
Cobalt-60	Cesium-137
Radium-226	Iodine-131

Nuclear Medicine

Nuclear medicine is one part of nuclear utilization readily accepted by the public. However, there is strong opposition to the siting of the low-level waste repositories needed for disposal of the wastes that result from the practice of nuclear medicine.

The medical applications of nuclear technology range from in vitro and in vivo injections for diagnostic tests to cobalt radiation for cancer therapy. A new medical specialty was created, a family of compact cyclotrons was developed to provide short-lived nuclides, and a sizable industry evolved to produce technetium. Until the nuclear industry was created, technetium had been missing from the chart of chemical elements because the half-life of the most stable member was too short, 21,000 years. Technetium and several other nuclides of importance here are discussed elsewhere in the chapter in connection with their production (see Table 21.18) [60, 61].

The two principal applications of nuclear medicine are for radiation treatment of tumors and for visualizing organ functions and tumors. These nuclear-medical procedures supplement other medical treatment and diagnostic procedures, and are available in most hospitals throughout the industrialized countries and in the major hospitals in other countries.

References

Status and Outlook

1. Finger H (1989) Need for nuclear energy, national energy strategy hearings. US Council for Energy Awareness, Washington, DC
2. World Nuclear Association Web site. Information and brief news. Accessed September 2012. http://www.world-nuclear.org/info/reactors.html
3. NRDC: Nuclear Notebook Global nuclear stockpiles, 1945–2002 By Robert S. Norris and Hans M. Kristensen November/December 2002 (vol. 58, no. 06) © 2002 Bulletin of the Atomic Scientists, pp. 103–104
4. Chernobyl Nuclear Power Plant Accident-Health and Environmental Consequences. DOE/ER-0332. US Department of Energy, June 1987

Nuclear Safety

5. US NRC (1979) The accident at Three Mile Island. Staff Report to the President’s Commission, Nuclear Regulatory Commission, Washington DC

Nuclear Safety

6. US NRC. Code of federal regulations, Title 10 Energy, Chapter 1. Nuclear Regulatory Commission, Washington, DC
7. Domenici PV (2004) A brighter tomorrow: fulfilling the promise of nuclear energy. Rowman & Littlefield, New York
8. Fact Sheet: United States NRC. Nuclear reactor risk. June 2003
9. USNRC (1975) Reactor safety study, WASH 1400 (NUREG 75/014). US Regulatory Commission, Washington, DC
10. From HP Society University of Michigan Web site, 1/18/05

The Earth’s Supply And Demand

11. USEIA (1989) Annual report to congress 1988. Energy Information Administration, Washington, DC
12. USEIA (1988) Annual energy review, DOE/EIA-0384[88]. Energy Information Administration, Washington, DC
13. Energy efficiency, The Energy Daily, 18; 64, Washington, DC, 1990
14. Parent JD (1977) A survey of United States and total world production, reserves and remaining recoverable fossil fuel and uranium. Institute or Gas Technology, Chicago
15. OECD/NEA-IAEA (1990) Uranium resources production and demand. OECD Nuclear Energy Agency and International Atomic Energy Agency, Paris, 1990
16. Energy Information Administration/Annual Energy Review, 2003

Nuclear Processes

17. Friedlander G, Kennedy J, Macias E, Miller J (1981) Nuclear and radiochemistry, 3rd edn. Wiley, New York
18. Nuclides and isotopes, 14th Ed, 202-37-000 GE Nuclear Operations, San Jose, CA

Fission

19. Lamarsh JR (1972) Introduction to nuclear reactor theory. Addison-Wesley, Reading
20. Advanced proliferation resistant, lower cost, uranium-thorium dioxide fuels for light water reactors. Nuclear Energy Research Initiative, Idaho National Engineering and Environmental Laboratory, Idaho Falls, ID, 2000

Fusion

21. A status report on controlled thermonuclear fusion, STU/PUB/872 International Fusion Research Council. International Atomic Energy Agency, Vienna, 1990
22. Starpower, the U.S. and the International Quest for Fusion Energy, OTA-E-338, Office of Technology Assessment, Congress of the United States, Oct 1987
23. Cleary, D. The American Association for the Advancement of Science Website. *Science Insider*, May 27, 2010, accessed August 2010. http://news.sciencemag.org/scienceinsider/2010/05/no-solution-yet-to-iters-budget-html

Nuclide Production

24. Kauffman G (1990) Beyond uranium. Chemical and Engineering News, Washington, DC, Nov. 1990
25. Meyer W, Plascjak P. Cyclotron group, National Institute of Health-Personal communication
26. Cohen BL (1971) Concepts of nuclear physics. McGraw-Hill, New York

Neutron Transmutation Products

27. Crandall J (1965) The Savannah River high flux demonstration, USAEC Report DP999. US Atomic Energy Commission, Washington, DC 28-June 1965

Charged Particle Transmutation Products

28. Lagunas-Solar M (1982) Cyclotron production of no-carrier-added medical radionuclides. In: 7th Conference on the applications of acceleration in research and industry, Denton

Isotope Enrichment

29. Isotec (1990) Stable isotopes for research and industry. Isotec, Maimisburg
30. Avona V, Spicer H (1987) Separation and applications of stable isotopes. American Laboratory, April 1987
31. Rae H (ed) (1978) Separation of hydrogen isotopes. American Chemical Society, Washington, DC

The Uranium Fuel Cycle

32. Nuclear Engineering International (1990) Fuel Cycle Review, Cornell University
33. Eister W, Stoughton R, Sullivan W (1955) Processing of nuclear reactor fuel. In: Glasstone S (ed) Principles of nuclear reactor engineering. Van Nostrand, New York

Fuel Preparation

34. Mantz E (1970) Production of uranium tetrafluoride and uranium metal, USAEC Report NCLO 1068. U.S. Atomic Energy Commission, Washington, DC
35. Olander D (1978) The gas centrifuge. Sci Am 239:2
36. United States Enrichment Corporation (1988) The case for restructuring. The Council for Energy Awareness, Washington, DC

Spent Fuel Reprocessing

37. Long J (1967) Engineering for nuclear fuel reprocessing. Gordon and Breach, New York
38. Logsdail D et al (1985) Solvent extraction and ion exchange. Wiley, New York

Radioactive Waste Management

39. USDOE (1989) Integrated data base for 1989: spent fuel and radioactive waste inventories. U.S. Department of Energy, Washington, DC

Storage of Spent Fuel

40. Matthew Bunn et al and Atsuyuki Suzuki et al (2001) Interim Storage of Spent Nuclear Fuel, Harvard University and University of Tokyo
41. Eister W (1977) Materials considerations in radioactive waste storage. Nucl Technol 1:6

Low-Level Waste Disposal

42. USDOE (1989) Annual report to congress. Office of Civilian Radioactive Waste Management, U.S. Department of Energy, Washington, DC
43. USDOE (1989) Environmental assessment of remedial action at the Monument Valley Uranium Mill Tailings Site, Monument Valley, AZ. U.S. Department of Energy, Washington, DC
44. USDOE (1989) Uranium mill tailings remedial action program, Annual Report, 1989. U.S. Department of Energy, Washington, DC

Transportation of Nuclear Materials

45. Rail transportation corridor analysis, BMI/ONWI-617. U.S. Department of Energy, Washington, DC, p. 186
46. USDOE (1986) Transporting spent nuclear fuel. U.S. Department of Energy, Office of Civilian Radioactive Waste Management, Washington, DC

The Nuclear Reactor

47. Benedict M, Pigford T, Levi H (1981) Nuclear chemical engineering. McGraw-Hill, New York
48. Forsberg C, Reich W (1991) Worldwide advanced nuclear power reactors with passive and inherent safety: what why, how and who, ORNL/TM-11907, September 1991. Oak Ridge, Oak Ridge National Laboratory

Light Water Reactors

49. Beckjord E et al (1989) International comparison of LWR performance, Report MIT EL 87–004. Massachusetts Institute of Technology, Cambridge, Feb 1989
50. Frost B (1982) Nuclear fuel elements. Pergamon, New York
51. Cohen P (1969) Water coolant technology of power reactors. Gordon and Breach, New York
52. USCEA (1988) Advanced LWRs, Nuclear industries, July 1988. U.S. Council for Energy Awareness, Washington, DC

CANDU Heavy Water Reactor

53. Veeder J, Didsbury R (1985) A catalogue of advanced fuel cycles in CANDU-PHW reactors. Chalk River Nuclear Laboratories, Chalk River

Liquid Metal Fast Reactor

54. Till C, Chang Y (1988) The integral fast reactor. Adv Nucl Sci Technol 20:127–154

Other Nuclear Reactors

55. See NASA Web site concerning Explorer program and nuclear reactors.

Radiation Processing

56. USDOE (1988) Machine sources for food irradiation. U.S. Department of Energy, Department of Energy, Washington, DC

Radioisotope Application

57. Eister W et a1 (1970) Radioisotope production in the U.S., radioisotope production study, Sao Paulo, Brazil. International Atomic Energy Agency, Vienna

Radiation Sources

58. Eichholz G (1972) Radioisotope engineering. Marcel Dekker, New York

Radioisotope Thermal Electric Generators

59. Arnold ED (1964) Handbook of isotopic power source characteristics. Oak Ridge National Laboratory, Oak Ridge

Nuclear Medicine

60. Eister W et al (1970) Radiopharmaceuticals and short-lived radioisotopes, radioisotope production study, Sao Paulo, Brazil. International Atomic Energy Agency, Vienna
61. Ester, W., et al., Radioisotope generators America (see above)

Synthetic Nitrogen Products

22

Gary R. Maxwell

Nitrogen products are among the most important chemicals produced in the world today. The largest quantities are used as fertilizers, but nitrogen products also find very important uses in the manufacture of nylon and acrylic fibers, methacrylate and other plastics, foamed insulation and plastics, metal plating, gold mining, animal feed supplements, and herbicides.

Nitrogen

Characteristics

Nitrogen is colorless, odorless, and slightly lighter than air with a density of 0.967 (air = 1.0). Some people consider nitrogen to be one of the most dangerous gases. This is because a person in an atmosphere of nitrogen can lose consciousness without any warning symptoms in as little as 20 s. Death can follow in 3–4 min. One deep breath of 100% nitrogen can be fatal because nitrogen will displace carbon dioxide in the body completely. In the absence of a carbon dioxide signal to the brain, breathing stops.

No one should work in or enter atmospheres containing less than 19.5% oxygen, unless equipped with a self-contained breathing apparatus or a breathing air mask. This is also true of rescue personnel who can be overcome by the same oxygen-deficient atmosphere as the initial victim [1].

The properties of nitrogen are listed in Table 22.1.

Although the ability of nitrogen to easily unite with other elements is quite limited, it does form some interesting compounds. When nitrogen combines with certain elements, it produces dyes that rival the rainbow in the brilliance of their colors. With other elements, nitrogen forms drugs such as morphine, quinine, and acetanilid that ease pain and combat diseases. With other compounds, nitrogen forms some of our most powerful explosives such as nitroglycerin and guncotton. Nitrogen can also be used to make some very deadly compounds such as ptomaines. Some nitrogen compounds provide pleasant-tasting flavors and sweet perfumes, and others are so vile in taste and odor that they are beyond description. In nitric acid, we find nitrogen in a strong acid, whereas in ammonium hydroxide, we have it in a well-known base [2].

Nitrogen Production Processes

Economical fixation of nitrogen from the atmosphere has been a never-ending quest. It has been a difficult task because elemental nitrogen is comparatively unreactive. It only combines with most elements under high pressure and/or high temperature. As a result, industry meets its demand for high-purity nitrogen by obtaining most of it from liquid air (see reference 1 and Chap. 27).

Nitrogen production via PSA (pressure swing absorption) is based on the principle that nitrogen and oxygen have different absorption rates on carbon molecular sieves (CMS). Some of the nitrogen production processes that use this technology are described in [3–7].

Nitrogen Fixation

Nature uses nitrogen fixation to combine free nitrogen with other elements. Nodules on the roots of many plants contain bacteria that extract nitrogen from the air and convert it into soluble nitrates that fertilize the soil. Most nitrogen products are based on nitrogen oxides and ammonia, as described on the following pages.

Nitrogen Oxides

The five oxides of nitrogen are [2]: nitrous oxide (N_2O), nitrogen dioxide (NO_2), nitric oxide (NO), nitrogen

---

G.R. Maxwell (✉)
DuPont Company, Memphis, TN, USA

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_22, © Springer Science+Business Media New York 2012

Table 22.1 Physical properties of nitrogen

Property	Value
Molecular weight	14.0067
Boiling point, °C (°F)	−195.8 (−320.5)
Freezing point, °C (°F)	−200.86 (−346.0)
Critical temperature, °C (°F)	−146.95 (−232.4)
Critical pressure, kPa (psi)	3,393 (492.3)
Critical volume, cc/g mol	89.5
Specific heat, 20°C and 147 psia	
c_p (BTU/lb°F)	0.247
c_v (BTU/lb°F)	0.176
21°C (J/kg K)	1,046
$k = c_p/c_v$	1.41
Heat of formation of gas, ΔH, kJ/mol	0.3604
Heat of fusion at melting point	
J/g	25.6
Heat of vaporization at boiling point	
J/g	199
BTU/lb	85.5
Solubility in water (g N_2 per 100 g H_2O at 760 mm)	
0°C	0.00239
20°C	0.00189
40°C	0.00139
60°C	0.00105
Specific gravity	
Relative to air	0.967
−195.8°C (liquid)	0.808
−252°C (solid)	1.026
Specific volume (70°F and 1 atm), ft^3/lb	13.80
Density, kg/m^3	
20°C	1.16
Liquid at boiling point	808.6
Solid at melting/freezing point	1,028
Thermal conductivity, W/m K	
Gas @ 21°C	0.026
Liquid @ −200°C	0.14456

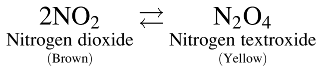
pentoxide (N_2O_5), and nitrogen trioxide (N_2O_3). Nitrogen trioxide and nitrogen pentoxide are the anhydrides of nitrous acid and of nitric acid.

Nitrous oxide (laughing gas) is a colorless gas. Since the 1840s, the major use of nitrous oxide has been as an anesthetic, especially by dentists. It is also used as a propellant in some aerosol cans, in atomic absorption spectrophotometry, in cryosurgery, and in racecar engines to provide extra power and acceleration.

Nitric oxide is a colorless gas that is insoluble in water. The most noticeable chemical characteristic of nitric oxide is the ease with which it combines with oxygen to form nitrogen dioxide. Its major use is in the industrial preparation of nitric acid.

In the late 1980s, nitric oxide (NO) was discovered as a product of enzymatic synthesis in mammals. And in 1998, the Nobel Prize in Physiology was awarded to the scientist that discovered the role of NO as a biological messenger. Nitric oxide uses range from its role as a critical endogenous regulator of blood flow and thrombosis to a principal neurotransmitter mediating erectile function to a major pathophysiological mediator of inflammation and host defense. These major discoveries have stimulated research into a vast array of fields. The Nitric Oxide Society has been formed to coordinate meetings, and they also publish a research journal. More information is available at http://www.apnet.com/no.

Nitrogen dioxide is a red-brown gas that is very soluble in water. It has an unpleasant odor and is quite poisonous. When nitrogen dioxide is cooled, its brown color fades to a pale yellow.



$$\underset{\substack{\text{Nitrogen dioxide}\\ \text{(Brown)}}}{2NO_2} \rightleftarrows \underset{\substack{\text{Nitrogen textroxide}\\ \text{(Yellow)}}}{N_2O_4}$$

The importance of nitrogen dioxide lies in the fact that it combines readily with water to form nitric acid.

$$3NO_2 + H_2O \rightarrow 2HNO_3 + NO$$

This is the final step in the synthetic production of nitric acid. Nitrogen dioxide also acts as an oxidizing agent in the manufacture of sulfuric acid.

Nitrous oxide production technology can be licensed from the Sanghi Organization in Worli, India (http://www.sanghioverseas.com). Their technology is based on heating ammonium nitrate to 250°C and then purifying the resulting gas stream.

Research continues on the formation of nitrogen oxides. The high-temperature fixation of nitrogen as oxides remains an environmental problem rather than a commercially attractive process, as large quantities of nitrogen oxides are produced by fixation in high-temperature combustion processes such as power plants, automobile engines, and home furnaces. The US EPA promulgated regulations in 2000 (NO_x SIP Call) that require several states in the Eastern United States to reduce their NO_x emissions by a combined total of 35–40% by 2005 (see http://www.epa.gov/ttn/otag).

Ammonia

Ammonia was prepared by the alchemists by distilling leather scraps, hoof, horns, and other animal refuse. Because they found that the best ammonia was obtained by distilling deer antlers, the alchemists called ammonia "spirits of hartshorn" [2]. Nature forms ammonia by decomposition of proteins. The odor of ammonia can often be detected around manure piles and other decaying organic matter.

A process for synthesizing ammonia from nitrogen and hydrogen, using high temperatures and pressures and an iron-containing catalyst, was invented by Fritz Haber at

BASF in 1908 [3]. In 1909, C. Bosch of BASF built a pilot plant using an osmium-based catalyst, and in 1913, a larger plant was built in Germany [8–13]. The chemistry for this process is

$$N_2 + 3H_2 \rightleftarrows 2NH_3$$

In 1883, the Mond gas process produced ammonia by gasifying coal at a relatively low temperature by using a mixture of air and steam [3]. It was invented primarily to produce ammonia that was needed for the Solvay process that makes sodium carbonate [14, 15].

Some other processes that have been used to make ammonia are

- The Casale ammonia process in Switzerland in 1921 [3, 11, 16].
- The high-pressure Claude–Casale ammonia synthesis process in the 1920s [3, 11, 17, 18].
- The Fauser process in Italy in 1924 [11].
- The Mont Cenis process that was similar to the Haber–Bosch process except that it used coke-oven gas [3]. It was started up in The Netherlands in 1929 [19, 20].
- The Braun process is a variation on the classic ammonia synthesis process in which the synthesis gas is purified cryogenically [3]. It has been widely used since the mid-1960s [21].
- The AMV process for making ammonia was invented by ICI and announced in 1982 [3]. It is described in [22–24].
- The LCA (*L*eading *C*oncept *A*mmonia) is essentially a simplified form of the standard ammonia synthesis process that is more suitable for smaller plants. It is described in [3, 14, 25, 26].
- The BYAS (*B*ypass *A*mmonia *S*ynthesis) process can be used for economical expansion of existing ammonia synthesis plants. It is described in [3].
- The KAAP (*K*ellogg *A*dvanced *A*mmonia *P*rocess) process is the first high-pressure ammonia synthesis process that makes ammonia from nitrogen and hydrogen without the aid of an iron-containing catalyst [3]. It is described in [27–30].

Ammonia and some of its salts and derivatives are used as fertilizers. Ammonia also is a very important building block for many other chemical commodities as shown in Fig. 22.1.

Nitrogen Consumption

The demand for nitrogen in a chemically fixed form (as opposed to elemental nitrogen gas) drives a huge international industry that encompasses the production of many nitrogen products. Nitrogen products had a total annual commercial value on the order of $50 billion in 1996. The cornerstone of this industry is ammonia. Virtually all ammonia is produced in anhydrous form via the Haber process. Anhydrous ammonia is the basic raw material in the manufacture of fertilizers, livestock feeds, commercial and military explosives, polymer intermediates, and miscellaneous chemicals [31].

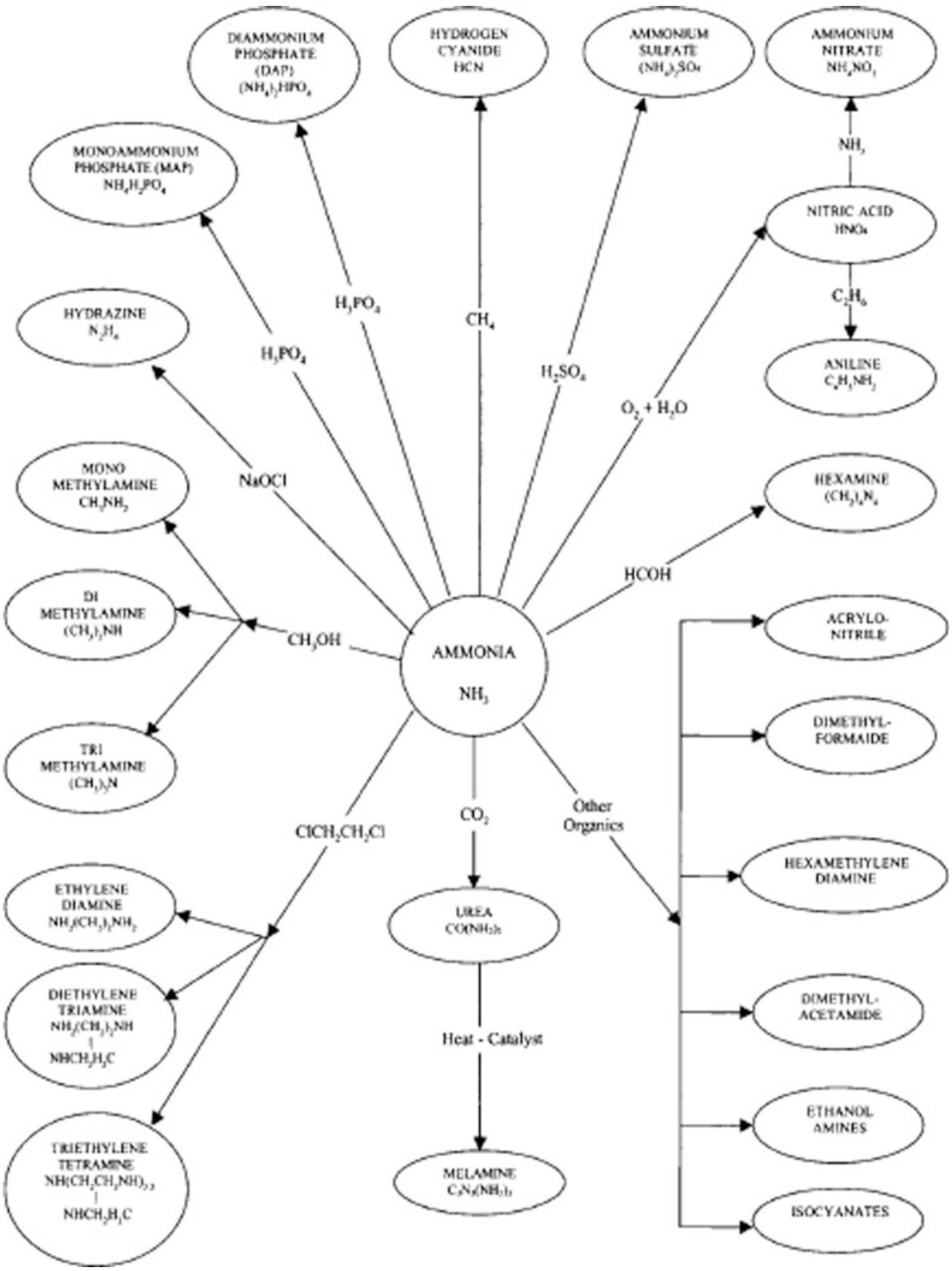
On a worldwide basis, nitrogen fertilizer consumption accounts for more than 85% of the total nitrogen industry. In the United States, with its significant industrial markets, fertilizer use accounts for only approximately 80% of total nitrogen consumption [31]. About 97% of nitrogen fertilizers are derived from synthetically produced ammonia. The rest are produced as by-product ammonium sulfate from caprolactam processes, acrylonitrile (ACRN) processes, metal refining operations, and small quantities of natural nitrates, especially from Chile [31]. Ammonia is the starting material for six primary nitrogen fertilizer products: urea, ammonium nitrate, ammonium phosphates, ammonium sulfate, nitric acid, and nitrogen solutions [32].

Up through the 1960s, the development of the nitrogen industry took place in the developed countries of Western Europe, North America, and Japan. However, in the 1970s and early 1980s, the construction of new plants shifted to the gas-rich countries of the Caribbean and Middle East. Additional plants were also built in some large consuming countries such as China, India, Indonesia, and Pakistan. At the same time, many plant closures occurred in Western Europe and Japan. Western European share fell from 20% in 1980/1981 to 11% in 1997/1998. In 1980/1981, the developing countries accounted for 31% of nitrogen fertilizer production. By 1997/1998, their share had increased to 55% [32] (see Fig. 22.2). The main regions that produced nitrogen fertilizers in 1997/1998 are summarized in Table 22.2. Although at least 60 countries produce urea, over one-third of the world's production is in just two countries: China and India [32].

World ammonia capacity increased by nearly 14% from 1984 to 1996, and capacity for urea, the primary downstream nitrogen product, increased by 45%. The increases were due primarily to (1) a desire by some major importing countries to become more self-sufficient and (2) the construction of export-oriented capacity in the Middle East and in the former Soviet Union (FSU). Ammonium nitrate capacity declined by 2% from 1984 to 1996, and ammonium sulfate capacity declined by 8% [31]. Ammonium phosphate capacity increased by 9% during this time.

In the future, developing nations are expected to continue to account for most of the increases in ammonia and urea capacity. Ammonia capacity is expected to increase by about 13.5 million tons and urea capacity by about 15 million tons between 1999 and 2004. The availability of relatively low-cost feedstock (usually natural gas) will be a major determinant as to where this new capacity is installed [31, 33].

Fig. 22.1 Synthetic nitrogen products



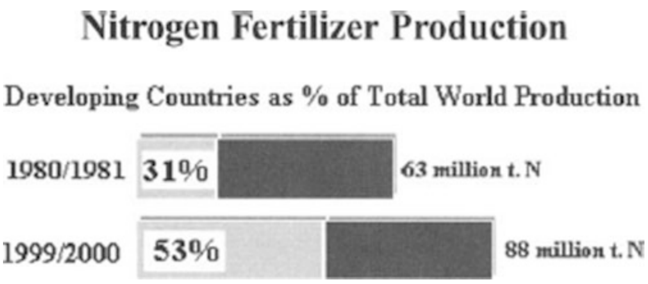


Fig. 22.2 Nitrogen fertilizer production 1980/1981 and 1999/2000 [32] (Reproduced by permission of the International Fertilizer Industry Association, IFA)

The apparent consumption of ammonia increased by about 15% between 1984 and 1996, and the apparent consumption of urea increased by about 54%. The developing countries are largely responsible for the increased consumption.

In 2002, urea accounted for almost 61% of worldwide consumption of the four major downstream nitrogen products. World urea consumption is forecast to increase at an average annual rate of 2.4% between 1999 and 2004. Only ammonium phosphate consumption, at 4.3% per year, is expected to grow more rapidly [31, 32].

Although world fertilizer consumption grew substantially between 1984 and 1996, it actually peaked at 79.6 million

Table 22.2 Main producing regions for nitrogen fertilizers

Region	% of world production
China	23
North America	18
South Asia	15
Western Europe	11
Former Soviet Union (FSU)	10
Other countries	7
Central Europe	5
Middle East	5
Indonesia and Japan	4
Mexico and the Caribbean	2

tons of nitrogen in 1989 and declined through 1994 because of a sharp drop in fertilizer consumption in Eastern Europe, the FSU, and Western Europe. Total world nitrogen fertilizer consumption resumed its long-term upward growth trend in 1995, and substantial growth is expected [31]. Japan is one of the few countries where the use of slow-release fertilizers is developing on field crops, especially in crops grown in paddies [32].

Estimates of world industrial nitrogen consumption by region show that the four largest markets in 1984 were North America, socialist Asia, Western Europe, and the FSU [31]. By 1996, South Asia had replaced the FSU as one of the four largest markets.

Trade is an important component of the world nitrogen industry, and trade in each of the nitrogen products analyzed (ammonia, urea, ammonium nitrate, ammonium sulfate, and ammonium phosphates) increased substantially between 1986 and 1998. The percentage of production that is traded internationally varies from 10% for ammonia up to 40% for ammonium phosphates [31].

Environmental Issues

Since 1980, the nitrogen industry has made major improvements in environmental controls to reduce atmospheric emissions and liquid effluents.

The primary pollution problem in nitric acid manufacture is the abatement of nitrogen oxides (NO_x) in tail gases. In the United States, gaseous emissions from newly constructed nitric acid plants must be limited to 1.5 kg of NO_x/ton of nitric acid (100% basis) produced, with a maximum stack opacity of 10%. Modern acid towers, with extended sections, can reduce NO_x emissions to less than 200 ppm [31].

The most important environmental concern about the use of nitrogen fertilizers is the loss of nitrogen to groundwater. This is having a significant impact on the use of nitrogen fertilizers, particularly in Western Europe. There is a lesser but still significant concern about nitrate levels in groundwater in the United States. Other than this runoff aspect, which most seriously affects ammonium nitrate, no serious environmental concerns exist with the use of the major nitrogen products as fertilizer materials [31].

Environmental issues play an important role in the ammonium sulfate industry. This is because a significant portion of the world's ammonium sulfate production is the direct result of the necessity to remove sulfur dioxide from stack gases at various metal smelting and refining operations in order to conform to government regulations on sulfur dioxide emissions. A large potential source of additional by-product ammonium sulfate production is sulfur dioxide recovery from coal-fired electrical generating stations. It is estimated that a 500-MW plant that is burning 3.5% sulfur coal could generate 250,000 tons of ammonium sulfate per year [31]. This is equivalent to ~7% of the total worldwide ammonium sulfate production in 2000.

Ammonia Production

Because no economical nitrogen fixation process that starts with nitrogen oxides has been discovered, ammonia has developed into the most important building block for synthetic nitrogen products. Anhydrous ammonia is produced in about 80 countries [32].

In the USA, only 80% of the ammonia is used to make fertilizers. Chemical intermediates (such as acrylonitrile and caprolactam) account for 19% of the ammonia use, and the remaining 1% is used in pulp and paper, metals, and refrigeration applications [34].

Worldwide ammonia capacity grew from 119 million tons in 1980 to a peak of 141 million tons in 1989 [32]. Ammonia demand grew at a rate of 3.2% per year from 1993 to 1998, but the growth rate is expected to slow to 0.5% per year from 1998 to 2002 [34]. US capacity declined in the 1970s, and it continued to decline during the 1990s by 5–10% to about 13 million tons/year. During the 1990s, US ammonia capacity was between 13 and 14 million tons/year with operating rates over 95%. The US capacity is forecast to reach 15 million tons/year by 2002, and worldwide ammonia capacity is forecast to grow at over 2% per year and exceed 140 million tons/year by 2002 [31].

By 2002, 30% of North American ammonia production was curtailed, and in 2004, four more plants were shut down. Although periodic downtimes of plants are not uncommon, the combination of import competition that depressed prices and very high feedstock costs in many industrialized countries has resulted in shutdowns of high-cost plants which resulted in a decline in world capacity [35–37].

In 1998, most of the ammonia production capacity was located in the regions shown in Table 22.3 [32, 36].

Table 22.3 Ammonia capacity by region (thousands of metric tons *N*)

	1987		1998	1999		2005	
Region	Quantity	Share (%)	Share	Quantity	Share (%)	Capacity	Share (%)
China	18,675	16.9	22	30,450	23.6	33,460	24.6
Former Soviet Union	21,725	19.7	17	19,340	15.0	18,455	13.6
North America	16,390	14.8	15	18,955	14.7	18,410	13.6
South Asia	8,935	8.1	12	15,750	12.2	16,705	12.3
Western Europe	15,635	14.1	9	11,870	9.2	11,255	8.3
Middle East	4,100	3.7	7	5,950	4.6	7,795	5.7
Central Europe	9,830	8.9	6	7,560	5.9	6,820	5.0
Indonesia and Japan	5,800	5.3	4	7,725	6.0	8,340	6.1
Mexico and the Caribbean (including Venezuela)	5,705	5.2	3	6,415	5.0	8,265	6.1
Other countries	3,700	3.4	5	4,790	3.7	6,330	4.7
Totals	*110,495*			*128,805*		*135,835*	

Table 22.4 Feedstock economics for ammonia production: relative consumptions and economics for different feedstocks

	Natural gas	Heavy oil	Coal
Energy consumption	1.0	1.3	1.7
Investment cost	1.0	1.4	2.4
Production cost	1.0	1.2	1.7

Source: EFMA. Reproduced by permission of European Fertilizer Manufacturers Association

About 90% of world ammonia production is processed or used in the countries where it is produced. The remaining 10% of production enters international trade. The major net exporters in 1997 were Russia and Ukraine (43% of world exports), Trinidad (22%), and the Middle East (13%). The major net importers were the United States (43% of world imports) and Western Europe (26%) [32].

In the mid-1990s, the ammonia industry accounted for about 5% of worldwide natural gas consumption. For economic and environmental reasons, natural gas is the feedstock of choice. However, processes for ammonia production can use a wide range of energy sources. For example, 60% of China's nitrogen fertilizer production is currently based on coal. At present, natural gas is the most economic feedstock for the production of ammonia as shown in Table 22.4 [38].

All commercial ammonia manufacturing processes are based on the synthesis of ammonia from nitrogen and hydrogen. Thus, the many processes are differentiated by the method of producing the hydrogen and nitrogen for the synthesis. Preparation of nitrogen can be by separation from air in a standard "air" plant. Or air can be fed to the ammonia unit where the oxygen is used to burn a portion of the hydrocarbon feed, leaving the nitrogen for ammonia synthesis. The latter is the dominant type. Most of the hydrogen for ammonia manufacture is prepared from hydrocarbon feeds.

Some physical properties of ammonia and densities of aqueous solutions of ammonia are given in Tables 22.5 and 22.6, respectively.

Table 22.5 Physical properties of ammonia

Property	Value
Molecular weight	17.03
Boiling point, °C	−33.35 @ 760 mmHg
Freezing point, °C	−77.7
Color	Colorless
Critical temperature, °C	133.0
Critical pressure, kPa (psi)	1,425 (1,657)
Specific heat, J/kg K	
0°C	2,097.2
100°C	2,226.2
200°C	2,105.6
Heat of formation of gas, ΔH, kJ/kg mol (BTU/lb mol)	
0 K	−39,222 (−16,862)
298 K	−46,222 (−19,872)
Solubility in water (wt.%)	
0°C	42.8
20°C	33.1
25°C	31.8
40°C	23.4
60°C	14.1
Specific gravity[a]	
−40°C	0.690
0°C	0.639
40°C	0.580
Vapor pressure, mmHg	7,500 @ 25°C
Vapor density	0.6 (air = 1.0) at 0°C
Flammable limits in air (% by volume)	
Lower explosion limit (LEL)	15
Upper explosion limit (UEL)	28
Autoignition temperature	651°C (1,204°F)

[a] Anhydrous ammonia

Hydrogen Production

Ammonia production is the single largest consumer of hydrogen for use in chemical manufacture. Hydrogen production is typically integrated into the overall plant design, so most ammonia producers regard themselves as consumers of natural gas (or another feedstock) rather than hydrogen.

Table 22.6 Densities of aqueous ammonia at 15°C

Ammonia (wt.%)	Density (g/L)
8	0.970
16	0.947
32	0.889
50	0.832
75	0.733
100	0.618

Worldwide, a few ammonia plants have been located near by-product sources of hydrogen. In the United States, Coastal Refining & Marketing, Inc. brought a new ammonia plant onstream in late 1997 in Freeport, Texas, using purchased pipeline hydrogen. Air Liquide America purifies 45 million SCF/day of by-product hydrogen from Dow Chemical Company for pipeline delivery to Coastal's Freeport facility. In 2002, this was the only US ammonia plant using purchased hydrogen [39]. Hydrogen requirements in the main world regions are shown in [39, 40], and more details are given in Chap. 27.

To help meet increased hydrogen requirements, a methodology has been developed for assessment of hydrogen sources that is based on an analogy with the problem of process heat recovery. A hydrogen surplus diagram allows the engineer to find the "hydrogen pinch" and to set targets for hydrogen recovery, hydrogen plant production, and import requirements. This method also gives insights into the effective use of hydrogen purification units. It has been shown that a purification unit should not be placed below the hydrogen pinch. Purifying gas above the pinch may have some benefits, but placing the purifier across the pinch is the best option [41].

Hydrogen pinch technology has been applied commercially by several companies including AspenTech, BP-Amoco, Engineers India, Exxon, ICI, Linnhoff-March, M. W. Kellogg, and UOP [42, 43].

Hydrogen is a nontoxic, colorless, odorless, and tasteless gas. It is the lightest and most abundant element (making up over 90% of the atoms in our universe), but it is present at only extremely low levels (0.1 ppm) as a pure element in the earth's atmosphere. More than 50% of the atoms in our environment are hydrogen.

Hydrogen is present in fossil fuels and water in sufficient amounts that it can be produced from these sources on a large scale by several different methods: (1) petrochemical processes, (2) coal-based chemical processes, and (3) electrochemical processes (electrolysis).

In Table 22.7, the percentage of hydrogen production is broken down by type of manufacturing process [44].

In the petrochemical processes, both natural gas and crude oil fractions can be converted into synthesis gas using two basically different methods [44]:

Table 22.7 World hydrogen production and type of manufacturing process (in %) [44]

Process	1974	1984	1988
Cracking of crude oil	48	77	80
Cracking of natural gas	30		
Coal and coke gasification	16	18	16
Electrolysis	3	4	4
Miscellaneous processes	3	1	

Source: Copyright by VCH Publishers and reproduced by permission of the copyright owner

Table 22.8 H_2/CO molar ratio in synthesis gas

Method of manufacture	H_2/CO molar ratio
Oxygen-coke-steam	0.6
Air-coke-steam	0.9
Oxygen-coal-steam	1.0
Oxygen-fuel oil-steam	1.0
Propane-steam	1.33
Methane-oxygen	1.7
Oil-steam	2.1
Petroleum ether-steam	2.4
Methane-steam	3.0–5.0

- With the allothermal steam reforming method, catalytic cracking takes place in the presence of water vapor. The necessary heat is supplied from external sources.
- With the autothermal cracking process, heat for the thermal cracking is supplied by partial combustion of the feed. Water is used and carbon dioxide (CO_2) may be recycled to attain a desired CO/H_2 ratio.

The processes for the manufacture of synthesis gas were originally based on the gasification of coke from hard coal and low-temperature coke from brown coal by means of air and steam. After World War II, the easy-to-handle liquid and gaseous fossil fuels—oil and natural gas—were also employed as feedstocks. Their value lies in their high hydrogen content that produces higher molar ratios of hydrogen (H_2) to carbon monoxide (CO) in the synthesis gas (see Table 22.8). By using excess steam in the reforming of methane, the H_2-to-CO molar ratio can be as high as 5.0 (see Table 22.8).

Selection of a process for hydrogen manufacture from hydrocarbons and coal therefore depends on the raw material and its cost, the scale of operation, the purity of the synthesis gas to be produced, the pressure level of the natural gas feed, and the number and type of downstream processes that will consume the carbon monoxide and hydrogen.

Hydrogen is also manufactured industrially by direct electrolysis of H_2O, HF, and 22–25% hydrochloric acid (HCl). However, the hydrogen produced by electrolysis accounts for a small percentage of the total H_2 production (see Table 22.7). In contrast to the steam reforming of

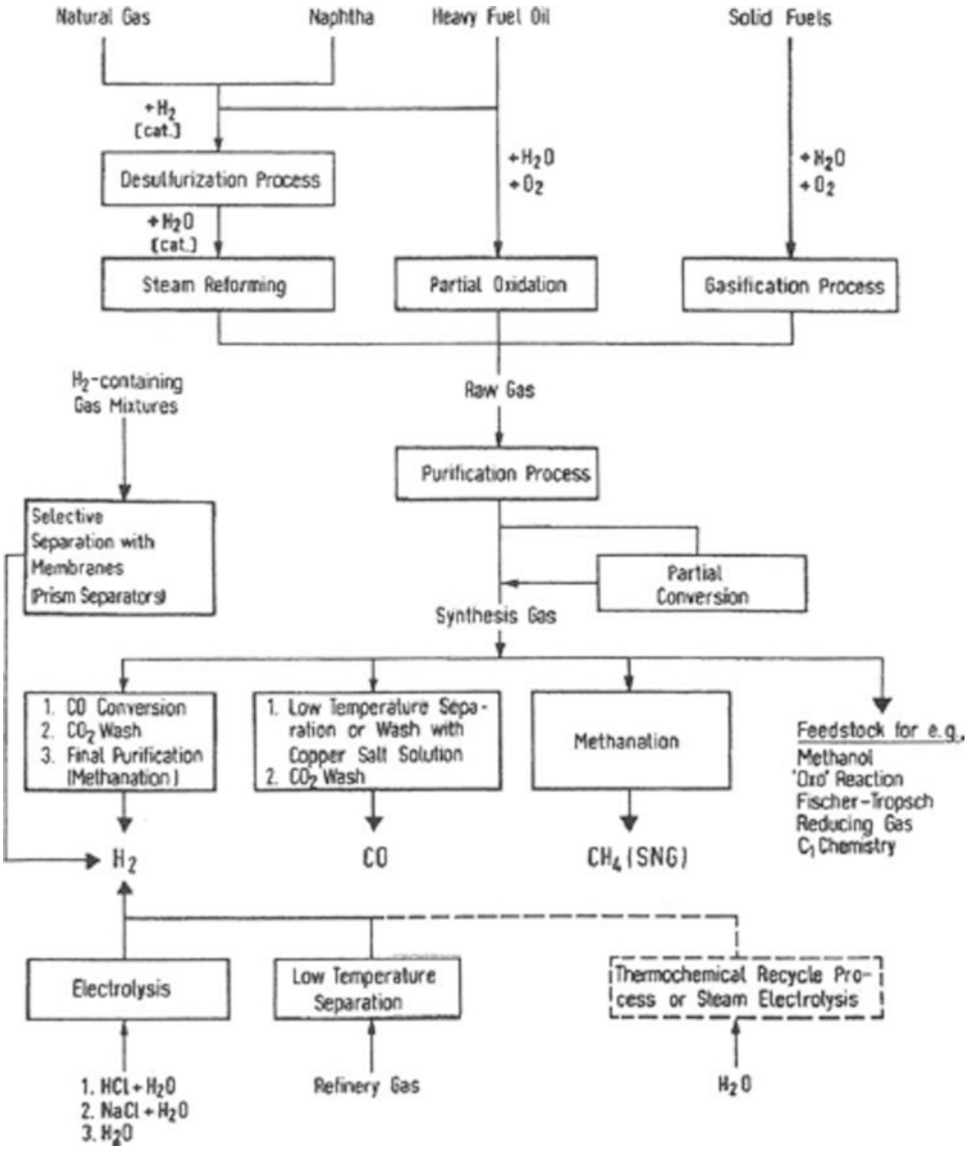


Fig. 22.3 Hydrogen manufacturing process steps [44] (Copyright by VCH Publishers and reproduced by permission of the copyright owner)

hydrocarbons, the hydrogen from electrolysis is very pure (>99 vol.%) which eliminates the costly purification steps.

Manufacturing Processes

Hydrogen is manufactured by four principal processes (see Table 22.7 and Fig. 22.3): steam reforming of natural gas, partial combustion of natural gas or oil with pure oxygen, gasification of coal or coke with air (or oxygen) and steam, and recovery of by-product hydrogen from petroleum refinery gases or other cracking operations. Small amounts of hydrogen also are manufactured by electrolysis. These processes are discussed in more detail in the following sections.

Reforming. Reforming is a general name for the reaction of a hydrocarbon, such as methane, with water and/or carbon dioxide, to produce a mixture of carbon monoxide and hydrogen. The different reforming processes can be divided into the following types [38]:

- Conventional steam reforming with a fired primary reformer and stoichiometric air secondary reforming (stoichiometric H/N ratio)
- Steam reforming with mild conditions in a fired primary reformer and excess air secondary reforming (under-stoichiometric H/N ratio)
- Heat exchange autothermal reforming (ATR), with a process gas heated steam reformer (heat exchange reformer) and a separate secondary reformer, or in a combined autothermal reformer, using excess or enriched air (under-stoichiometric or stoichiometric H/N ratio)

The typical steps in the reforming process and how they tie into ammonia production are shown in Fig. 22.4.

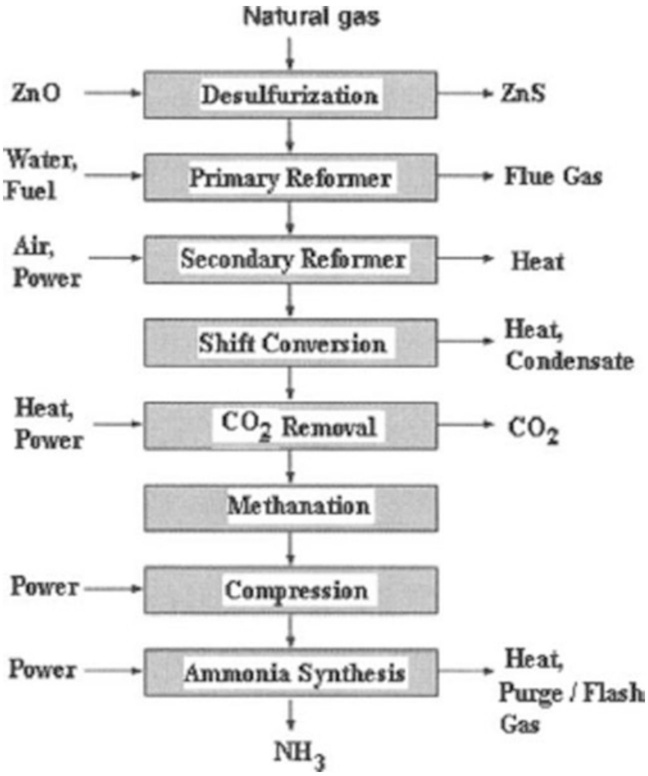


Fig. 22.4 Block diagram of the steam/air reforming process [38] (Used by permission of European Fertilizer Manufacturers Association)

If water is used, the process is called steam reforming or steam cracking. The reforming reaction (22.1) is endothermic and requires a catalyst [3].

$$CH_4 + H_2O \rightleftarrows 3H_2 + CO \quad \Delta H = +49 \text{ kcal/mol or } 205 \text{ kJ/mol} \tag{22.1}$$

Other reactions that proceed at the same time as the reforming reaction are [44]

$$CO + H_2O \rightleftarrows H_2 + CO_2 \quad \Delta H = -10 \text{ kcal/mol or } 42 \text{ kJ/mol} \tag{22.2a}$$

(The Homogenous Water–Gas Reaction or Water–Gas Shift)

$$CH_4 \rightleftarrows 2H_2 + C \quad \Delta H = +17.9 \text{ kcal/mol or } 75 \text{ kJ/mol} \tag{22.3}$$

$$2CO \rightleftarrows CO_2 + C \tag{22.4}$$

Boudouard reaction

$$\Delta H = -41.4 \text{ kcal/mol or } 173 \text{ kJ/mol}$$

The equilibrium composition of the synthesis gas depends on the steam-to-gas ratio entering the reactor, the reaction temperature, the reaction pressure, and the quantity of inerts in the reaction mixture. To avoid carbon formation as indicated by (22.3) and (22.4), the steam-to-gas ratio must be kept high enough to favor the reforming reaction (22.1) and the water–gas shift reaction (22.2a) over the reactions that form carbon.

In 2003, Synetix, a subsidiary of Johnson Matthey, offered for license a large-scale steam reforming process. In this process, hydrocarbon feeds with boiling points up to 200°C can be treated. The process consists of three steps: feed pretreatment, catalytic reforming, and reforming of residual methane. The advantage of the Synetix process is that no soot forms, even with liquid crude oil fractions as feed. This makes catalyst regeneration unnecessary [44]. Because of these advantages, the Synetix technology is used in over 400 reformers in over 30 countries. When North Sea gas and other natural gas reserves around the world were developed, the use of naphtha as a feedstock declined and the need for this technology was reduced.

Feed Pretreatment

The catalysts used in the steam reforming process are poisoned by trace components in the hydrocarbon feed—particularly sulfur, chlorine, and metal compounds. The best way to remove sulfur compounds is to convert the organic sulfur species to H_2S over a hydrodesulfurization catalyst. The next step is sulfur removal with an absorbent. The same catalyst can usually convert any organochloride species to give HCl and also act as an absorbent for most problematic metal species. A second absorbent is used for chloride removal [45].

Prereformer

A prereformer partially completes the steam reforming reactions upstream of the main steam reformer at a much lower temperature by using a more active catalyst. One advantage in new plant designs is that the conventional steam reformer furnace can be smaller [45]. All the ethane, the heavier hydrocarbons, and some of the methane are decomposed endothermically and at temperatures below those in the main reformer. The product gas from the prereformer is an equilibrium mixture of hydrogen, methane, steam, and carbon oxides. The prereformer can reduce fuel consumption by up to 10% and increase throughput by up to 15% [46].

Another advantage is that a prereformer may be preferred in flowsheets for processing heavier feedstocks.

The prereformer converts heavier hydrocarbons to methane before they are fed to the steam reformer. This allows the steam reformer to be designed for methane service. The prereforming catalysts are very sensitive to poisons. Therefore, a good feed purification section is essential [45].

Reformers

Steam reforming refers to the endothermic catalytic conversion of light hydrocarbons in the presence of steam. Catalytic reforming takes place across a nickel catalyst that is packed in tubes in a fired furnace (the "primary reformer") and in the refractory-lined chamber reactor (the "secondary reformer"). This is where hot process air is added to introduce nitrogen into the process. Typical reaction conditions in the primary reformer are 700–830°C and 15–40 bar. Additional details about the primary reformer can be found in [44]. A portion of the primary reformer discharge gas is burned in the secondary reformer with added air or oxygen to give a gas mixture at a temperature of over 1,200°C. Almost all of the methane reacts with steam at this temperature. The methane concentration depends upon the pressure, the temperature, and the quantity of nitrogen and steam present [44].

Reformer Catalysts

The secondary reformer catalyst does not need to be as active as that in the primary reformer. The usual nickel concentration is about 15% in the secondary reformer and 25% in the primary reformer catalyst.

The steam reforming catalyst selection depends on furnace design, feedstock types, and the operating conditions. The main categories of catalyst are for (1) light-duty reforming, (2) intermediate-duty reforming, and (3) heavy-duty reforming. The light-duty catalysts are for natural gas, refinery off-gas, and preformed feeds [45].

Intermediate-duty catalysts are for feeds with a significant content of components from ethanes up to liquid petroleum gas (LPG). Heavy-duty catalysts are primarily for naphtha feeds that have even more tendency for carbon deposition [45].

Silica and silica-bearing materials cannot be used with the shift catalyst. This is because the silica volatilizes and migrates from the hotter zone to lower-temperature zones downstream. Usually it deposits on the waste heat boiler tubes after the secondary reformer.

Normally, the nickel oxide is reduced to nickel and water by the hydrogen that is produced in the operation. In some cases, the reduced nickel can be reoxidized to nickel oxide when large amounts of steam and small amounts of H_2 are present [47]:

$$\mathrm{Ni} + \mathrm{H_2O} \rightleftarrows \mathrm{NiO} + \mathrm{H_2} \tag{22.5}$$

The addition of H_2 to the natural gas feed keeps the nickel in the reduced state, which makes it more active. Also, the hydrogen will retard the formation of nickel sulfide and prevent or minimize poisoning of the catalyst:

$$\mathrm{NiS} + \mathrm{H_2} \rightleftarrows \mathrm{Ni} + \mathrm{H_2S} \tag{22.6}$$

Carbon can also reduce the effectiveness of the catalyst. When conditions are favorable for the following reactions, carbon will deposit on the catalyst:

$$\begin{gathered}\mathrm{CH_4} \rightleftarrows 2\mathrm{H_2} + \mathrm{C} \\ \Delta H = +17.9\ \text{kcal/mol or } 75\ \text{kJ/mol}\end{gathered} \tag{22.3}$$

$$\begin{gathered}2\mathrm{CO} \rightleftarrows \mathrm{CO_2} + \mathrm{C} \\ \Delta H = -41.4\ \text{kcal/mol or } 173\ \text{kJ/mol}\end{gathered} \tag{22.4}$$

Thus, if insufficient steam is present, carbon will be deposited on the catalyst and reduce its effectiveness. Carbon will gasify with steam at 800°C or higher in the reformer:

$$\begin{gathered}\mathrm{C} + \mathrm{H_2O} \rightleftarrows \mathrm{H_2} + \mathrm{CO} \\ \Delta H = +28\ \text{kcal/mol or } 119\ \text{kJ/mol}\end{gathered} \tag{22.2b}$$

Reformer operation and changes in operating rate should always be carried out in a way that ensures sufficient steam is over the catalyst. Thus, if the feed rate is to be increased, the steam rate should be increased first. Conversely, if the feed rate is to be decreased, the natural gas rate should be decreased first.

Reformer Materials of Construction

The primary reformer is a steam–hydrocarbon reforming tubular furnace that is typically externally fired at 25–35 bar and 780–820°C on the process side. From the 1950s through the 1960s, SS 304, SS 310, SS 347, HK 40 alloy, and HP 25/35 modified alloys were used as tube materials. However, these materials developed various operating problems as rates increased and longer service lives were needed for economical operation [48].

HP microalloys were developed during the 1990s. The microalloys enhanced carburization resistance and improved high-temperature creep–rupture resistance [48].

For reformer outlet manifolds, the normal metallurgy choice is a wrought type of Alloy 800 H. Hot reformed-gas transfer lines are usually refractory-lined with an interior of Alloy 800 sheathing [48].

In the secondary reformer, air is added to the process stream at operating conditions of 28–30 bar and 955–1,025°C. The refractory-lined vessel has an outer shell of a low-alloy steel containing 0.5 month. Metal dusting occurs in the secondary reformer outlet sections. With hot gases containing a high CO content, carbon will diffuse into the Fe–Cr–Ni alloy. This phenomenon can lead to local mechanical fracturing of surface layers and failures by pitting [48].

Materials such as SS 304 and Alloy 800 are very susceptible to metal dusting in the range of 500–800°C. Besides temperature, carbon activity (the CO/CO_2 ratio in the gas) and CO partial pressure also affect metal dusting. Severe attacks occur when the carbon activity is in the range of 3–10. Recirculating CO_2 into the primary reformer along with feedstock can maintain a low CO/CO_2 ratio and avoid the severity of this attack. By maintaining a high steam-to-hydrogen ratio in the gas, the metal dusting can also be minimized [48].

Hydrogen embrittlement is another important corrosion problem that is encountered in reformed-gas pipelines. The Nelson curves list the operating limits that should be followed to avoid decarburization and fissuring of steel in hydrogen service [48].

Waste Heat Recovery

The waste heat recovery system is associated with flue gas from the reformer furnace and process gas from the secondary reformer. It generates high-pressure steam in specially designed boilers. Proper material selections and stringent water quality control are two proactive loss-prevention methods [48].

Hydrogen Production Costs and Capital Costs

The theoretical energy requirement per mole of hydrogen produced for the overall steam reforming process is 40.75 kJ/mol of hydrogen [49]. The capital cost for a 60 million SCF/day hydrogen plant, based on steam methane reforming technology, is US$ 1.00 to US$ 1.50 per SCF of hydrogen/day [50].

The cost for making hydrogen by steam reforming of natural gas depends primarily on the cost of natural gas. Several sources estimate the hydrogen production cost, excluding capital charges (in US$ per million BTU of H_2), to be 150% of the natural gas cost (in US$ per million BTU) [50–52].

Other Reforming Processes

The RKN process uses steam reforming to make hydrogen from hydrocarbon gases. This process was developed by Haldor Topsøe in the 1960s. By 1974, 24 plants based on this technology were operating [53].

The SMART (*S*team *M*ethane *A*dvanced *R*eformer *T*echnology) process makes hydrogen by the steam reforming of methane, and it houses the catalyst in a proprietary heat exchanger. The process was developed by Mannesmann KTI in 1996, and the first installation was started up in Maryland in 1998 [53].

The ATR (*a*utothermal *r*eforming) process makes CO-enriched syngas. It combines partial oxidation with adiabatic steam reforming. It was developed in the late 1950s for ammonia and methanol synthesis and then further developed in the 1990s by Haldor Topsøe [53].

The CAR (*c*ombined *a*utothermal *r*eforming) process is used to make syngas from light hydrocarbons, and the heat is provided by partial oxidation in a section of the reactor. It was developed by Uhde and commercialized in Slovakia in 1991 [53].

The KRES (*K*ellogg *r*eforming *e*xchanger *s*ystem) is a reforming process that provides syngas to the KAAP process. The KAAP process is a high-pressure process that makes ammonia from its elements and does not use an iron-containing catalyst. MW Kellogg developed the process in 1990, and Ocelot Ammonia installed the first plant in British Columbia, in 1994 [53]. Another KRES plant started up in 2003 at an NH_3 plant owned by Liaotong in Liaoning Province, China [54].

The *H*aldor *T*opsøe *c*onvection *r*eformer (HTCR) is a relatively small piece of equipment that combines the radiant and waste heat sections of the conventional reformer. It uses PSA to make 99.9% hydrogen purity. It is best for small- and medium-sized hydrogen plants (500–10,000 Nm^3/h) [55].

Reliability and Revamps

Some of the revisions to conventional reforming are listed below.

Decreased Firing in the Primary Reformer. Decreased heat supply in the primary reformer means that the process outlet temperature is lowered to about 700°C, the firing efficiency increases, and the size and cost of the primary reformer are reduced. The milder operating conditions prolong catalyst life, catalyst tube life, and outlet header service life [38].

Increased Process Air Supply to the Secondary Reformer. Decreased heat supply in the primary reformer means that increased internal firing is necessary to achieve approximately the same degree of total reforming. A somewhat higher methane slip (and thus a lower secondary reformer outlet temperature) is acceptable and preferable in this type of process. This is because methane is removed in the final purification [38].

The process air requirement is about 50% higher than in the conventional process. This means increased compression capacity and energy. The process air compressor is usually driven by a gas turbine with the exhaust gas from the turbine being used as combustion air in the primary reformer. Some excess steam is available for export when using a gas turbine [38].

Reformer. Where an increase of 10–50% of existing steam reforming capacity is required, revamping is often more cost effective than a new plant or the purchase of hydrogen over the fence. In most revamps, the reaction area is only slightly affected by operating rate, and existing reactors are often able to operate at increased rate. Existing heat exchangers are often adequate as capacity is increased. However, air coolers in a CO_2 removal unit often need to be modified by adding surface area or increasing fan horsepower [56].

Separation equipment tends to have fixed limits and can be costly to change. However, more capacity can often be obtained by changing column internals or the solvent composition [56].

Radiant Section. Changes in the radiant section can be difficult to justify based on capacity alone. However, if the radiant tubes are near the end of their useful life, an upgrade to this section may provide more capacity. Changes in metallurgy since the mid-1970s have allowed changes in radiant section operating conditions.

Time period	Tube material	Operating pressure (psig)	Operating temperature (°F/°C)
1970s	HK 40	300	1,450/790
1980s	HP	325	1,550/845
1990s	Microalloy	450	1,575/855

The microalloy tubes allow increased flux rates and higher reformer outlet temperatures. This in turn can make it possible to reduce the steam-to-carbon ratio, while the hydrogen purity remains the same [56].

Convection Section. The effects on the reformer of increased gas flow and temperature are multiplied in the convection section because additional load is placed on convection coils, fans, and the steam system. These areas commonly limit the operating rate of the reformer. The tube supports also need to be considered because they are exposed to hot flue gas without the cooling effect of process fluids [56].

One of the most effective reformer modifications is to use heat from the convection section to preheat radiant section feed. This will reduce radiant section heat load, reduce radiant section firing rate, and potentially unload other areas such as steam generation. This option has been used to increase capacity by 10% without increasing the arch temperature in the radiant section [56].

Combustion Air Preheat. Combustion air preheat reduces the flue gas flow through the furnace which unloads the fans and the entire convection section. The best known option is to use heat exchange with flue gas. For new units, this can have the advantage of very high efficiency because the heat sink is ambient air. Another option is the use of steam to preheat air. This option was used to obtain a 12% increase in capacity at the same fuel firing rate and combustion air flow. A third option is to use circulating boiler feedwater to preheat the air [56].

Heat Exchange Autothermal Reforming. A new development in the late 1980s was heat exchange autothermal reforming. In this process, the heat content of the secondary reformer gas is used in a primary reformer with a new design. This reformer is a gas-heated, heat exchange reformer rather than the conventional fired furnace design. Surplus air or oxygen-enriched air is required in the secondary reformer to meet the heat balance in this autothermal concept [38].

Emissions to the atmosphere are reduced significantly by eliminating the flue gas from the primary reformer. NO_x emissions may be reduced by 50% or more compared to conventional steam. The level of NO_x reductions depends on the extent of auxiliary combustion in the plant [38].

Two processes of this kind are in operation, and some others are at the pilot stage. As of 1999, single-train capacities have been limited to 1,000 tons/day or less [38], and four plants had been installed that use Synetix gas-heated reformer (GHR) technology [57].

Synetix developed a modified design of the GHR known as the advanced GHR or AGHR. BHP Petroleum of Australia became the first operator of an AGHR in 1998. The key difference in the AGHR is that it uses a single-pass tube arrangement, whereas the GHR uses a bayonet tube reformer. Based on the experience at BHP Petroleum, the AGHR design results in a reformer that is lower in cost, easier to operate, easier to fabricate, and allows scale-up to capacities in excess of current world-scale throughputs (see Fig. 22.5) [57].

A hydrogen plant based on steam reforming with PSA was installed in Venezuela at the Compagnia Hidrogeno de Paraguana (CHP) plant to supply hydrogen to a refinery. The plant produces 50 million SCFD of hydrogen at 99.5% purity and 400 psig. The plant is a joint venture between BOC and Foster Wheeler and was started up in September 1997. This plant uses 4% less in total fuel for firing and has an installed cost that is approximately 23% below that of other plants of the same size [58].

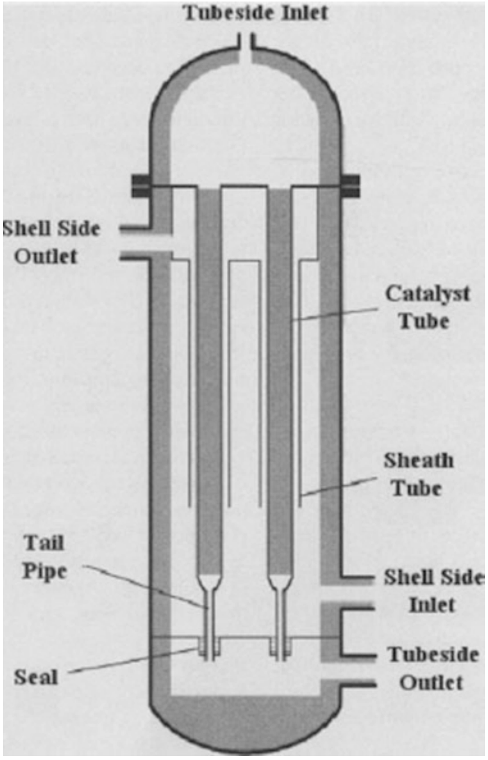


Fig. 22.5 Advanced Gas Heated Reformer (AGHR) from Synetix (Reproduced by permission of Johnson Matthey Catalyst. Copyright Johnson Matthey PLC)

Small-Scale Reforming

Small-scale reforming systems are being pursued by a number of companies developing fuel cell electrical generation systems for home and automotive fuel cell applications. Before these systems become affordable and simple enough to be used as home electrical system components, they may be cost effective for industrial hydrogen supply systems [59].

Small-scale reforming systems are relatively complex because they need fuel and air-feed systems, the reformer, a hydrogen purification system, and various cooling and water processing ancillary systems to make it all work. The systems also have to employ a specific hydrocarbon that is available at a reasonable cost at a customer's location. These systems probably work best for customers with hydrogen consumption rates in the 1,500–10,000 SCF/h range. They are most cost effective when employed to produce hydrogen with a 99.9% or lower assay. As of 2007, small-scale reforming systems have not been commercialized. But they are in the advanced development stage [59].

Technology Suppliers

Companies that provide technology to produce hydrogen via steam reforming are in the United States, Europe, and Japan. Many of these companies are listed in [39].

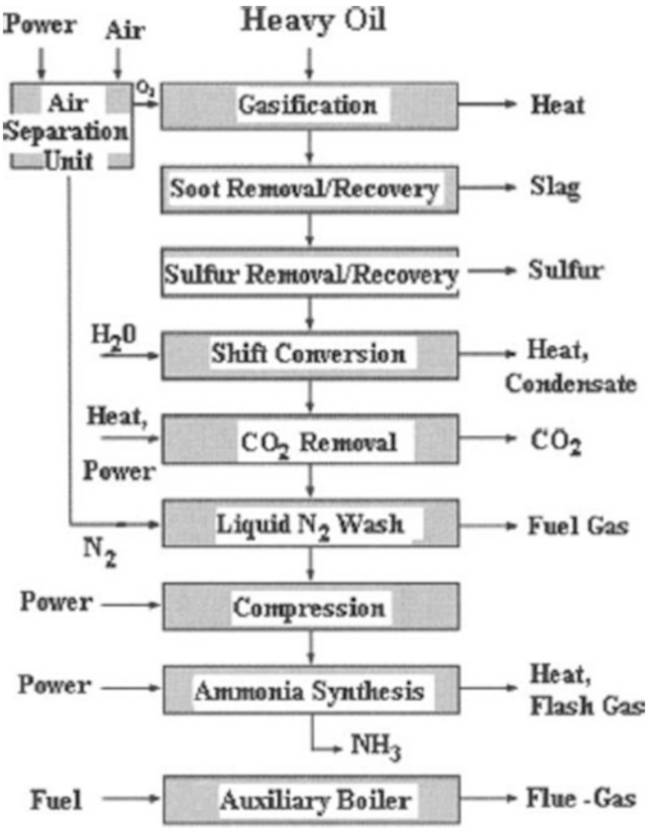


Fig. 22.6 Block diagram of partial oxidation process [38] (Used by permission of the European Fertilizer Manufacturers Association)

Partial Oxidation. Noncatalytic partial oxidation processes (see Fig. 22.6) react hydrocarbons with gaseous oxygen at very high temperatures in a refractory-lined, pressurized reactor to produce synthesis gas. Although the process can handle a wide range of feedstocks, it was developed mainly to utilize the low-value "bottom of the barrel" and waste feedstocks that typically have high-sulfur and high-metal contents [39].

Synthesis gas manufacture by partial oxidation or autothermal cracking of crude oil fractions was developed by BASF/Lurgi, Texaco and Hydrocarbon Research. Heat for the thermal cracking is supplied by partial combustion of the feed in the presence of water. Recycled CO_2 may also be added to the combustion to attain a desired CO/H_2 ratio. Shell developed a modified version (which does not use a catalyst) that is referred to as the gasification process [44].

As illustrated in Table 22.8, the H_2-to-CO molar ratio in the synthesis gas product stream depends on the raw materials and the operating conditions. By burning natural gas with a limited quantity of oxygen, a synthesis gas that has an H_2-to-CO molar ratio of approximately 2:1 can be produced:

$$CH_4 + \frac{1}{2}O_2 \rightarrow 2H_2 + CO \qquad (22.7)$$

Table 22.9 Product gas compositions from various feeds by partial combustion (percent by volume—dry)

Feedstocks and gas components	Nat. gas	Light naphtha	Heavy fuel oil	Vacuum resid.	Propane asphalt	Coal
Hydrogen	61.80	52.09	46.74	44.80	43.69	34.39
Carbon monoxide	33.75	42.59	48.14	49.52	50.09	44.22
Carbon dioxide	3.10	4.88	3.80	4.16	4.45	18.64
Methane	1.00	0.30	0.30	0.30	0.30	0.38
N_2 + A	0.27	0.13	0.23	0.18	0.30	0.68
H_2S	–	0.01	0.76	1.00	1.12	1.46
COS	–	–	0.03	0.04	0.05	0.10
NH_3	–	–	–	–	0.13	–
H_2/CO mol/mol	1.83	1.22	0.97	0.09	0.87	0.78

If steam is introduced into the reaction, the H_2-to-CO molar ratio can be increased to over 2.0. This reaction is exothermic and can reach temperatures of 1,200–1,500°C [60]. Typical operating costs are shown in [39].

If fuel oil or crude oil is used as the raw material in partial combustion, care must be taken to provide for the removal of sulfur compounds and ash-containing materials in the crude oil (see Fig. 22.6). With crude-oil or fuel-oil partial combustion, the carbon content is much greater than in the case of natural gas, and special design considerations are necessary to produce a satisfactory gas.

Typical gas compositions from the partial combustion of various fuels are given in Table 22.9. The nitrogen that is needed to produce the desired H_2/N_2 ratio for ammonia production is usually introduced later in the processing sequence.

Partial Oxidation Processes

The Multipurpose Gasification (MPG) process uses partial oxidation with oxygen to generate synthesis gas from different hydrocarbon feeds. These feeds include natural gas, tars, and other coal gasification residues, refinery residues, asphalts, coal and coke slurries, and chemical wastes. MPG is an updated development based on technology that Lurgi acquired in 1997 from SVZ [61].

The MPG process preheats a gaseous feedstock with hot raw gas and an optional fired preheater. Preheat level and installation of the fired heater are determined by energy and cost optimization, which is based on the relative values of feed, fuel gas, and oxygen. Feed gas and oxygen enter the reactor with a minor amount of steam via the gas burner. Gasification occurs in a refractory-lined reactor at temperatures between 1,200 °C and 1,400°C [61].

The noncatalytic partial oxidation of hydrocarbons by the Shell gasification process (SGP) takes place in a refractory-lined reactor that uses a specially designed burner. The oxidant is preheated and then mixed with steam before it is fed to the burner. The feedstock gasification conditions range from 1,200°C to 1,400°C at a pressure of 50–70 bar, depending on the feedstock [62, 63].

The THGP (*T*exaco *H*ydrogen *G*eneration *P*rocess) makes pure, high-pressure hydrogen from various gaseous and light hydrocarbons. This process can be licensed from the Texaco Development Corporation [3, 64].

Technology Suppliers

The companies that license partial oxidation technology can be found in the United States and Europe. Some of these companies are listed in [39].

Coal and Coke Gasification. In coal gasification, the exothermic partial combustion of carbon and the endothermic water gas formation represent the actual gasification reactions [44]:

$$C + O_2 \rightleftarrows 2CO \quad \Delta H = -60\ \text{kcal/mol or 246 kJ/mol} \\ \text{Partial combustion} \tag{22.8}$$

$$C + H_2O \rightleftarrows H_2 + CO \quad \Delta H = +28\ \text{kcal/mol or 119 kJ/mol} \\ \text{Heterogeneous water gas reaction} \tag{22.2b}$$

Some other important reactions are

$$C + CO_2 \rightleftarrows 2CO \quad \Delta H = +38\ \text{kcal/mol or 162 kJ/mol} \\ \text{Boudouard reaction} \tag{22.9}$$

$$CO + H_2O \rightleftarrows H_2 + CO_2 \quad \Delta H = -10\,\text{kcal/mol or 42 kJ/mol} \\ \text{Homogeneous water gas reaction (water gas shift)} \tag{22.2a}$$

$$C + 2H_2 \rightleftarrows CH_4 \quad \Delta H = -2\,\text{kcal/mol or }87\,\text{kJ/mol} \tag{22.10}$$

Hydrogenative gasification

$$CO + 3H_2 \rightleftarrows H_2O + CH_4 \quad \Delta H = -49\,\text{kcal/mol or }206\,\text{kJ/mol} \tag{22.11}$$

Methanation

Gasification processes can be characterized by the type of coal used and by the coal's physical and chemical properties [44]. Some processes use external heating and others use self-heating. The reactors may be fixed bed, fluidized bed, or entrained bed. The coal gasification processes are described in more detail under "Coal Technology" (Chap. 19).

Sasol (*S*uid-*A*frikaans *S*intetiese *O*lie) in South Africa has one of the largest coal gasification production operations in the world with the three plants [3]. One plant produces only chemicals and the other plants produce both liquid fuels and chemical feedstocks from coal. In 1994, Sasol's total ammonia capacity was over 500,000 tons/year [65].

One of the first coal gasification plants in the United States was the Great Plains Synfuels plant that is operated by Dakota Gasification in Beulah, ND. This plant consumes over 5.5 million tons of coal per year and produces over 54 billion SCF (1.4 billion Nm^3) of natural gas, 365,000 tons/year of ammonia, 24 million gal of liquid nitrogen per year (68 million kg or 150 million lb) plus several other chemicals [66] (see http://www.dakota.com for more details).

In 2000, Farmland Industries in Kansas started up a petroleum coke gasification plant to make ammonia. The plant uses Texaco's process to make 1,000 tons/day of ammonia [67, 68]. In 2004, Rentech purchased an 830-tons-per-day ammonia plant in Illinois and plans to convert it from natural gas to coal feedstock. Startup of the conversion is expected in 2007 [69].

By-product Hydrogen. Large volumes of by-product hydrogen are generated from a variety of production processes. Some of these processes are listed in [39]. Some of the processes that can be used to obtain purified hydrogen are summarized below.

The HyTex (*H*ydrogen *T*exaco) process makes pure hydrogen from waste gases in oil refineries in a three-step process [3]. The process was developed by Texaco and announced in 1991. The first commercial unit started up in Anacortes, WA, in 1993.

Platforming operations make a gas that contains as much as 90–95% hydrogen. This gas is usually purified by low-temperature fractionation or washing with liquid nitrogen.

Kvaerner Oil & Gas, Norway, has introduced a process for manufacturing carbon black that generates substantial quantities of by-product hydrogen. In 1998, the process was installed in Montreal, Québec (Canada), to produce 20,000 tons of carbon black and 50 million Nm^3/year (5.2 million SCF/day) of hydrogen. The process appears to be economic only when the carbon black can also be sold as an end product [39].

Air Products and Chemicals, Inc. has been selected to supply a hydrocarbon and nitrogen recovery system for a new polyethylene manufacturing plant in Baytown, TX. The plant will be owned by Chevron Phillips Chemical Company and Solvay Polymers, Inc. The recovery system uses partial condensation in conjunction with Air Products' PSA technology to recover hydrocarbons in the polyolefin plants and recycle nitrogen with a purity of greater than 99% [70].

A few ammonia plants have been located where a hydrogen off-gas stream is available from a nearby methanol or ethylene operation (e.g., Canadian plants at Kitimat, BC, and Joffre, AB). The capital cost of such a plant is about 50% of the cost of a conventional plant of similar capacity because only the synthesis portion of the plant is required. However, by-product carbon dioxide is not produced, and downstream urea production is therefore not possible [71].

Electrolysis. Electrolytic hydrogen production yields the highest-purity hydrogen (up to 99.999%), benefits from widespread raw materials availability (electricity and water), boasts simple system architecture, and can be scaled economically to serve applications ranging from the smallest hydrogen uses to the larger-volume uses. The main disadvantage of water electrolysis is that electricity is an expensive "fuel." As a result, the technology generally is practical only for systems with hydrogen requirements of 2,000 SCF/h and less [59].

By 2002, several manufacturers had introduced advanced water electrolysis systems that are standardized, compact in size, need minimal operator intervention, and require little maintenance. New cell designs, materials of construction, standardized designs, and manufacturing techniques have enabled manufacturers to decrease the fixed costs of electrolysis technology. In addition, these new systems operate automatically and require very little maintenance [59, 72].

Developing regions are the largest market for electrolysis systems. Electrolysis currently accounts for a very small portion of the hydrogen generated in developed countries that have a commercial hydrogen infrastructure. However, electrolysis can be economic for small-scale generation in areas with inexpensive electricity, and a few plants still produce hydrogen for small-volume ammonia production [36]. In 1997, Messer-MG Industries announced two hydrogen gas plants based on water electrolysis in the United States. In Canada, a portion of the capacity installed at HydrogenAl's merchant hydrogen plant in Québec is based on electrolytic cells [39]. Some of the companies that offer electrolysis technology are listed in [39].

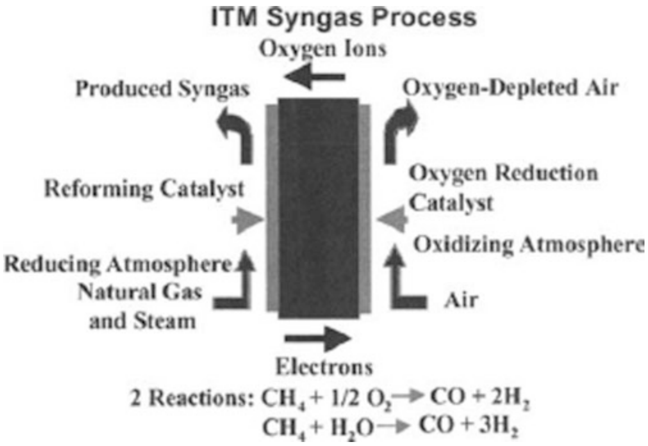


Fig. 22.7 ITM syngas process

Other Processes [3]. The Hypro process makes hydrogen by catalytically decomposing hydrocarbons to carbon and hydrogen. The carbon is burned to provide the heat for the reaction. This process was developed by UOP.

The Lane process makes hydrogen by passing steam over sponge iron at approximately 650°C. The iron is converted to magnetite.

The MRH (*m*ethanol *r*eformer *h*ydrogen) process for generating hydrogen from methanol and separating it by PSA was developed by the Marutani CPE Company.

The Proximol process makes hydrogen by reforming methanol. This technology is offered by Lurgi.

The SBA-HT (*S*ociété *B*elge de l' *A*zote–*H*aldor *T*opsøe) process is a combination of both steam reforming and partial oxidation. The process converts LPG to syngas that is rich in hydrogen. This process was operated in France and Belgium in the 1960s.

New Developments. Several research organizations are developing low-cost methods of hydrogen production. Much of this work is sponsored by the US Department of Energy. Some of the research programs can be found in [73].

The ITM Syngas process involves the direct conversion of methane to synthesis gas (see Fig. 22.7). The process utilizes a mixed, conducting ceramic membrane and partial oxidation to produce the synthesis gas [39]. The goal of this advanced reformer technology is to reduce the cost of hydrogen production by over 25% [73, 74].

Two projects were announced in the late 1990s to develop the ITM process and another related technology. Air Products will lead an 8-year, US$ 90 million research project that is supposed to culminate in the construction in 2005 of a 500 million SCFD ITM process development unit [39]. A separate Oxygen Transport Membrane (OTM) Syngas alliance was formed in 1997 to develop ceramic membrane technology for conversion of natural gas to synthesis gas [75].

Another process is the CO_2-free production of hydrogen via thermocatalytic decomposition of hydrocarbon fuels. The process involves a single-step decomposition of hydrocarbons over carbon catalysts in an air- and water-free environment. Preliminary assessments of the process indicated that hydrogen could be produced at a cost of US$ 5.00 per million BTU (if carbon sold at US$ 100/ton). This production cost is less than that for a steam reforming process coupled with CO_2 sequestration [49].

Some additional new developments in hydrogen production are summarized below.

NGK Insulators of Japan received US Patent 5,741,474 in 1998 for a process for production of high-purity hydrogen. This process is a combination of reforming, partial oxidation, and membrane separation.

Northwest Power Systems obtained US Patent 5,997,594 in 1999 for a steam reformer with internal hydrogen purification.

BP and Kvaerner process are finalizing the demonstration of their compact reformer technology in 2001. This technology involves the integration of combustion, heat transfer, and catalytic reaction within a simple tubular module [76].

Battelle Pacific Northwest National Laboratories are developing microreactors that produce synthesis gas. These reactors can be mass-produced to yield efficient, compact, and cost-effective systems [77].

A related effort is FORSiM (*f*ast *o*xidation *r*eaction in *S*i-technology-based *m*icroreactors) which is funded by the Dutch Technology Foundation and is a cooperative venture between the University of Twente and the Technical University of Eindhoven. The objective of this work is to build and operate the first microreactor for catalytic partial oxidation for small-scale and on-demand hydrogen production [77].

Initial Purification of Synthesis Gas

Synthesis gas from the gasification of fossil fuels is contaminated by several gaseous compounds that would affect its further use in different ways. The purification steps depend on the type of syngas process and the feedstock. Table 22.9 illustrates how the synthesis gas stream can vary with different feedstocks. Figure 22.8 illustrates the different process steps that may occur before ammonia synthesis takes place [78].

If sulfur is present as H_2S or COS, it is a poison for many catalysts and will partly or completely inhibit the catalyst activity [44]. Carbon monoxide (CO) and carbon dioxide (CO_2) can poison the ammonia synthesis catalyst, so both of these compounds must be removed [38].

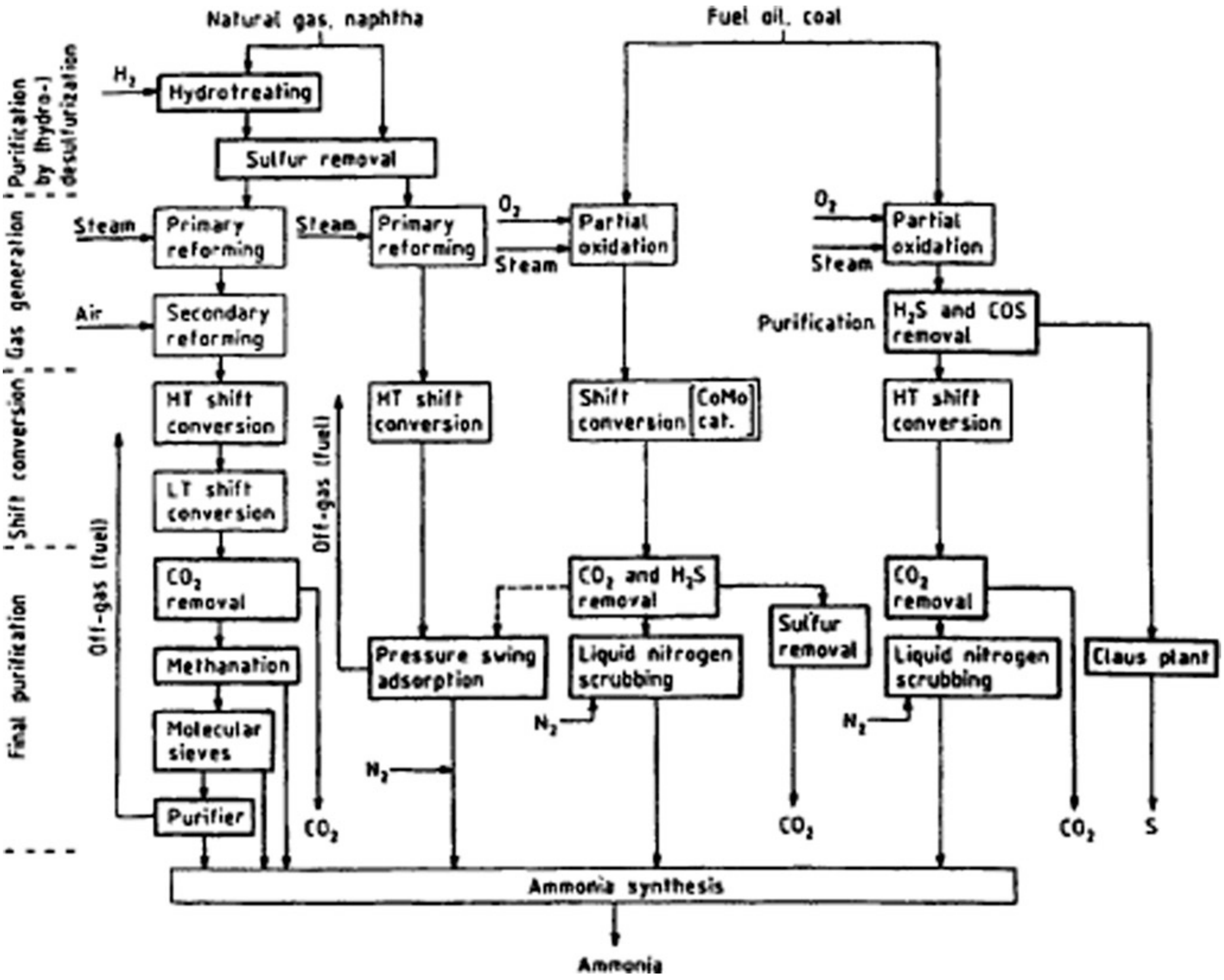


Fig. 22.8 Alternative process steps for generation and purification of synthesis gas (Courtesy of Wiley-VCH. Bakemeier H, Huberich T et al (1985) Ammonia. In: Ullmann's encyclopedia of industrial chemistry, 5th edn, vol A2. VCH Verlagsgesellschaft, Weinheim, pp 143–242)

Table 22.10 Hydrogen recovery technology characteristics [39, 184]

Characteristic	Membrane	Absorption	Cryogenics
Hydrogen purity, %	<95	99.9+	95–99
Hydrogen recovery, %	<90	75–90	90–98
Hydrogen product pressure	<Feed pressure	Feed pressure	Variable
By-products available	No	No	Yes
Feed pressure, psig	250–1,800		250–500

After the sulfur is removed, most traditional ammonia processes have employed the purification steps that are shown in Figs. 22.4 and 22.6. However, ammonia plants have been built that use hydrogen purification via PSA, membrane separation with polymeric membranes, and cryogenic separation. PSA achieves the greatest product purities (about 99.999% pure hydrogen), but it is capital-intensive. Polymeric membranes are the least capital-intensive, but they achieve the lowest purity (about 96–98%). In 1998, PSA was considered the standard process for high-purity hydrogen recovery from raw synthesis gas. It has replaced the traditional steps of the shift conversion of carbon monoxide followed by carbon dioxide removal by absorption and then final purification by methanation. Characteristics of these three purification technologies are outlined in Table 22.10 [39].

Purification with PSA and Polymeric Membranes. The PSA process is based on the selective adsorption of gaseous compounds on a fixed bed of solid adsorbent in a series of identical adsorption beds. The adsorbent is an active carbon or a carbon molecular sieve. Each bed undergoes a repetitive cycle of adsorption and regeneration steps [79, 80]. PSA provides the hydrogen at about the same pressure as the feed, but recoveries are typically lower than from other technologies [51].

The first purification plant that used PSA was developed by Union Carbide Corporation (UCC) and was built at the Yokkaichi Plant of Mitsubishi Petrochemical Industries in 1971 [5]. The process is now licensed by UOP, and more than 400 units were operating worldwide in 1992 [3].

The HYSEC process was developed by Mitsubishi Kakoki K. and Kansai Coke & Chemicals. It has basically the same PSA unit as the UCC process. After the main PSA beds, trace amounts of remaining oxygen are removed by a Deoxo catalytic converter followed by a zeolitic dehumidifier [5].

The LO-FIN (*L*ast *O*ut–*F*irst *I*n) includes a unique gas-retaining vessel that preserves the concentration gradient in one stream before using it to repressurize another bed. This process was jointly developed by Toyo Engineering and Essex Corporation [5].

The Sumitomo-BF PSA process uses CMS as the selective adsorbent [5].

Monsanto and Ube (Japan) developed membrane processes for purification of hydrogen gas mixtures. This process is based on the selective diffusion of hydrogen through semipermeable membranes in the form of hollow fibers. The Monsanto PRISM® separator process (owned by Air Products as of 2004) uses a polysulfone fiber, whereas Ube uses an aromatic polyimide fiber [44].

Conventional polymeric hydrogen separation membranes yield hydrogen at low pressure. Air Products has demonstrated a carbon membrane on an alumina support that removes hydrocarbons from hydrogen/hydrocarbon mixtures and leaves the hydrogen at high pressure [39].

For ultra-high-purity hydrogen (up to 99.9999% pure), high-pressure gas passes over heated palladium membranes. Johnson Matthey is believed to be the leading supplier of these membranes. In 1996, Generex (Tryon, NC) licensed newly developed palladium membrane technology from Los Alamos National Laboratories. In 1998, Wah Chang, an Allegheny Teledyne Company, purchased all patents and rights to the hydrogen metal membrane technology that Bend Research developed [39].

US Patent 6,183,542 was issued in 2001 for a palladium membrane process. This process provides an apparatus that can handle high flow rates of gas while using a minimal amount of hydrogen-permeable material.

Proton Energy Systems (Rocky Hill, CT) received US Patent 6,168,705 in 2001 for an electrochemical gas purifier system that can purify and simultaneously compress hydrogen from a dirty gas stream without relying on moving parts. Final hydrogen gas pressures can exceed 2,000 psig.

The companies that supply PSA and membrane separation systems are listed in [39].

Carbon Monoxide Shift. The water–gas shift conversion or the carbon monoxide shift reaction is one of the traditional purification steps that are found in many ammonia plants. The CO must be removed because it acts as a poison to the catalyst that is used in ammonia synthesis.

The carbon monoxide shift removes most of the CO from the synthesis gas and also produces more hydrogen:

$$\mathrm{CO} + \mathrm{H_2O} \rightleftarrows \mathrm{H_2} + \mathrm{CO_2} \quad \Delta\mathrm{H} = -10\,\mathrm{kcal/mol\ or\ 42\,kJ/mol}$$
$$\text{Homogenous water gas reaction (water gas shift)} \tag{22.2a}$$

The "shift" from CO to CO_2 occurs in two steps. In the high-temperature shift (HTS) conversion, the synthesis gas is passed through a bed of iron oxide/chromium oxide catalyst at around 400°C. The CO content of the gas is reduced to about 3% (on a dry-gas basis). Additional details about the HTS step are given in [38, 45].

The gas from the HTS is cooled to increase the conversion, and then it is passed through the low-temperature shift (LTS) converter. The LTS converter is filled with copper oxide/zinc oxide–based catalyst and operates at about 200–220°C. The residual CO content is about 0.2–0.4% (on a dry-gas basis) [38].

In some plants, the two shift reactions are combined in a medium-temperature shift conversion. When the feed gas is not desulfurized, the CO conversion is called sour–gas shift, and a sulfur-resistant catalyst is used [81–86].

Removal of Sulfur Compounds and Carbon Dioxide. If sulfur is present as H_2S or COS or if CO_2 is present, any of these compounds will be a poison for many catalysts and will partly or completely inhibit catalyst activity.

The point at which sulfur removal is employed depends on the synthesis gas process that is used. Table 22.11 lists many of the processes that are available [44].

The Amine Guard is a corrosion inhibitor that was developed by Union Carbide. It is added to the monoethanolamine (MEA) solvent and allows MEA concentration to be increased. The higher MEA concentration leads to a lower circulation flow rate and a lower energy demand during regeneration [78].

As shown in Figs. 22.4, 22.6, and 22.8, the CO_2 removal step is normally after the shift conversion step. The process gas from the LTS converter contains mainly hydrogen, nitrogen, CO_2, and excess process steam. The gas is cooled, and most of the excess steam is condensed before it enters the CO_2 removal system. This condensate normally contains

Table 22.11 Processes for removal of carbon dioxide and sulfur compounds from synthesis gas

Process name	Key chemical(s)
The Alkazid process	Alkali salts of amino acids (*N*-methylaminopropionic acid)
The aMDEA process	Activated methyl diethanolamine
The Benfield process	Hot potassium carbonate
The Carsol process	Potassium carbonate
The Catacarb process	Hot potassium carbonate
The Fluor Solvent process	Propylene carbonate
The Giammarco-Vetrocoke process	Potassium carbonate with arsenite
The Hi Pure process	Two-stage hot potassium carbonate washing
The Purisol process from Lurgi	*N*-Methyl-2-pyrrolidone (NMP)
The Rectisol process	Low temperature methanol
The Selexol process	Dimethyl ethers of polyethylene glycol
The Shell Sulfinol process	Di-isopropanolamine dissolved in sulfolane and water
Pressurized washing	Monoethanolamine (MEA) or diglycolamine

1,500–2,000 ppm of ammonia and 800–1,200 ppm of methanol. Therefore, it should be stripped or recycled.

The heat that is released during the cooling/condensation can be used to regenerate CO_2 scrubbing solution, to drive an absorption refrigeration unit, or to preheat boiler feedwater. The amount of heat released depends on the process steam-to-carbon ratio (see Tables 22.8 and 22.9). If all this low-level heat is used for CO_2 removal or absorption refrigeration, then high-level heat has to be used for the feedwater system. An energy-efficient process should therefore have a CO_2 removal system with a low-heat demand [38].

The CO_2 is removed in either a chemical, a hybrid, or a physical absorption process. Residual CO_2 contents are usually in the range of 50–1,000 ppmv, depending on the type and design of the removal unit. The physical absorption processes may be designed for zero heat consumption. But for comparison with the chemical processes, the mechanical energy requirements have to be considered [38].

Alkazid Process

The Alkazid process removes sulfur compounds from gas streams. All the sulfur compounds are first catalytically hydrogenated to hydrogen sulfide (H_2S) using a cobalt/molybdena catalyst. The H_2S is then absorbed in an aqueous solution of a potassium salt of either methylamino propionic acid ("Alkazid M") or dimethylamino acetic acid ("Alkazid DIK"). This solution is heated to regenerate the hydrogen sulfide as a concentrate. This concentrate is then treated by the Claus process to recover the sulfur [3].

aMDEA Process

The aMDEA (*a*ctivated *m*ethyl*d*iethanolamine) process removes CO_2, H_2S, and trace sulfur compounds from natural gas and syngas via a pressurized wash with activated diethanolamine. This process was developed by BASF, and in 2002, it was used in more than 140 plants [3].

Benfield Process

The Benfield (*B*enson and *F*ield) process removes carbon dioxide, H_2S, and other acid gases from industrial gas streams by scrubbing with hot aqueous potassium carbonate that contains activators. The chemical reactions are [3, 44]

$$K_2CO_3 + CO_2 + H_2O \rightleftarrows KHCO_3 \quad (22.13)$$

$$K_2CO_3 + H_2S \rightleftarrows KHS + KHCO_3 \quad (22.14)$$

Process details are given in [48, 87]. The process was invented in 1952 and is now licensed by UOP. More than 700 plants were operating in 2000 [3].

Catacarb Process

The Catacarb (*cata*lyzed removal of *carb*on dioxide) process removes carbon dioxide and hydrogen sulfide from gas streams by adsorption in a hot potassium carbonate solution that contains a proprietary catalyst. The process was developed and licensed by Eickmeyer and Associates based on work at the US Bureau of Mines in the 1950s. More than 100 plants were operating in 1997 [3].

Fluor Solvent Process

The Fluor Solvent process removes CO_2 from natural gas and various industrial gas streams by dissolving the CO_2 in a propylene carbonate solvent [3]. Additional details are given in [88, 89]. The process was invented in 1958 by the Fluor Corporation and can be licensed from Fluor Daniel. By 1985, 13 plants were operating [3].

Giammarco-Vetrocoke Process

Two processes are known by the Giammarco-Vetrocoke name. Both processes use an aqueous solution of sodium or potassium carbonate and arsenite to absorb acid gases. Some variations of the process use glycine instead of arsenite to activate the potassium carbonate solution. In one process, the solution is used to extract carbon dioxide from natural gas or synthesis gas. In the other, hydrogen sulfide is extracted from coke-oven or synthesis gas and yields elemental sulfur from a complex sequence of reactions. In 1992, more than 200 plants were operating [3, 90].

Hi Pure Process

The Hi Pure process is a variation of the Benfield process. It uses two stages of scrubbing by hot potassium carbonate solution to reduce the CO_2 content of gases to very low levels [3].

Purisol Process

The Purisol process removes H_2S from gases by selective absorption in *N*-methyl-2-pyrrolidone (NMP). It was developed and licensed by Lurgi, particularly for desulfurizing waste gases from IGCC (*i*ntegrated *g*asification *c*ombined *c*ycle) coal gasification plants. However, the technology might be applied to some synthesis gas processes. In 1996, seven Purisol units were either in operation or under construction [3].

Rectisol Process

The Rectisol process was originally developed to remove sulfur and acid gas compounds from gas mixtures that are produced from the partial oxidation of hydrocarbons. It is based on pressurized washing with low-temperature methanol, which results in the physical absorption of the sulfur compounds in the methanol. The process was originally developed in 1951 by Lurgi for the SASOL coal gasification plant in South Africa. The process was further developed by Linde and is now offered for licensing by both companies for the removal of CO_2, H_2S, HCN, C_6H_6 (benzene), and gum-forming hydrocarbons from syngas and fuel gas. In 2003, over 70 units were in operation or under construction [3].

Selexol Process

The Selexol process removes acid gases from hydrocarbon gas streams by selective absorption in polyethylene glycol dimethyl ether (DMPEG). It absorbs H_2S, CO_2, COS, and mercaptans. The process has been used to remove carbon dioxide from syngas, natural gas, and coal gas. The process was developed by Allied Chemical in the 1960s and is now offered for license by UOP. Over 55 units were operating in 2002 [3].

Shell Sulfinol Process

The Shell Sulfinol process removes H_2S, CO_2, COS, and organic sulfur compounds from natural gas by scrubbing with di-isopropanolamine dissolved in a mixture of sulfolane ($C_4H_8SO_2$) and water. It was developed in the 1960s by Shell. In 1997, over 200 commercial units were operating or under construction [3].

Pressure Washing with Monoethanolamine

In this process, a 15–30% solution of MEA in water is used to absorb the CO_2 under pressure. The solution is then regenerated by heating it in a stripper to release the CO_2. This process is characterized by good CO_2 absorption properties at low pressure. However, high regeneration energy consumption limits its use.

If the Amine Guard corrosion inhibitor is used, the MEA concentration in the circulating solution can be increased to 30% from a normal 20%. Hence, the circulating rate can be decreased by 33%, and the heat requirements are decreased by 43% [91]. The Amine Guard technology is licensed by UPO [92].

Retrofits of CO_2 Removal System

In a CO_2 removal system that uses wet scrubbing, the existing towers are the major limit to more capacity because they are expensive to replace. In an amine system, absorption increases as amine concentration increases. But a higher amine concentration requires (1) more filtration to clean the solution and (2) the addition of corrosion inhibitors. Another option is to change from MEA to methyl diethanolamine (MDEA) [56].

In a potassium carbonate system, different additives can be used to increase the CO_2 absorption rate. In any wet scrubbing system, a change from random to structured packing can lead to higher solvent circulation rates and improved mass transfer [56].

Final Purification of Synthesis Gas

Before the synthesis gas enters the ammonia synthesis loop, essentially all of the oxygen compounds must be completely removed to (1) avoid poisoning the ammonia synthesis catalyst and (2) keep CO_2 from forming carbamates and ammonium carbonate in the synthesis loop. It is also advantageous to remove the inert gases (methane, argon, etc.) to achieve a higher synthesis conversion per pass [78].

Methanation. In a steam reforming process that includes CO shift conversion and CO_2 removal, the synthesis gas still contains 0.1–0.2 mol.% CO and 100–1,000 ppmv of CO_2. The following reactions are the simplest method for eliminating these small concentrations of oxygen compounds:

$$\begin{aligned} &CO + 3H_2 \rightleftarrows H_2O + CH_4 \\ &\quad \Delta H = -49.27\,\text{kcal/mol or } 206\,\text{kJ/mol} \end{aligned} \tag{22.11}$$

$$\begin{aligned} &CO_2 + 4H_2 \rightleftarrows 2H_2O + CH_4 \\ &\quad \Delta H = -39.44\,\text{kcal/mol or } 165\,\text{kJ/mol} \\ &\quad \text{Methanation} \end{aligned} \tag{22.12}$$

The normal methanation operating temperature is 250–300°C, and a large excess of hydrogen is present. The equilibrium lies far to the right side of the above reactions, so the CO and CO_2 impurities can be reduced to about 5 ppm [45]. Additional catalyst and process details are given in [55, 78, 93, 94].

Cryogenic Purification. In the cryogenic purifier, all the methane and the excess nitrogen are removed from the synthesis gas as well as a part of the argon. The cooling is produced by depressurization, and no external refrigeration is needed. The purified syngas is then practically free of all impurities, except for a small amount of argon. The cryogenic unit also receives the purge from the ammonia synthesis section and delivers an off-gas for fuel. The combination of higher conversion per pass and reduced purge flow results in a more efficient process [38]. The KBR Purifier is an example of this process [95]. Fifteen KBR Purifier plants have been built since 1966. These plants range in capacity from 680 to 1,750 tons/day, and as of 2001, all 15 plants are still operating [95].

Dehydration. If the makeup gas to the ammonia synthesis loop is absolutely free of catalyst poisons, such as H_2O and CO_2, it can flow directly to the ammonia synthesis converter. This leads to the most favorable arrangement from a minimum energy point of view. This can be accomplished by allowing the gas that leaves the methanation step to pass through beds of molecular sieves to remove water and CO_2 [78].

Liquid Nitrogen Wash. In many partial oxidation syngas processes, liquid nitrogen scrubbing is used to remove the carbon monoxide that remains after the shift conversion step. The CO content may be as high as 3–5% in plants that have only an HTS conversion. The liquid nitrogen wash (1) delivers a gas to the ammonia synthesis loop that is free of all impurities (including inert gases) and (2) adds all or part of the nitrogen that is required for ammonia synthesis. The nitrogen is obtained from the air-separation plant that provides the oxygen for the partial oxidation process (see Fig. 22.6) [78].

Adjust Hydrogen-to-Nitrogen Ratio. The optimum ammonia synthesis reaction rate depends on several factors including pressure, temperature, H_2-to-N_2 molar ratio, and catalyst activity. Therefore, the H_2-to-N_2 molar ratio is adjusted to suit the requirements in ammonia synthesis. This adjustment occurs before the compression step.

Compression

Ammonia synthesis is normally carried out at a pressure that is higher than that for synthesis gas preparation. Therefore, the purified synthesis gas to the ammonia synthesis loop must be compressed to a higher pressure [78].

Due to several major developments in ammonia process technology, ammonia plants with 1,000–1,500 tons/day capacities have become the industry standard for new plant construction. In 2001, plants as large as 2,000 tons/day have become common. These plants have much lower production costs than the earlier generation of smaller plants mainly because steam-driven, centrifugal compressors are used rather than electrically driven, reciprocating compressors [36, 78].

Ammonia Synthesis

The ammonia synthesis reaction is [78]

$$\begin{aligned} &N_2 + 3H_2 \rightleftarrows 2NH_3 \\ &\quad \Delta H_{298} = -45.72\,\text{kJ/mol or } 11.04\,\text{kcal/mol} \end{aligned} \tag{22.15}$$

The reaction normally takes place on an iron catalyst. The reaction pressure is in the range of 100–250 bar, and temperatures are in the range of 350–550°C. At the usual commercial converter operating conditions, the conversion achieved per pass is 20–30% [38]. In most commercial ammonia plants, the Haber recycle loop process is still used to give substantially complete conversion of the synthesis gas. In this process, the ammonia is separated from the recycle gas by cooling and condensation. Next, the unconverted synthesis gas is supplemented with fresh makeup gas and returned as feed to the ammonia synthesis converter [78].

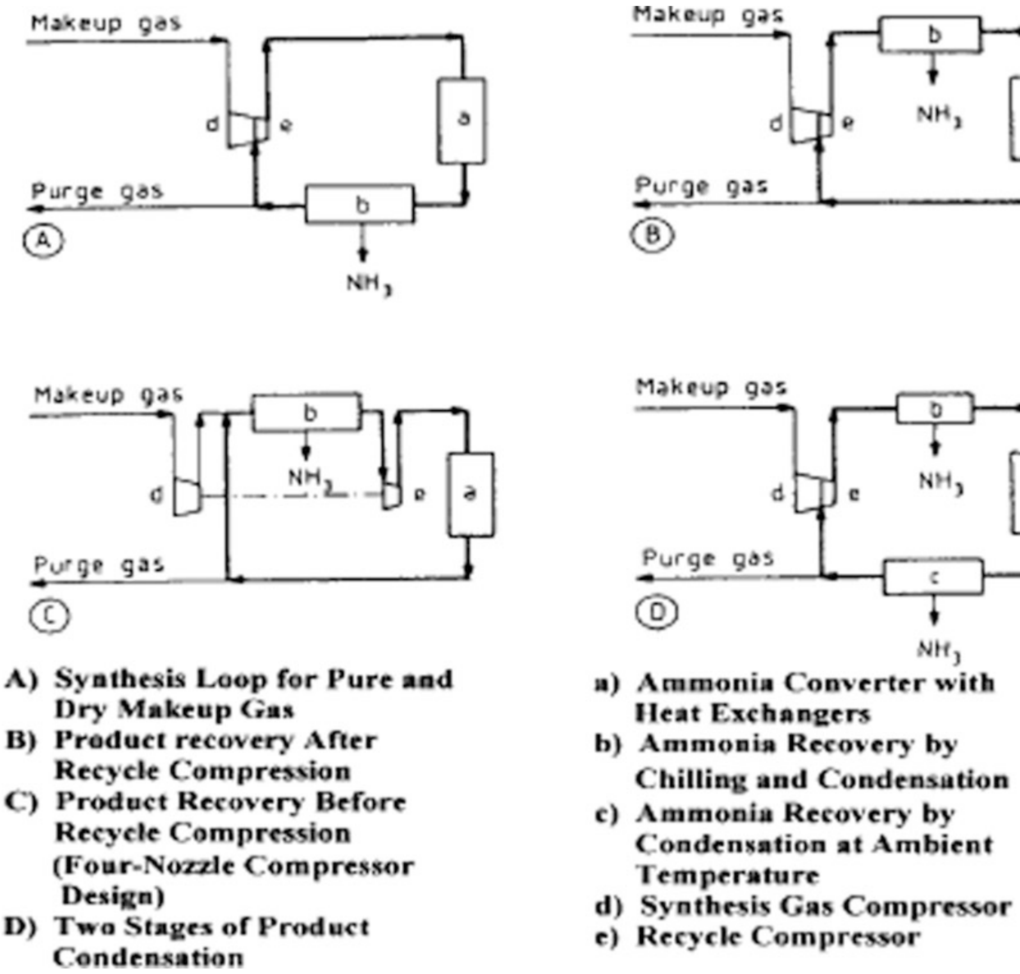


Fig. 22.9 Schematic flow diagrams of typical ammonia synthesis loops [78] (Courtesy of Wiley-VCH. Bakemeier H, Huberich T et al (1985) Ammonia. In: Ullmann's encyclopedia of industrial chemistry, 5th edn, vol A2. VCH Verlagsgesellschaft, Weinheim, pp 143–242)

Synthesis loop arrangements differ with respect to (1) the points in the loop at which the makeup gas is delivered, (2) where the ammonia is taken out, and (3) where the purge gas is taken out [38] (see Fig. 22.9). The best arrangement is shown in Fig. 22.9a. After the gas leaves the ammonia synthesis converter, ammonia is condensed/removed by cooling and the recycle gas is returned to the recycle compressor. This represents the most favorable arrangement from a minimum energy point of view. It results in the lowest ammonia content at the entrance to the converter and the highest ammonia concentration for condensation [78]. The advantages and disadvantages of the other arrangements in Fig. 22.9 are described in [78].

Conventional reforming with methanation as the final purification step produces a synthesis gas that contains inerts (CH_4 and argon) in quantities that do not dissolve in the condensed ammonia. Most of the inerts are removed by taking a purge stream out of the synthesis loop. The size of this purge stream controls the level of inerts in the loop at about 10–15%. The purge gas is scrubbed with water to remove ammonia, and then it can be used as fuel or sent to hydrogen recovery. The best point at which to take the loop purge is discussed in [38].

Reaction Rate

Knowledge of the macrokinetics is important for solving the industrial problem of designing ammonia synthesis reactors, for determining the optimal operating conditions, and for computer control of ammonia plants. Some of the considerations are as follows: high pressure promotes a high rate of ammonia formation, high ammonia concentration in the synthesis gas (recycle gas) restricts ammonia formation (see Fig. 22.10), the rate of formation initially increases with rising temperature but then goes through a maximum as the system approaches thermodynamic equilibrium (see Fig. 22.11), and with lower temperatures, the maximum rate shifts to a lower hydrogen–nitrogen ratio (see Fig. 22.12) [78].

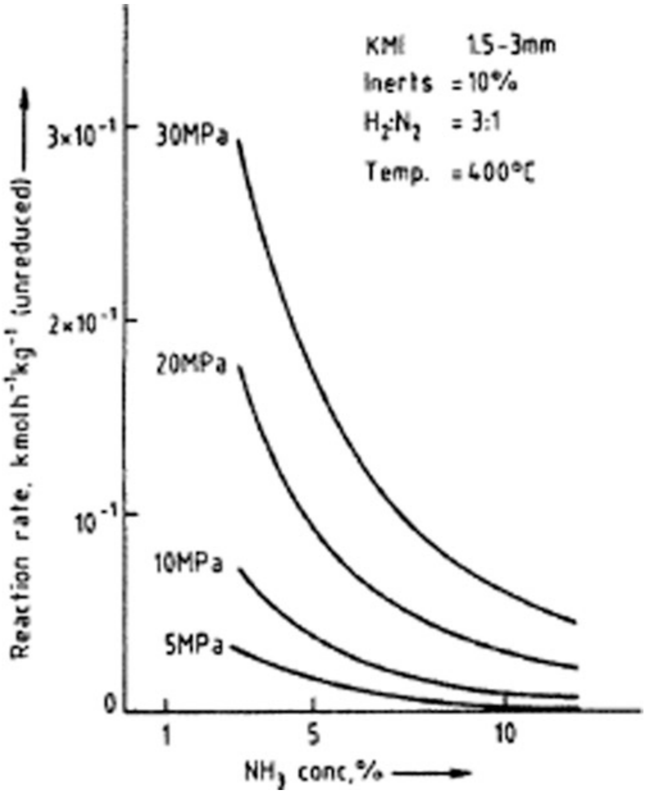


Fig. 22.10 Reaction rate for NH_3 synthesis. Dependence on the ammonia concentration at various pressures (Courtesy of Wiley-VCH. Bakemeier H, Huberich T et al (1985) Ammonia. In: Ullmann's encyclopedia of industrial chemistry, 5th edn, vol A2. VCH Verlagsgesellschaft, Weinheim, pp 143–242)

In 2001, Hyprotech and Synetix announced an ammonia plant simulation that can be used for modeling, online monitoring, and optimization of the plant. The simulation includes Synetix reactor models, customized thermodynamic data, and information to simulate the performance of a range of catalysts. The reactor models in the simulation include primary and secondary reformers, high-temperature shift converter, low-temperature shift converter, methanator, and ammonia synthesis converter [96].

Catalysts

When ammonia is made from natural gas using the steam reforming process, several reaction stages are needed and catalysts are key to the economic operation of each stage. The chemistry of the process and the basic materials that make up the catalysts are shown in Fig. 22.13 [45].

For a given operating pressure and a desired production rate, the catalyst determines (1) the operating temperature range, (2) recycle gas flow, and (3) refrigeration requirements. It also indirectly influences the makeup gas purity requirements [78].

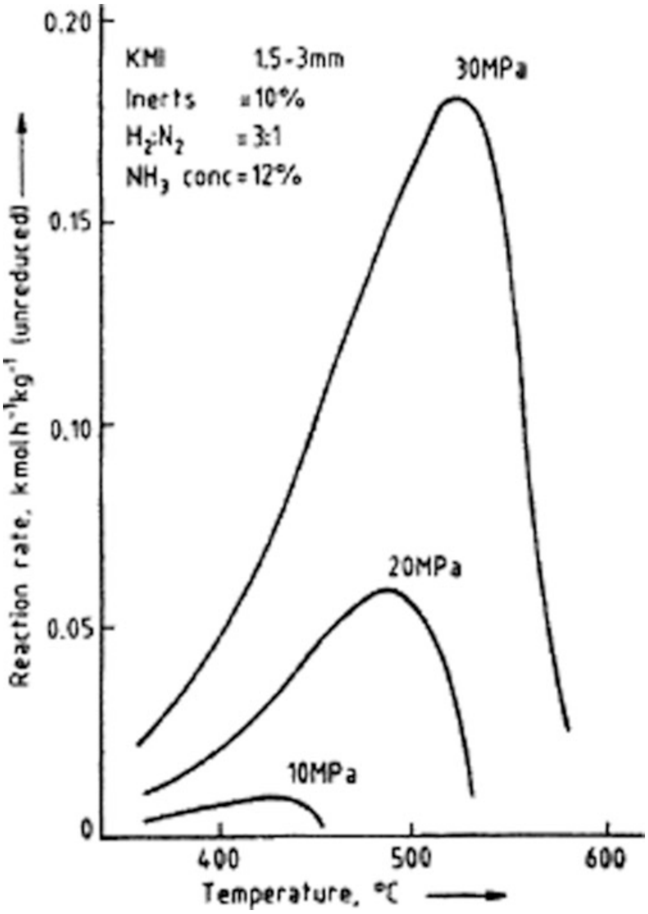


Fig. 22.11 Reaction rate for NH_3 synthesis. Dependence on the temperature at various pressures (Courtesy of Wiley-VCH. Bakemeier H, Huberich T et al (1985) Ammonia. In: Ullmann's encyclopedia of industrial chemistry, 5th edn, vol A2. VCH Verlagsgesellschaft, Weinheim, pp 143–242)

Industrial catalysts for ammonia synthesis must satisfy the following requirements: (1) high catalyst activity at the lowest possible reaction temperatures, (2) the highest possible insensitivity to oxygen- and chlorine-containing catalyst poisons, (3) long life, and (4) mechanical strength.

The choice of particle size and shape of commercial ammonia catalysts is determined mainly by the catalyst performance and the pressure drop. From the standpoint of space–time yield, it is desirable to use the finest possible particle, which is about 1–2 mm [78].

For processes operating at pressures of 25–45 MPa (250–450 bar) and space velocities of 8,000–20,000 m^3/m^3 h (STP), a grain size of 6–10 mm is preferred. (Space velocity is defined as the number of reactor volumes of feed—at specified conditions—that can be treated in unit time.) Larger granulations (8–15 or 14–20 mm) are used only in plants where the lowest possible pressure drop is essential because of very high gas velocities. In catalyst zones in which the ammonia formation rate is so high that

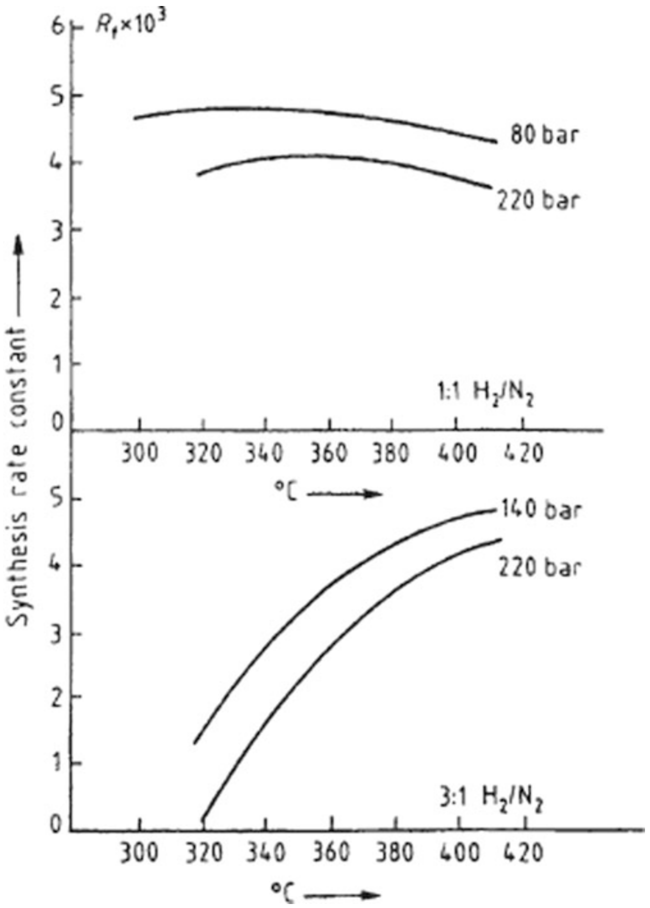


Fig. 22.12 Ammonia synthesis rate constant dependence on hydrogen–nitrogen ratio (Courtesy of Wiley-VCH. Bakemeier H, Huberich T et al (1985) Ammonia. In: Ullmann's encyclopedia of industrial chemistry, 5th edn, vol A2. VCH Verlagsgesellschaft, Weinheim, pp 143–242)

the allowable temperature limits are exceeded, it may be advantageous as well to use coarse particles for suppressing the reaction [78]. Some of the poisons of ammonia catalysts are discussed in [78].

The activity of damaged catalysts may be almost completely restored by reduction with clean synthesis gas at a relatively low temperature. It has also been found that the degree of poisoning rises as the partial pressure ratio, p_{H_2O}/p_{H_2}, increases and the degree of poisoning falls with increasing temperature [78].

Energy Efficiency

Due to increased feedstock costs, some new ammonia plant designs use fuel more efficiently, but their capital cost may be higher. The recovery of hydrogen and ammonia from the synthesis purge gas by a cryogenic unit or a membrane system results in an ammonia capacity increase of about 5% [36].

Substantial improvements have been made in the energy efficiency of CO_2 removal systems. The first large-scale ammonia plants in the 1960s typically used MEA as a solvent. Energy input was over 50,000 kcal/kg mol of CO_2 removed. In 2001, plants use improved solvents and designs that can reduce the energy input to about 10,000 kcal/kg mol of CO_2 removed [36].

A typical world-scale plant that was built in the 1970s consumed about 42 billion BTU of natural gas per ton of ammonia produced. Retrofitting such a plant to improve fuel efficiency can reduce gas consumption to about 36 million BTU/ton. Ammonia plants that were built in the late 1990s use only about 30 million BTU/ton of ammonia, are easier to

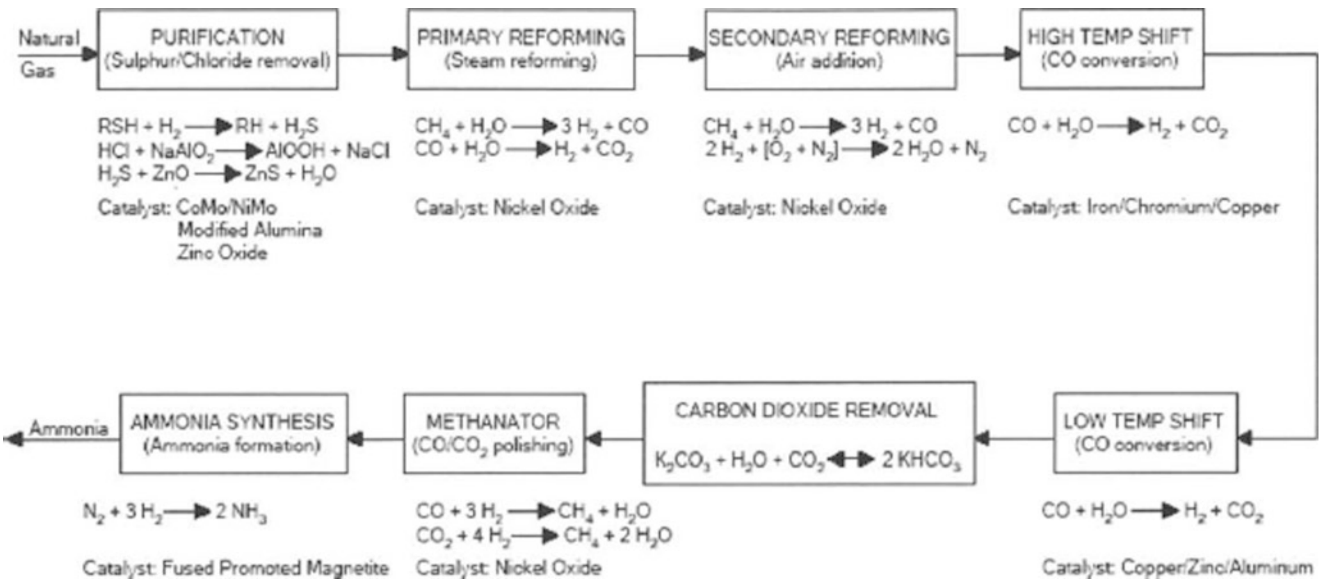


Fig. 22.13 Basic catalyses materials and chemistry of the ammonia process (Copyright Johnson Matthey plc. Reproduced by permission of Johnson Matthey plc)

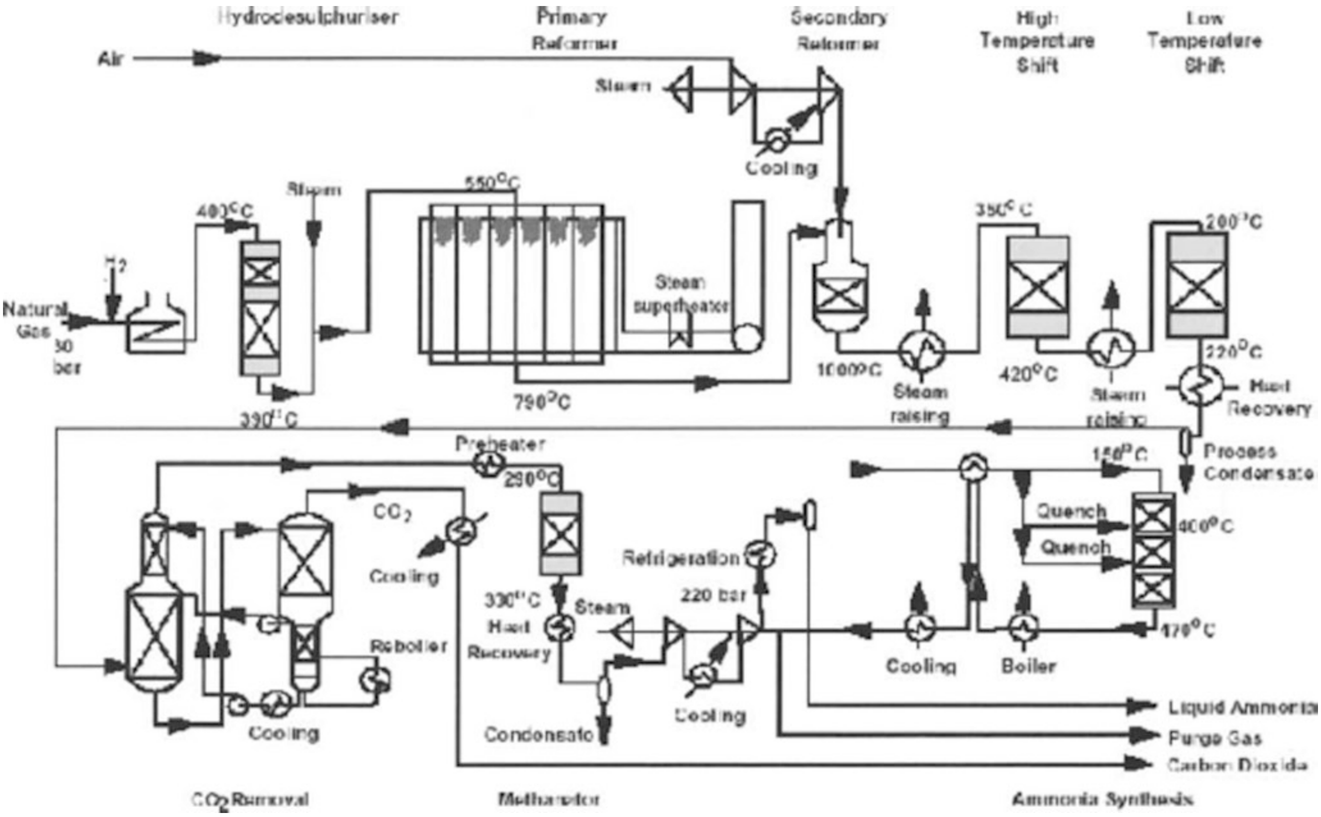


Fig. 22.14 Simplified ammonia flowsheet (Copyright Johnson Mattey plc. Reproduced by permission of Johnson Mathey plc)

operate, and have slightly lower conversion costs. Some new plants also recover more than 1 million BTU/ton by generating electricity from waste heat [36].

Ammonia Plant Design

A simplified flowsheet for an ammonia plant that processes natural gas via steam reforming is shown in Fig. 22.14. A block diagram of this same plant is shown in Fig. 22.15. This diagram lists typical stream compositions, typical operating conditions, catalyst types (recommended by Synetix), and catalyst volumes.

The KAAP*plus*™ (*K*ellogg *B*rown & *R*oot *A*dvanced *A*mmonia *P*rocess *plus*) is an ammonia process design made up of commercially proven technologies: KBR Reforming Exchanger System (KRES), KBR Purifier, and KBR Advanced Ammonia Process (KAAP) for synthesis. This process uses a ruthenium-based catalyst that is 20 times more active than the iron-based catalysts and improves synthesis efficiency. This is because synthesis pressure is lowered from a typical pressure of 150 bar (2,175 psig) to 90 bar (1,305 psig) [36, 95, 97]. Capital cost savings of 3–8% are achieved with KAAP*plus*™ [95].

Haldor Topsøe's ammonia synthesis technology is based on the S-200 ammonia converter. This is a two-bed radial flow converter with indirect cooling between the beds. This converter concept has been used extensively to upgrade existing converters (Topsøe or other designs) in modification projects to achieve higher capacity (up to 20%) and/or better energy efficiency [98].

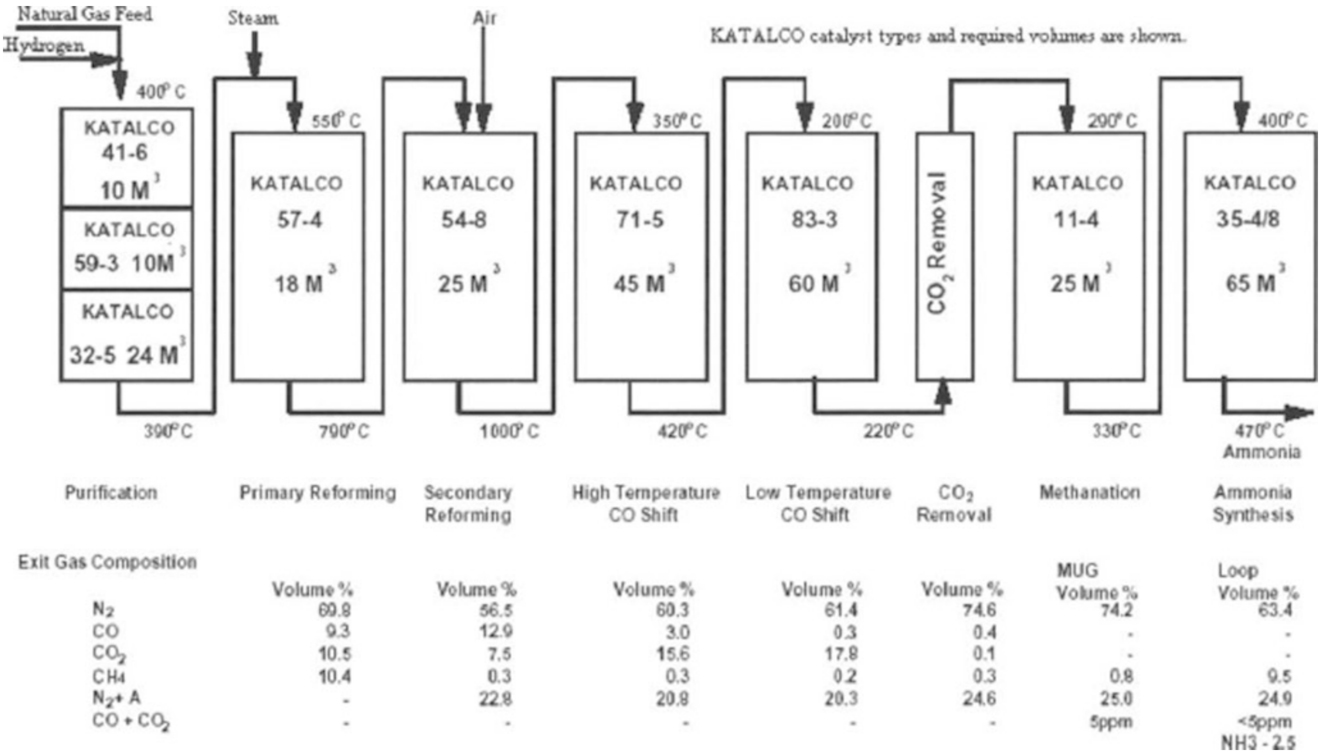
Other companies that offer ammonia process technology are Linde [99, 100], ICI [23, 24], Ammonia Casale [101–104], and Uhde [105].

Ammonia Separation

The removal of ammonia product is accomplished via mechanical refrigeration or absorption/distillation. The choice is made by examining the fixed and operating costs. Typically, refrigeration is more economical at synthesis pressures of 100 atm (1,470 psia) or more. At lower pressures, absorption/distillation is usually favored [106].

Ammonia Synthesis

If the synthesis gas contains traces of carbon oxides, ammonium carbamate will form upon mixing with the ammonia in the recirculating gas from the synthesis loop. The carbamate will clog and/or corrode downstream equipment. To avoid



	Volume %	Volume %	Volume %	Volume %	Volume %	MUG Volume %	Loop Volume %
N_2	69.8	56.5	60.3	61.4	74.6	74.2	63.4
CO	9.3	12.9	3.0	0.3	0.4	-	-
CO_2	10.5	7.5	15.6	17.8	0.1	-	-
CH_4	10.4	0.3	0.3	0.2	0.3	0.8	9.5
N_2+ A	-	22.8	20.8	20.3	24.6	25.0	24.9
CO + CO_2	-	-	-	-	-	5ppm	<5ppm NH3 - 2.5

Fig. 22.15 Block diagram of 1,000 tons/day ammonia plant (Reproduced by permission of Johnson Matthey Catalysts. Copyright Johnson Matthey plc)

this condition, the carbon oxide levels in fresh makeup gas should be less than 5 ppm [48].

Many ammonia synthesis converters are subject to nitriding and hydrogen embrittlement. The internal catalyst baskets are made of SS 321 material [48]. The nitriding effect is more pronounced in low-alloy steels above 450°C. Austinic steels with a high nickel content offer considerably more resistance. Alloys of the Cr–Ni–Mo type are usually used for the gas side. Atomic hydrogen is absorbed in the metal by diffusion. The subsequent reaction with carbon in the steel results in blistering and cracking from decarburization [48].

Large-Capacity Ammonia Plants

In 2002, the design and construction of ammonia plants with capacities of 3,000 tons/day or more were underway. These large-capacity plants will be at least 50% larger than the largest plants that were operating in 2002. The benefit of bigger plants is to realize economy of scale. Doubling the size of a single-train plant has the potential to reduce the capital-related cost of production by about 20% [107].

The companies that are developing large-capacity plants are KBR [107], Haldor Topsøe [107, 108], Uhde [107, 109], plus Lurgi and Ammonia Casale [110].

Ammonia Production Costs

Ammonia production costs are very dependent on feedstock cost, which is normally natural gas. Natural gas prices can vary significantly from region to region and by the cycles in the world energy market. Typical ammonia production costs are listed in [36]. The natural gas cost makes up 70% to over 85% of the cash production costs as the natural gas cost varies from $2.00 per million BTU up to $6.00 per million BTU. In general the ammonia cash production costs in $ per ton are a factor of 40–50 times the natural gas price in US dollars per million BTU.

Uses of Ammonia

Anhydrous ammonia was produced in about 80 countries in 2001. About 85% is used for nitrogen fertilizer production, including about 4% that is directly applied to the fields [32, 111]. In the United States, the distribution of ammonia use differs slightly from the worldwide uses. In the United States, only 80% of the ammonia is used to make fertilizers. Chemical intermediates account for 19% of the ammonia use, and the remaining 1% is used in pulp and paper, metals, and refrigeration applications [34].

Chemical Production and Other Uses

The uses of ammonia are listed below:

- *Acrylonitrile.* ACRN is used in the production of acrylic fibers and various resins. ACRN is produced by the catalytic oxidation of propylene and ammonia. About 0.48 ton of ammonia is needed to make 1 ton of ACRN [36].
- *Caprolactam.* Caprolactam is used principally in the production of nylon-6 fibers. In the production of caprolactam, ammonia is used to maintain a reaction pH of 7 in one step, and in another step, it is used to neutralize sulfuric acid. Both of these steps result in the production of a low-grade ammonium sulfate fertilizer. In the classical process, 5 kg of $(NH_4)_2SO_4$ is produced per kilogram of caprolactam. In 1992, about 33% of the worldwide caprolactam capacity was based on DSM's caprolactam process that only produces 1.8 kg of $(NH_4)_2SO_4$ per kg of caprolactam [44].
- *Miscellaneous Amines and Nitriles.* A small but significant amount of ammonia is consumed in the production of various amines and nitriles [36].
- *Metallurgy.* In this application, ammonia is cracked to produce a "dissociated ammonia" that consists of 75% hydrogen and 25% nitrogen. Dissociated ammonia is used in a number of metal treatment processes [36].
- *Refrigeration.* Anhydrous ammonia is used as a refrigerant in industrial, closed-circuit refrigeration systems. Ammonia's high latent heat, low vapor density, chemical stability, and low iron corrosivity promote its use. Ammonia is used in large industrial systems (rather than halocarbon refrigerants) because ammonia absorbs about six times as much heat per unit weight and energy costs of operation are substantially lower [36].
- *Pulp and Paper.* In the production of ammonia-based sulfite pulp, the amount of ammonia used per ton of pulp varies from mill to mill. The industry average is 125 lb of ammonia per ton of sulfite pulp. Ammonia is also used for stack gas scrubbing and wastewater treatment [36].
- *Phosphate Ore Flotation.* A small amount of ammonia is used as a modifying reagent in the froth flotation of phosphate ores.
- *Household Cleaners.* Household ammonia, liquid window cleaners, liquid all-purpose cleaners, and various other household cleaning products constitute this small market for ammonia. Ammonia is also used to make surfactants that go into liquid dishwashing detergents [36].
- *Air Pollution Control.* A number of processes use ammonia for scrubbing sulfur oxides (SO_x) and/or nitrogen oxides (NO_x) from industrial and electrical power plant stack gases [36].
- *Uranium Concentrate Production.* Ammonia is used to precipitate the final uranium concentrate salt (yellowcake) prior to drying or calcination. It is generally economically preferred over sodium hydroxide or hydrogen peroxide [36].
- *Water Purification.* Ammonia is combined with chlorine to purify some municipal and industrial water supplies [36].
- *Waste Treatment.* Small amounts of ammonia are added when needed as the nitrogen source for the bacteria in industrial and municipal biological waste treatment systems. Other industrial plants use minor amounts of ammonia to neutralize acid in plant wastes [36].
- *Duplicating Processes.* Ammonia is used as a process chemical in photographic copy machines [36].
- *Fabric Treatment.* Permanent-press, 100% cotton fabrics are produced via a licensed process involving mercerization in a bath of liquid anhydrous ammonia. This Sanfor-Set process was commercialized in 1975 [36].
- *Semiconductor Industry.* High-purity ammonia (99.99995%) is used in gallium nitride (GaN) manufacturing processes to provide high brightness blue and white LEDs (light emitting diodes), in high-performance optoelectronics (such as liquid crystal displays and flat panel displays), and in high-power electronic devices (such as lasers and laser diodes).

Distribution and Storage

Ammonia Toxicity

Ammonia is a strong local irritant, and the primary target organ is the pulmonary system. Ammonia or ammonium hydroxide can penetrate the cornea rapidly and lead to damage of the iris, cataract, and glaucoma. Oral ingestion of aqueous ammonia can corrode the mucous membranes of the oral cavity, and constrictions of the esophagus may result [78].

Ammonia is not considered to be carcinogenic nor is it mutagenic. The effects of different ammonia concentrations are summarized in Table 22.12 [78].

Ammonia Quality

Most industrial uses of ammonia require a higher-purity level of product than agricultural uses. Typical ammonia specifications are shown in Table 22.13 [78].

Commercial grade is produced by ammonia synthesis, while refrigeration grade is normally made from industrial raw ammonia by distillation. For ammonia shipped or pipelined in the United States, water content must be at least 0.2 wt.% to inhibit stress corrosion cracking of the carbon steel [36].

Various concentrations and purities of aqueous ammonia are on the market. The typical concentration is 25–30% ammonia, and the iron content is less than 10 ppm. If the ammonia content is above 25%, it must be shipped in a pressure vessel because of its elevated vapor pressure [78].

Distribution and Storage

Ammonia production requires storage facilities to smooth over fluctuations in production, usage, and shipments. Ammonia is stored and distributed to point of use almost exclusively as a liquid [78].

Three methods have been used for storing liquid ammonia [78]:

- Pressure storage at ambient temperature in spherical or cylindrical pressure vessels having capacities up to about 1,500 tons
- Atmospheric storage at −33°C in insulated cylindrical tanks for up to about 50,000 tons per vessel
- Reduced pressure storage at about 0°C, usually in insulated, spherical pressure vessels for quantities up to about 2,500 tons per sphere

Table 22.12 Health effects of different ammonia concentrations

Ammonia concentration (ppm)	Health effect[a]
5	Threshold detection limit
50	Easily perceived
50–72	No significant impairment to respiration
100	Irritation to nose and throat
	Burning sensation in eyes
200	Headache and nausea
250–500	Rapid heart beating
700	Immediate onset of burning sensations in the eyes
1,000	Immediate coughing

[a] The TLV has been set at 25 ppm (17 mg/m^3) as an 8 h time-weighted average (TWA) with a short-term exposure limit (STEL) of 35 ppm (24 mg/m^3)

Ammonia is delivered in small containers, tank trucks, tank cars, barges, and via pipeline. The most common small containers are cylindrical steel bottles and pressurized flasks that contain about 20–200 kg and polyethylene canisters and metal casks [78]. Trucks have ammonia capacities up to 100 m^3, whereas jumbo rail cars hold up to 150 m^3. Liquid ammonia shipments by barge constitute a larger volume than by road or rail. Additional guidelines about ammonia shipments are given in [36, 78].

Several ammonia pipeline systems were built in the United States in the 1960s. Their construction resulted from the construction of large plants near the sources of natural gas and from developing a large market for liquid anhydrous ammonia in direct fertilization. Pipelines are the lowest cost method to move ammonia over the long distances between the producers along the Gulf Coast and the consumers in the upper Midwest. Some pipelines are up to 3,000-km long and up to 10 in. in diameter [78].

An integrated pipeline system has not been developed in Europe. Only a few pipelines with lengths of less than 50 km have been installed. However, in the FSU, a 2,424-km-long pipeline is in use [78].

Ammonia Price

Direct sales by basic producers are generally at wholesale prices. Basic producers also sell to distributors and jobbers for resale to end users buying smaller quantities. Distributors of ammonia to industrial end users have their own storage facilities, packaging plants, and distribution networks. They are able to supply end users that buy in less than truckload quantities (a minimum of 1,000–2,000 lb or 450–910 kg) or in cylinders or bottles [36].

Another important aspect of the ammonia business is seasonality. The major end use for anhydrous ammonia is as a direct application fertilizer. The application season, particularly in Canada and the Northern United States, is limited and significant storage capability is needed. As a result, it is normal to expect a plant to produce a downstream nitrogen product (such as urea) that is easier to store and

Table 22.13 Ammonia quality specifications

	Commercial grade		Refrigeration grade	
	USA	Germany	USA	Germany
Purity, wt.%, minimum	99.5	99.5	99.98	99.98[a]
Water, wt.%, maximum	0.5	0.2	0.015	0.02
Inerts[b], mL/g, maximum	Not specified		0.1	0.08
Oil, ppm by weight	5.0	5.0	3.0	Not specified
Free of H_2S, pyridine, and naphthenes				

[a] Allowable boiling point change on vaporization of 5–97% of the test sample, 0.9°C

[b] The noncondensable gases dissolved in ammonia are H_2, N_2, CH_4, and Ar. Their amounts depend on the methods of synthesis and storage. The inerts amount to about 50 mL/kg for atmospheric storage

handle and does not have such a limited application season. Industrial markets are normally nonseasonal [36].

Long-term fixed-price contracts are not considered workable due to the volatility in the market and the uncertainty regarding future natural gas prices. The normal long-term contract simply guarantees continuity of supply and calls for prices to be negotiated every 3–6 months [36]. Historical prices are listed in [36].

Nitric Acid

Nitric acid is a strongly acidic, corrosive liquid that is produced by the oxidation of ammonia and subsequent reaction of the oxidation products with water. Pure nitric acid is colorless. The characteristic yellow-brown color generally associated with concentrated solutions is due to dissolved nitrogen dioxide [112].

Nitric acid is an active compound, and its salts are found in all fertile soils. The alchemists obtained nitric acid by heating alum and copper sulfate with nitrate in a retort. Owing to its powerful corrosive action, they named it aqua fortis or "strong water" [2].

Being a powerful oxidizing agent, nitric acid reacts violently with many organic materials, and the reactions may be explosive. As a general rule, oxidizing reactions occur primarily with the concentrated acid and favor the formation of nitrogen dioxide (NO_2). The acidic properties tend to dominate the dilute acid, and this results in the preferential formation of nitrogen oxide (NO) [38].

Nitric acid reacts with all metals except the precious metal series and certain alloys. Although chromium, iron, and aluminum readily dissolve in dilute nitric acid, the concentrated acid forms a metal oxide layer that protects the metal from further oxidation [38].

Physical Properties

Pure anhydrous nitric acid (100%) is a colorless liquid that solidifies at −41.6°C to form white crystals. It boils at 84.1°C. When it boils in light, a partial decomposition occurs with the formation of NO_2 via the following reaction:

$$HNO_3 \rightleftarrows H_2O + 2NO_2 + \frac{1}{2}O_2 \qquad (22.15)$$

Anhydrous nitric acid should be stored below 0°C to avoid decomposition. The nitrogen dioxide remains dissolved in the nitric acid and creates a yellow color at room temperature and a red color at higher temperatures. The pure acid tends to give off white fumes when exposed to air; however, acid with dissolved nitrogen dioxide gives off reddish-brown vapors which leads to the common name "red fuming acid" [38].

Nitric acid is miscible with water, and distillation results in an azeotrope with a concentration of 68.4% HNO_3 and a boiling temperature of 121.9°C at atmospheric pressure. Two solid hydrates are known the monohydrate ($HNO_3 \bullet H_2O$) and the trihydrate ($HNO_3 \bullet 3H_2O$) [38].

Nitrogen oxides are soluble in nitric acid, and this property influences all the physical characteristics that depend on the concentration of the oxides. This mainly includes the vapor pressure above the liquid and the boiling temperature as well as the color. "Red fuming nitric acid" is a term that, according to *Sax's Dangerous Properties of Industrial Materials*, applies to nitric acid that contains more than 17% NO_2 and has a density of more than 1.48 g/cm [4, 38]. White fuming nitric acid usually contains 90–99% by weight HNO_3, from 0 to 2% by weight dissolved NO_2, and up to 10% by weight water.

The TLV for nitric acid has been set at 2 ppm as an 8-h time-weighted average (TWA) with a short-term exposure limit (STEL) of 4 ppm.

Some properties of nitric acid are given in Tables 22.14 and 22.15.

Processes

The Birdeland–Eyde or Arc process made nitric acid by passing air through an electric arc, forming nitric oxide, oxidizing NO with air, and absorbing the resulting oxides of nitrogen in water. The reactions are

$$N_2 + O_2 \rightarrow 2NO \qquad (22.16)$$

$$2NO + O_2 \rightarrow 2NO_2 \qquad (22.17)$$

$$2NO_2 + H_2O \rightarrow HNO_2 + HNO_3 \qquad (22.18)$$

$$3HNO_2 \rightarrow HNO_3 + 2NO + H_2O \qquad (22.19)$$

The process was operated in Norway from 1905 to 1930, and it was first used in the United States in 1917 [3]. It used hydroelectric power but was made obsolete by the ammonia oxidation process.

The Ostwald process is the basis for the modern processes that make nitric acid by the catalytic oxidation of ammonia. It was discovered in 1900 by Wilhelm Ostwald, a German physical chemist. The process was used by Germany during World War I to make explosives after the Allied blockade cut off the supply of nitrites from Chile and other places [113].

Some of the processes that have been based on the Ostwald process are

Table 22.14 Physical properties of nitric acid solutions [85]

wt.% HNO$_3$	Density (at 20°C) (g/cm^3)	Freezing point (°C)	Boiling point (°C)	Specific heat (at 20°C) (J/(g K))[a]	Partial press. (at 20°C) (Pa)[b]		Viscosity (@ 20°C) (mPa s (=cP))	Thermal conduct (at 20°C) (W/(m K))[c]
					HNO$_3$	H$_2$O		
0.0	0.99823	0	100.0	4.19	–	2,333	1.0	0.61
10.0	1.0543	−7	101.2	3.73	–	2,266	1.1	0.57
20.0	1.1150	−17	103.4	3.39	–	2,026	1.2	0.54
30.0	1.1800	−36	107.0	3.18	–	1,760	1.4	0.50
40.0	1.2463	−30	112.0	3.01	–	1,440	1.6	0.47
50.0	1.3100	−20	116.4	2.85	27	1,053	1.9	0.43
60.0	1.3667	−22	120.4	2.64	120	653	2.0	0.40
70.0	1.4134	−41	121.6	2.43	387	347	2.0	0.36
80.0	1.4521	−39	116.6	2.22	1,400	120	1.9	0.35
90.0	1.4826	−60	102.0	1.97	3,600	27	1.4	0.31
100.0	1.5129	−42	86.0	1.76	6,000	0	0.9	0.28

[a] To convert J/(g K) to cal/(g°C), divide by 4.184

[b] To convert Pa to atm, divide by 1.013×10^5

[c] To convert W/(m K) to BTU (h ft°F), divide by 1.7307

Table 22.15 Thermodynamic properties of nitric acid and its hydrates [185]

	HNO$_3$	HNO$_3$ • H$_2$O	HNO$_3$ • 3H$_2$O
Nitric acid, wt.%	100.0	77.77	53.83
Freezing point, °C	−41.59	−37.62	−18.47
Heat of formation (at 25°C), kJ/mol[a]	−173.35	−472.07	−888.45
Free energy of formation (at 25°C), kJ/mol[a]	−79.97	−329.29	−810.99
Entropy (at 25°C), kJ/(mol K)[a]	155.71	217.00	347.17
Heat of fusion, kJ/mol[a]	10.48	17.52	29.12
Heat of vaporization (at 20°C), kJ/mol[a]	39.48	–	–

[a] To convert J to cal, divide by 4.184

- The CNA (*c*oncentrated *n*itric *a*cid) process is a general name for processes that make nitric acid more concentrated than the 70% made in conventional processes.
- The CONIA process makes nitric acid simultaneously at two different concentrations.
- The DSN (*d*irect *s*trong *n*itric) process is a general name for processes that concentrate nitric acid from the 50 to 70% range up to the 98% range.
- The HOKO process also makes nitric acid simultaneously at two different concentrations.
- The Hycon process makes nitric acid at a range of concentrations. It was developed in 1968 and patented (US Patent 3,542,510) by the Chemical Construction Corporation.
- The SABAR (*s*trong *a*cid *b*y *a*zeotropic *r*ectification) process makes nitric acid by the atmospheric oxidation of ammonia. Davy McKee developed the process and built plants based on this technology from 1974 to 1986.

Chemistry. In the Ostwald process, ammonia is catalytically oxidized. The key steps in the process are (1) oxidation of ammonia to nitric oxide (NO), (2) oxidation of NO to nitrogen dioxide (NO_2), and (3) absorption of NO_2 in water to produce nitric acid. Overall, the principal reactions may be summarized as follows [112, 114].

Ammonia reacts with air on platinum/rhodium alloy catalysts in the oxidation section of nitric acid plants:

$$\begin{gathered} 4NH_3 + 5O_2 \rightarrow 4NO + 6H_2O \\ \Delta H_{298} = -54\,\text{kcal/mol}\ (-226\,\text{kJ/mol}) \end{gathered} \tag{22.20}$$

Simultaneously N_2O, N_2, and H_2O are formed in side reactions according to (22.21) and (22.22):

$$\begin{gathered} 4NH_3 + 3O_2 \rightarrow 2N_2 + 6H_2O \\ \Delta H_{298} = -75.8\,\text{kcal/mol}\ (-317\,\text{kJ/mol}) \end{gathered} \tag{22.21}$$

$$4NH_3 + 4O_2 \rightarrow 2N_2O + 6H_2O \tag{22.22}$$

Some of the process energy recovery options are discussed in [114]. The yield of nitric oxide in (22.20) depends on pressure and temperature as indicated in Table 22.16 [114].

The space velocity in the ammonia oxidation reactor should be kept high to maximize yield and avoid

Table 22.16 Effect of reaction pressure and temperature on nitric oxide yield

Pressure (bar)	Temperature (°C)	Nitric oxide yield (%)
Below 1.7	810–850	97
1.7–6.5	850–900	96
Above 6.5	900–940	95

Source: Reproduced by permission of European Fertilizer Manufacturers Association

Table 22.17 Effect of reactor residence time on nitric oxide yield

Reactor residence time (s)	Nitric oxide yield, percent NH_3 oxidized
0.28	82.1
0.11	85.7
0.061	90.2
0.023	91.8

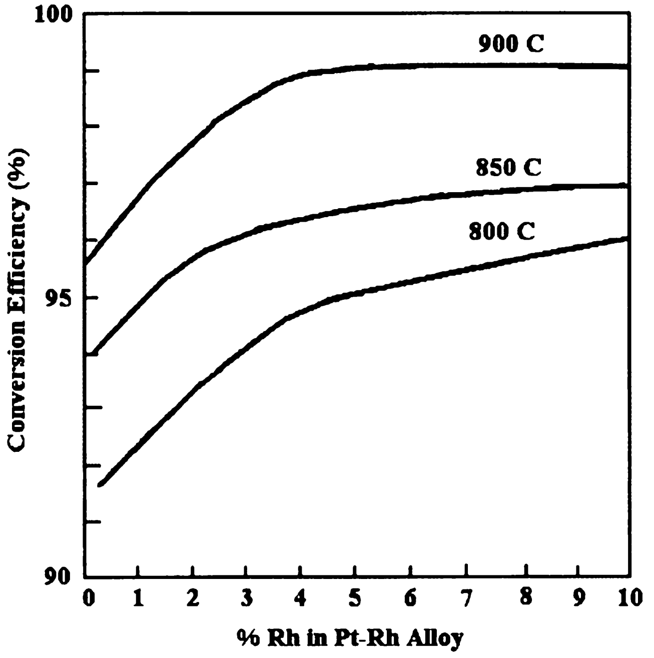


Fig. 22.16 Effect of rhodium content on catalyst efficiency

decomposition of ammonia in the feed gas. Another potential yield loss is shown in (22.23):

$$4NH_3 + 3NO_2 \rightarrow \frac{7}{2}N_2 + 6H_2O \qquad (22.23)$$

The effect of reactor residence time on nitric oxide yield is shown in Table 22.17 [115].

NO is oxidized to NO_2 as the combustion gases are cooled, but some side reactions (such as (22.24) and (22.25)) may also occur.

$$\begin{array}{c} 2NO + O_2 \rightarrow 2NO_2 \\ \Delta H_{298} = -13.6\,\text{kcal/mol}\ (-57\,\text{kJ/mol}) \end{array} \qquad (22.17)$$

$$4NO \rightarrow 2N_2O + O_2 \qquad (22.24)$$

$$\begin{array}{c} 2NO_2 \rightleftarrows N_2O_4 \\ \Delta H_{298} = -6.8\,\text{kcal/mol}\ (-28.6\,\text{kJ/mol}) \end{array} \qquad (22.25)$$

For the reaction in (22.17), secondary air is added to the gas mixture obtained from the ammonia oxidation to increase the oxygen content to such a level that the waste gas leaving the plant has an oxygen content between 2 and 4% by volume. The rate of reaction is proportional to pressure cubed (see (22.26)), and the volume or size of equipment needed to oxidize the nitric oxide is inversely proportional to pressure cubed:

$$d(P_{NO})/dt = -k(P_{NO})^2(P_{O_2}) \qquad (22.26)$$

The absorption reaction (22.27) is exothermic, and continuous cooling is required within the absorber:

$$\begin{array}{c} 3NO_2 + H_2O \rightarrow NO + 2HNO_3 \\ \Delta H_{298} = -6.8\ \text{kcal/mol}\ (-28.6\ \text{kJ/mol}) \end{array} \qquad (22.27)$$

The nitric acid produced in the absorber contains dissolved nitrogen oxides and is then bleached by the secondary air.

The overall reaction is

$$\begin{array}{c} NH_3 + 2O_2 \rightarrow HNO_3 + H_2O \\ \Delta H = -98.7\ \text{kcal/mol} \end{array} \qquad (22.28)$$

Several process variations are described in [112].

Catalysts. The catalyst typically consists of several woven or knitted wire gauze sheets with a composition of about 90% platinum, 5% rhodium, and 5% palladium. The primary functions of rhodium are to add mechanical strength to the gauze and to improve yield. The palladium is present to replace a portion of the more expensive rhodium [112]. Figure 22.16 shows that conversion efficiency does not improve if more than 5% rhodium is present in the catalyst [116].

Catalyst life depends on several variables. High-pressure oxidation operations (which also operate at higher temperatures, per Table 22.16) require more frequent catalyst regeneration. The presence of trace amounts of iron, calcium, MoS_2 (lubricants), and phosphorus in the ammonia feed has been shown to shorten catalyst life. Deposits of iron oxides tend to catalyze the conversion of ammonia and oxygen to nitrogen and water, rather than to nitric oxide. These effects, as well as poor ammonia–air mixing and poor gas distribution across the catalyst, may reduce the yield by up to 10% [112, 114].

Some weight loss due to catalyst erosion and vaporization occurs and is more pronounced at the higher converter temperatures that are employed in the higher-pressure processes. Lost catalyst is usually recovered by "getter" systems that use palladium alloy gauzes and recover 60–80% of the lost catalyst. The "getter" gauze is placed as close to the platinum catalyst pack as possible so that the catalyst components can be recovered while they are in the vapor form. Although pure palladium is the most effective "getter," it becomes too brittle, so a palladium alloy is used. Additional catalyst is recovered during periodic cleaning of the ammonia oxidation reactor and during cleaning of downstream heat exchangers [112, 114].

When the catalyst pack no longer provides the desired reaction rates or product composition, new sheets of catalyst may be added to the old sheets. Or the old sheets may be replaced by a set of new sheets, and the old sheets sent back to the supplier to be refurbished. If any catalyst that is lost from the gauze is trapped in the downstream filter or recovered during equipment cleaning, it is also returned to be reprocessed [112].

US Patent 6,073,467 was issued in 2000 for a three-dimensional, knitted, noble metal gauze for nitric acid production. This catalyst design claims to provide elevated product yields and longer catalyst service life with the small noble metal losses.

Operating Pressure. All nitric acid plants are based on the same basic chemical operations: oxidation of ammonia with air to give nitric oxide, oxidation of the nitric oxide to nitrogen dioxide, and absorption in water to give a solution of nitric acid. The efficiency of the first step is favored by low pressure, whereas that of the second step is favored by high pressure. Therefore, both single-pressure and dual-pressure nitric plants are operating [114].

In the single-pressure plant, the oxidation and absorption steps take place at essentially the same pressure. In dual-pressure plants, absorption takes place at a higher pressure than the oxidation stage. The oxidation and absorption steps can be classified as [114]: low pressure (below 1.7 bar), medium pressure (between 1.7 and 6.5 bar), and high pressure (between 6.5 and 13 bar).

The main unit operations in nitric acid plants are [114] ammonia evaporation, ammonia filtration, air filtration, air compression, air/ammonia mixing, catalytic reaction of ammonia and air, energy recovery by steam generation and/or gas reheating, gas cooling, dual pressure only →NO_x compression, absorption with nitric acid production, and tail gas energy recovery.

Most new plants built in the 1990s have a dual-pressure design. Ammonia oxidation is operated at a medium pressure, and absorption is operated at a high pressure. High-pressure systems offer the advantage of lower capital costs (as much as 30% lower than some lower-pressure processes) and are preferred in the United States for this reason. Relative disadvantages are higher catalyst loss and lower ammonia conversion. Lower-pressure oxidation designs have the best efficiencies in ammonia conversion and catalyst conservation. These plants have been preferred in Europe and other areas where ammonia has been relatively expensive and where allowable payout times have been longer [112].

The single-pressure design (see Fig. 22.17) is generally used in smaller plants or in larger plants where minimization of capital is a critical design consideration. Ammonia oxidation and absorption of NO_2 occur at the same relative pressure. This reduces the complexity and capital cost compared to the dual-pressure plant. However, efficiency may be lower in both the front and back ends of the plant. Single-pressure plants can be designed to operate at low to medium pressure to favor the ammonia oxidation reaction, at high pressure to favor the absorption reactions, or at any operating pressure in between [117].

The dual-pressure design (see Fig. 22.18) is generally use in larger plants or in mid-size plants where higher utility/raw material costs dictate a minimization of operating expense. Ammonia oxidation occurs at low or medium pressure. The result is an increase in efficiency of the ammonia oxidation reaction and lower catalyst loss. Absorption of NO_2 occurs at high pressure to maximize the partial pressure of the gas reactants. A nitrous gas compressor boosts the pressure from the front end of the process (ammonia oxidation) to the back end (absorption) [112].

In Table 22.18, the raw material and utility consumption figures are given for both single- and dual-pressure processes [117].

One type of NO_x abatement system, the DCN (*D*estruction by *C*atalysis of NO_x) reactor, is shown in Fig. 22.19 [117].

Concentration of Nitric Acid. Most production of CNAs (i.e., greater than the azeotrope of 68.4%) is based on weak-acid hydration via extractive distillation with a drying agent. Common drying agents are sulfuric acid and magnesium nitrate [112, 118].

A process exists that can directly produce CNA from dehydrated and concentrated nitrogen dioxide, weak acid, and oxygen. Generally the process operates at high pressure and follows these chemical reactions [112]:

$$2NO_2 \rightleftarrows N_2O_4 \tag{22.25}$$

$$N_2O_4 + H_2O + \frac{1}{2}O_2 \rightarrow 2HNO_3 \tag{22.28}$$

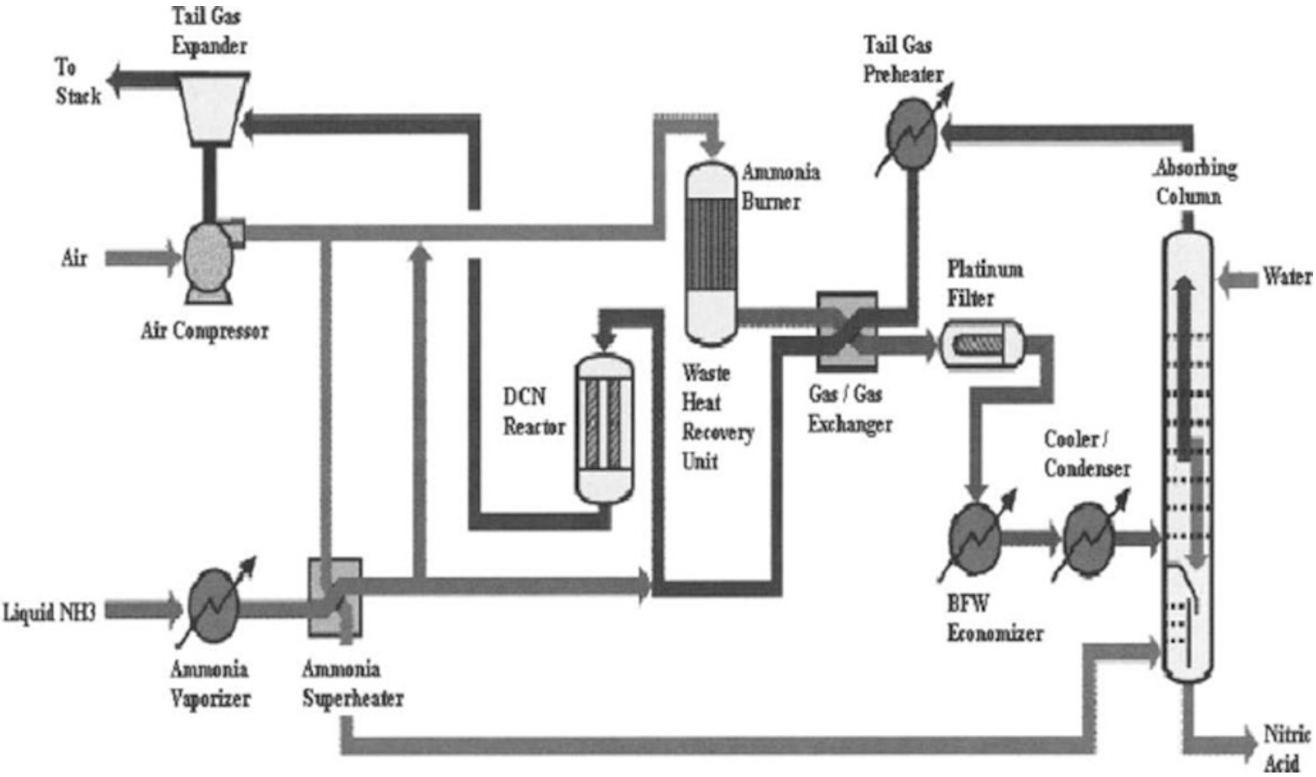


Fig. 22.17 Single-pressure nitric acid plant (Reproduced by permission of Monsanto Enviro-Chem Systems, Inc.)

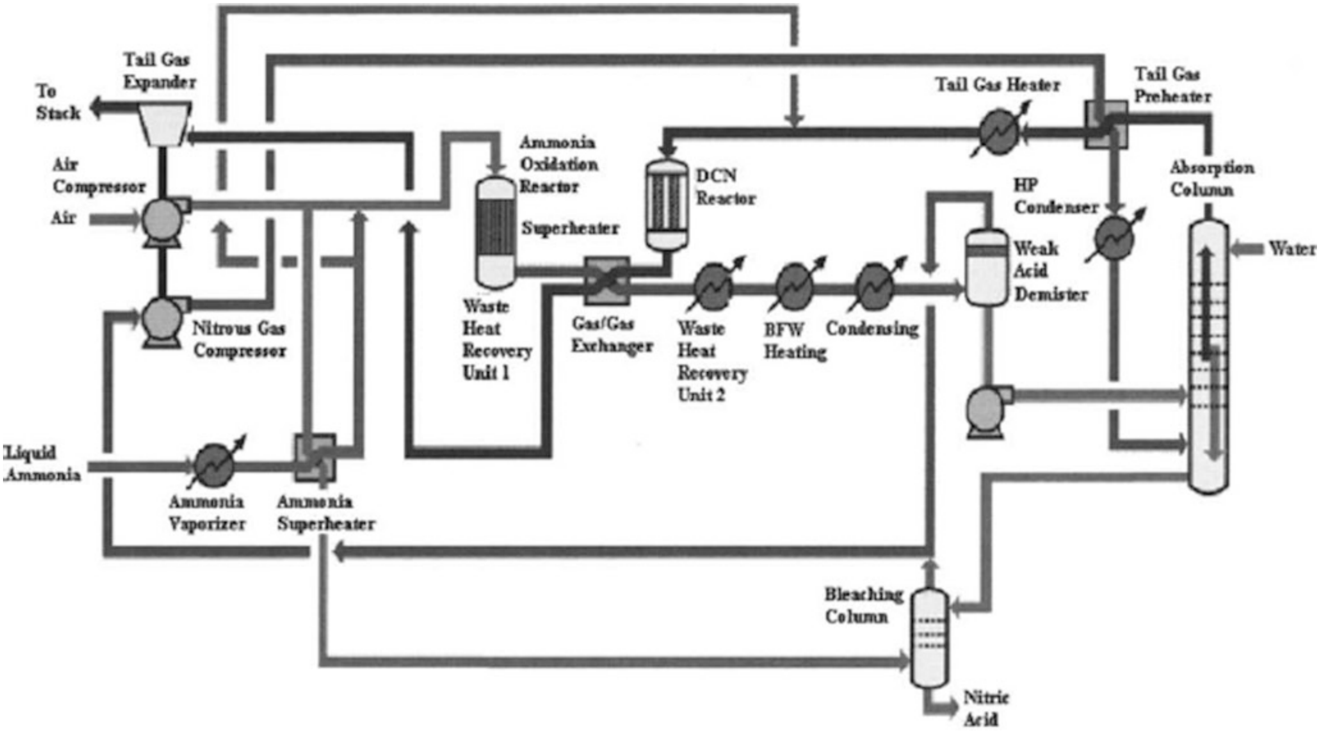


Fig. 22.18 Dual-pressure nitric acid plant (Reproduced by permission of Monsento Enviro-Chem System, Inc.)

Few US plants employ this method for production of concentrated acid. Most production of concentrated HNO_3 (i.e., concentrations above the azeotrope) is based on weak-acid dehydration via distillation with a drying agent. The most common dehydrators have been concentrated sulfuric acid and magnesium nitrite [112].

Table 22.18 Raw material and utility consumptions in different nitric acid plant designs

	H&H	M&M	M&H
Operating pressure, bar a	11	9	4.5/12
Typical run time, days	70	120	210
Ammonia usage[a], ton/ton	0.291	0.285	0.279
Platinum usage[b], mg/ton	50	35	30
Power usage[c], kWh/ton	1	2	2
HP steam export[d], ton/ton	0.51	0.60	0.54
LP steam import[e], ton/ton	0.27	0.21	−0.06
Cooling water[f], m^3/ton	130	120	110

Source: Reproduced by permission of Monsanto Enviro-Chem Systems, Inc.

[a] Includes catalytic NO_x abatement to 200 ppmv and average ammonia conversion efficiency over the entire gauze run length

[b] Described as net of platinum recovery systems

[c] Add 25 kW for HVAC, lighting and instrumentation

[d] Steam conditions are 42 kg/cm^2 g and 400°C

[e] Steam conditions are saturated at 3.5 kg/cm^2 g

[f] At 11°C temperature rise and 2.5 kg/cm^2 pressure allowance

An example of the dehydration process is the Kvaerner Chemetics process that uses concentrated sulfuric acid to eliminate the azeotrope and enable colorless, strong nitric acid up to 99% to be distilled from the mixed acids. The residual acid leaving the bottom of the processing tower is approximately 70% sulfuric acid [119].

Alternative methods of integrated concentration have been developed that do not require separate dehydrating agents. These processes differ from the direct process in that weak acid reacts with concentrated nitrogen dioxide to produce an acid that is sufficiently superazeotropic that distillation into concentrated acid is economically feasible. The weaker azeotropic acid may be recycled for concentration or used as it is [112]. The magnesium nitrate process is described in [120].

Uhde developed another process for the 98–99% that uses oxygen but no dehydrating agents. The NO from ammonia oxidation is converted to NO_2 by highly concentrated HNO_3. The NO_2 is cooled while it is physically absorbed in chilled, highly CNA. Then the NO_2 is separated from the CNA in a distillation column and liquefied by refrigeration. The liquid NO_2 is dimerized in a $N_2O_4/HNO_3/H_2O$ mixture that reacts with supplemental oxygen at a pressure of about 720 psi to form highly CNA. The acid that leaves the reactor contains about 20% dissolved N_2O_4, which is separated in a distillation column

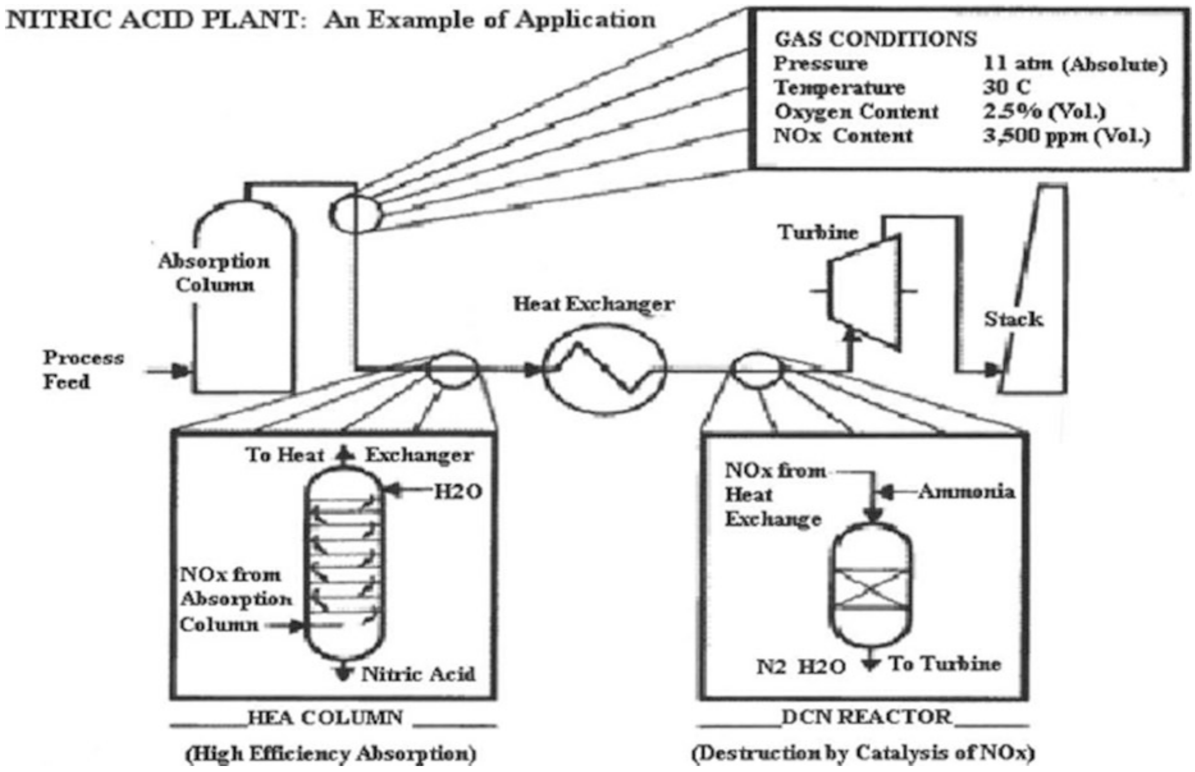


Fig. 22.19 NO_x abatement system with DCN (destruction by catalysis of NO_x) (Reproduced by permission of Monsanto Enviro-Chem System, Inc.)

Table 22.19 Operating conditions in NO_x abatement systems

	DCN only	HEA plus DCN	HEA only
NO_x gas to HEA, ppmv	N/A	3,500	3,500
NO_x gas to DCN, ppmv	3,500	700	N/A
NO_x gas to turbine, ppmv	200	200	200
Nitric acid recovered, kg/ton HNO_3	N/A	27	31
Ammonia consumed, kg/HNO_3	8.5	1.3	N/A
Gas temperature rise, °C	43	6.5	0
Pressure drop, kg/cm^2	0.07	0.2	0.35
Capital cost, Ratio	1.0	2.5	3.5

Source: Reproduced by permission of Monsanto Enviro-Chem Systems, Inc.
DCN Destruction of catalysis of NO_x; *HEA* high efficiency absorption

and returned to the reactor. The CNA product is bleached and sent to storage, and a portion is recycled to the absorber.

Stabilizers. Over a period of time, CNAs tend to decompose according to (22.29), and pressure will build up in storage vessels:

$$HNO_3 \rightarrow 4NO_2 + H_2O + O_2 \qquad (22.29)$$

Because nitric acid is also very corrosive, some stabilizers and/or corrosion inhibitors are used.

Corrosion of aluminum by red fuming nitric acid is reduced by adding 4% by weight of hydrogen fluoride. Decomposition of concentrated acid is reduced by such substances as quaternary ammonium compounds, organic sulfonium compounds [121].

Pollution Abatement. The primary pollution problem in nitric acid plants is the abatement of NO_x in tail gases. The processes that have been developed to reduce emissions at existing and new plants can be classified into four general categories: absorption, adsorption, selective catalytic reduction, and nonselective catalytic reduction [112, 120].

The main environmental factor that affects nitric acid process selection is the concentration of NO_x in the tail gas. In the United States, gaseous emissions from newly constructed nitric acid plants are limited to 1.5 kg NO_x/ton of nitric acid produced with a maximum opacity of 10%. In Western Europe, NO_x emissions are limited by EU regulations to 200 ppm [112].

Absorption abatement refers to modifications that involve the addition of increased absorption capacity or optimization of the existing absorption system [112].

Adsorption abatement uses acid-resistant molecular sieves to absorb the NO_x from the tail gas. The adsorbent is periodically regenerated and the NO_x recovered, converted to NO_2, and recycled for recovery as nitric acid [112].

Selective catalytic reduction is normally used in new nitric acid plants. In this process, ammonia reacts with nitric oxide and nitrogen dioxide but to a lesser extent with oxygen to selectively reduce the NO_x compounds to N_2 [114, 120] as shown below:

$$6NO + 4NH_3 \rightarrow 5N_2 + 6H_2O \qquad (22.30)$$

$$6NO_2 + 8NH_3 \rightarrow 7N_2 + 12H_2O \qquad (22.31)$$

$$3O_2 + 4NH_3 \rightarrow 2N_2 + 6H_2O \qquad (22.32)$$

Nonselective abatement uses a catalyst and fuel (usually a gaseous hydrocarbon) to reduce nitrogen oxides to nitrogen and combust any remaining free oxygen in the tail gas. This process consumes significantly more fuel than a selective reduction system, but the energy from burning the fuel is mostly retrieved as power in an expander [120].

The gas at the outlet of the absorber may vary within the following limits during stable operation: NO_x = 100–3,500 ppmv, N_2O = 300–3,500 ppmv, O_2 = 1.5–4% by volume, H_2O = 0.3–2% by volume, flow = 3,100–3,400 Nm^3/ton of HNO_3.

The minimum emission levels achieved in 2001 without added pollution abatement are

- Medium-pressure absorption: 1,000–2,000 ppmv
- High-pressure absorption: 100–200 ppmv

Whereas a dual-pressure plant or a high-pressure, single-pressure plant (with high-pressure absorption) may give acceptable emission levels, the medium-pressure absorption plant must be followed by an abatement system [114].

Monsanto Enviro-Chem offers NO_x abatement technology that is licensed from Rhodia of France. It includes a high-efficiency absorption (HEA) section for extended absorption and a catalytic reduction section (SCR) for catalytic destruction of NO_x (i.e., the DCN technology) [117]. Additional process details are given in [117]. The operating conditions for the steps in the Monsanto technology are compared in Table 22.19.

Production. An estimated 65% of the worldwide nitric acid production is used to make ammonium nitrate (AN) for use in fertilizers. Other AN uses (mainly explosives) account for about 15% of the market. The balance is consumed in a variety of non-AN industrial applications. The production of AN fertilizers and most industrial explosives require acid concentrations less than the azeotropic concentration of 68.4%.

World nitric acid production in 1999 is estimated to be 53 million tons/year. Captive use consumes most of the production, and the merchant market probably involves only about 10% of the total [112].

World demand for nitric acid will continue to be largely dependent upon demand for solid ammonium nitrate fertilizer and nitrogen fertilizer solutions that incorporate ammonium nitrate. Since the 1980s, urea has been replacing solid ammonium nitrate as a fertilizer. This has been partially offset by increased use of AN in explosives, polyurethane foams, declined by about 5% between 1987 and 1999, but it is projected to increase marginally by 2005. The major producing regions are Western Europe and the United Eastern Europe [112, 120].

Nitric acid is sold commercially in various concentrations that are expressed in degrees Baumé (Be) which is an alternate specific or fuming nitric acid (95% or more HNO_3/NO_2) are available [112, 120].

Storage and Distribution. Nitric acid is normally stored in flat-bottomed, roofed tanks that are made from low-carbon, austenitic acid are transported in tank cars and by truck. Stainless steel is necessary for concentrations up to 80–85%. Stronger solutions are less corrosive and may be stored in aluminum.

In the United States, the Department of Transportation (DOT) defines three categories of nitric acid: nonfuming, more than 70 wt.% acid; nonfuming, less than 70 wt.% acid; and red fuming nitric acid. All must be labeled “corrosive.”

Nitric acid is subject to self-decomposition, depending on the concentration and temperature. The general rule is that the higher the concentration or the temperature, the faster the decomposition rate. Nitric acid must be transported from the manufacturer to the user in the summer months [114].

Uses of Nitric Acid

Some of the major uses of nitric acid are

- *Ammonium nitrate (AN)*: Production for fertilizers and explosives.
- *Adipic acid*: Nitric acid oxidizes cyclohexanone–cyclohexanol mixtures to adipic acid, which goes into nylon 6,6, other resins, and plasticizers.
- *Nitrobenzene*: Nitrobenzene is made by the direct nitration of benzene with nitric/sulfuric acid mixtures.
- *Toluene diisocyanate (TDI)*: TDI (2,4- and 2,6 isomers) are produced from toluene diamine derived from dinitrotoluene, which is produced by the nitration of toluene with nitric/sulfuric acid mixtures. TDI is used mainly in
- *Potassium nitrate*: Production for tobacco fertilizer and industrial markets.
- *Sodium nitrate*: Production for fertilizers, explosives, and as a heat transfer medium.

Some of the other uses of nitric acid are as follows: other metal nitrates, in the steel industry, in the electronics industry, to make nitrochlorobenzene; cellulose nitrate, in several nitrate paraffins (e.g., nitromethane, nitroethane, and nitropropane), in nonmilitary explosives, to digest crude uranium concentrates, in mixed fertilizers, and in CNA.

Ammonium Nitrate

Ammonium nitrate (AN) was first prepared in the sixteenth century. Its early industrial development was primarily for use in explosives. However, after 1940, its use as a fertilizer exploiter developed rapidly. It is made by the reaction of gaseous ammonia with aqueous nitric acid:

$$\begin{aligned} &HNO_3(l) + NH_3(g) \rightarrow NH_4NO_3(aq) \\ &\Delta H = +26\ \text{kcal/g mol} \end{aligned} \tag{22.33}$$

The resulting AN solution may be handled in various ways [122]. It can be stored as a solution used in downstream plants or sold as a solution, it can be formed into solid AN by prilling or granulation, or it can be mixed with a solid filler. The most common filler is calcium carbonate in the form of ground limestone, dolomite, or by-product calcium carbonate. This product is known as calcium ammonium nitrate (CAN) and can be prilled or granulated.

Processes

As shown in Fig. 22.20 [123], the typical AN production process has three main unit operations: neutralization, evaporation, and solidification (prilling and granulation) [122, 124]. Additional details about AN production are given in [125].

Neutralization. The exothermic neutralization of nitric acid with ammonia gas produces AN solution and steam. The nitric acid is commonly preheated if the available concentration of nitric acid is in the lower end of the 50–70% range [122].

Neutralization can be performed in a single stage or in two stages. A two-stage neutralizer operates with a low pH in the first stage and a neutral pH in the second stage.

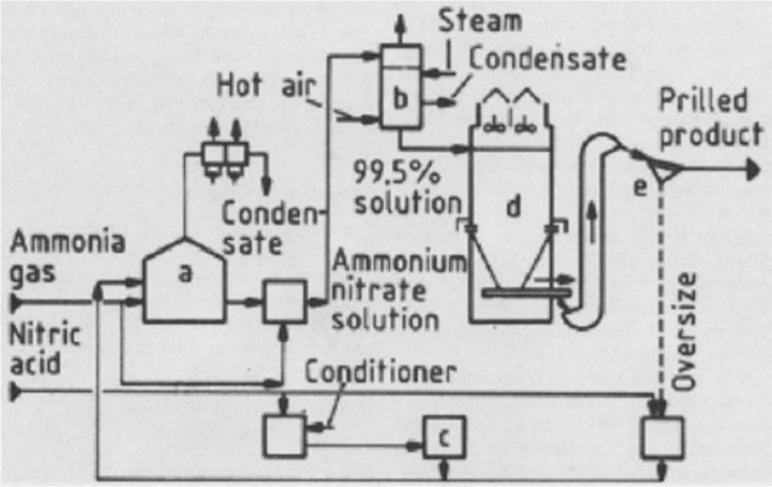


Fig. 22.20 Ammonium nitrate process sketch (Courtesy of Wiley-VCH [123])

In most neutralizers, the pressure, temperature, and concentration are linked by the boiling point characteristics of AN solutions with only two of these variables being independent [122].

Neutralizers may be free-boiling vessels, circulating systems, or pipe reactors. The water in the nitric acid is evaporated by the heat of reaction (550–620 BTU/lb of AN, depending on the acid strength). At least ten different types and designs of neutralizers are used in Europe. The factors that influence the choice of neutralizer design are described in [122]. Different neutralizer designs are discussed in [123, 126, 127].

Steam Purification. The steam that leaves the neutralizer can be purified, or it can be condensed and then purified. Some steam purification techniques are described in [122]. The steam may be used in the evaporator, be used to preheat and evaporate ammonia, or be used to preheat the nitric acid [122].

Evaporation. The evaporator is normally needed to remove water from the AN solution. It must produce a solution with the required concentration at a temperature that avoids crystallization. The acceptable water content is normally below 1% for a prilled product and up to 8% for the feed to some granulation processes. Evaporators in commercial use include circulatory systems, shell and tube heat exchangers, and falling film-type evaporators [122].

Prilling. Prilling refers to the formation of granules by the solidification of droplets of AN. The AN solution is concentrated to 96–99% at a pH above 4.5 to form a "melt." The melt may be mixed with an additive or nucleating agent that stabilizes the prills against temperature cycling through the crystal transition phases. Then the melt is transported to the top of the prilling tower and sprayed through a system of fixed orifices or dropped through a rotating perforated bucket into a rising air stream. The droplets crystallize into hard spherical "prills" that are dried, cooled, and sized for shipment. If CAN is made, ground calcium carbonate (limestone or dolomite) is added to the melt prior to the formation of the droplets [122, 124, 128].

SASOL in South Africa produces a porous, prilled ammonium nitrate (PPAN) that finds its widest application in a mix with fuel oil. This mixture is used as an explosive and is commonly known as ANFO (*a*mmonium *n*itrate–*f*uel *o*il). Additional details about PPAN are given in [129].

Granulation. Granulation refers to techniques using processes such as agglomeration, accretion, or crushing to make a granular fertilizer. In contrast to the prilling technique, granulation requires a more complicated plant, and a variety of equipment is used: rotating pans and drums, fluidized beds, and other equipment. Granular products can be made in a wider choice of particle sizes than prills [122].

The AN is added in the granulator as a spray of hot concentrated solution. No further drying of the granules will normally be required. The granules are screened, and the fines and crushed oversize returned to the granulator [122].

Typical CAN granulators include drums and pugmills. The calcium carbonate may be mixed with the AN solution before granulation or in the granulator itself. Granules from this process will normally require drying in a fluidized bed or rotary drier [122].

Air Abatement Equipment. Emissions from the prilling and granulation sections of AN and CAN plants can be treated by a range of abatement equipment. Particulate material from some granulation plants is relatively coarse in particle size, whereas the prill tower emissions contain very fine particles.

Candle filters are normally required for prill tower emissions. They can abate particulate emissions down to 15 mg/m^3 of air. For coarser material, dry devices such as bag filters or dry cyclones are used. Particulate emissions can have higher concentrations, perhaps up to a range of 30–50 mg/m^3, but the recovered material is a solid that can more readily be recycled.

Production

Production of fertilizer grade AN is concentrated mainly in Europe and North America. In 2001, US production was 7.1 million tons, which was 18% lower than the peak production in 1998. US plants were operating at 70% of nameplate capacity in 2001 [128]. International trade amounts to about 4 million tons (on a nitrogen equivalent basis) per year. Additional information can be found in Chap. 24.

Some of the physical and chemical properties of ammonium nitrate are given in Table 22.20 [122].

Table 22.20 Ammonium nitrate physical and chemical properties

Property	
Appearance	White or off-white granules or prills
Odor	Odorless
pH (10 g/100 mL)	>4.5
Melting point	160–170°C (depending on moisture content)
Decomposition point	>210°C
Molecular weight	80.05
Solubility in water	
0°C	118.3 g/100 g
20°C	190.0 g/100 g
80°C	576.0 g/100 g
Bulk density	830–1,100 kg/m^3
Nitrogen content	35% (100% ammonium nitrate)

Storage and Distribution. Solid AN in packages must be stored in a general warehouse that has been approved for AN duty. Bulk AN and CAN must be protected from moisture as both products are hygroscopic. Large bulk warehouses may be air-conditioned due to the local climate. Some additives can also reduce the water uptake. The anticaking agents may be used that are internal to the finished particle or applied as a coating [122, 128].

AN solution must be stored at a temperature above the crystallizing temperature of the solution. Gaseous ammonia is normally added in small quantities to maintain the solution at the correct pH because AN solutions lose ammonia during storage [122]. Steps should also be taken to avoid contamination by decomposition catalysts such as chlorides and organic materials.

Ammonium nitrate decomposes in two ways. Controlled decomposition with careful heating is the commercial process for producing nitrous oxide:

$$\begin{aligned} &\mathrm{NH_4NO_3} \xrightarrow{200-260^\circ} \mathrm{C\,N_2O + H_2O} \\ &\Delta \mathrm{H} = -6.7\ \mathrm{kcal/g\,mol} \end{aligned} \tag{22.34}$$

Explosive decomposition by heating or by explosive shock is the basis for its use as an explosive:

$$\begin{aligned} &\mathrm{2NH_4NO_3 \rightarrow 2N_2 + 4H_2O + O_2} \\ &\Delta \mathrm{H} = -28.2\ \mathrm{kcal/g\,mol} \end{aligned} \tag{22.35}$$

The oxygen that is released in (22.35) can be used to more than triple the explosive effects by mixing an organic, such as fuel oil, with AN. This mixture, known as ANFO, is much cheaper and much safer to handle than dynamite:

$$\begin{aligned} &\mathrm{3NH_4NO_3 + (CH_2)_{\mathit{n}} \rightarrow 3N_2 + 7H_2O + CO_2} \\ &\Delta \mathrm{H} = -102.5\ \mathrm{kcal/g\,mol} \end{aligned} \tag{22.36}$$

Urea

Urea was discovered in 1773, and it was first synthesized from ammonia and cyanic acid in 1828 [130]. In the early 1900s, urea was produced on an industrial scale by the hydration of cyanamide, which was obtained from calcium cyanamide [131]:

$$\mathrm{CaCN_2 + H_2O + CO_2 \rightarrow CaCO_3 + CNNH_3} \tag{22.37}$$

$$\mathrm{CNNH_3 + H_2O \rightarrow CO(NH_2)_2} \tag{22.38}$$

After development of the ammonia process by Haber and Bosch in 1913, the production of urea from ammonia and CO_2 developed rapidly. In 2001, urea is prepared on an industrial scale exclusively by this method [131].

$$\underset{\text{Ammonia}}{\mathrm{2NH_3}} + \underset{\text{Carbon dioxide}}{\mathrm{CO_2}} \rightleftarrows \underset{\text{Ammonium carbamate}}{\mathrm{NH_2COONH_4}} \tag{22.39}$$

$$\underset{\text{Ammonium carbamate}}{\mathrm{NH_2COONH_4}} \rightleftarrows \underset{\text{Urea}}{\mathrm{CO(NH_2)_2}} + \underset{\text{Water}}{\mathrm{H_2O}} \tag{22.40}$$

Reaction (22.39) is fast and exothermic and essentially goes to completion under the high-pressure reaction conditions. Reaction (22.40) is slower, is endothermic, and does not go to completion. The conversion (on a CO_2 basis) is usually 50–80%. The conversion increases with increasing temperature and NH_3/CO_2 ratio. It decreases with increasing H_2O/CO_2 ratio [132].

Table 22.21 Physical properties of urea [130]

Chemical formula	NH_2CONH_2
Molecular weight	60.06
Freezing/melting point	132.6°C
Boiling point	Decomposes
Density, d_4 [17]	1.3230 g/cm^3
Heat of solution in water	+251 J/g (60 cal/g)
Bulk density	0.74 g/cm^3
Specific heat, J/(kg K)	
0°C	1.439
50°C	1.661
100°C	1.887
150°C	2.109

Properties of saturated aqueous solutions of urea

Temperature (°C)	Solubility in water (g/100 g solution)	Density (g/cm^3)	Viscosity (mPa s = cP)	Water vapor pressure (kPa)
0	41.0	1.120	2.63	0.53
20	51.6	1.147	1.96	1.73
40	62.2	1.167	1.72	5.33
60	72.2	1.184	1.72	12.00
80	80.6	1.198	1.93	21.33
100	88.3	1.210	2.35	29.33
120	95.5	1.221	2.93	18.00
130	99.2	1.226	3.25	0.93

When urea melt is subjected to heat, some biuret is formed:

$$\underset{\text{Urea}}{2CO(NH_2)_2} \rightleftarrows \underset{\text{Biuret}}{NH_2CONHCONH_2} + \underset{\text{Ammonia}}{NH_3} \qquad (22.41)$$

Biuret is undesirable for some industrial and fertilizer applications. Prilled urea normally contains more biuret (0.8–1.5%) than granular urea because prilled urea goes through an extra evaporation step.

Urea is an odorless white solid that contains 46.65% nitrogen. It is not flammable, but it will melt and decompose in a fire to give off ammonia. When it is dissolved in water, it hydrolyzes very slowly to ammonium carbamate and eventually decomposes to ammonia and carbon dioxide. Some of the other properties of urea are given in Table 22.21 [130, 132].

Processes

In 2001, most new urea plants are licensed by Snamprogetti (SNAM), Stamicarbon (STAC), or Toyo. SNAM utilizes thermal stripping, whereas STAC and Toyo use CO_2 stripping. At the end of 1996, about 70 SNAM plants, 125 STAC, and 7 Toyo plants had been built. STAC will design plants for over 3,000 tons/day; SNAM designs plants for about 2,800 tons/day; and Toyo designs plants for about 2,300 tons/day as single-train units [130].

Because urea is made from ammonia and carbon dioxide, all urea plants are located adjacent to or in close proximity to an ammonia plant. Figure 22.21 [133] shows an example of an ammonia plant and a urea plant that are part of the same complex.

The SNAM process, the synthesis section of the STAC process, and the synthesis section of the Toyo ACES (*a*dvanced *p*rocess for *c*ost and *e*nergy *s*aving) process are described in [130, 132]. An updated version of the STAC process, Urea 2000plus, is offered by DSM (STAC's parent company) [130, 132].

The *i*sobaric *d*ouble-*r*ecycle (IDR) urea process was developed by Montedison. In this process, most of the unconverted material leaving the reactor is separated by heating and stripping at synthesis pressure using two strippers in series. The unconverted ammonia, CO_2, and carbamate in the urea solution are recycled to the synthesis loop [130, 132].

The *H*eat *R*ecycle *U*rea *P*rocess (HRUP) was developed by Urea Technologies in the 1970s and is offered by Monsanto Enviro-Chem. This process is described in [132, 134]. Several urea-producing reactors have been updated with Casale high-efficiency trays. These trays increase conversion, which increases plant capacity and reduces energy use [102].

Prilling. A prilling plant is shown in Fig. 22.22 and is described in [130, 132].

Granulation. Almost all new plants that make granules use the Hydro-Agri process, which can be built to make over 3,000 tons/day. The basic principle of the process involves the spraying of the melt onto recycled seed particles or prills circulating in the granulator. The process is described in more detail in [130–132].

Another process is the C&I Girdler drum system, but it cannot compete in today's market because of capacity restrictions. Toyo has developed a spouting-fluid-bed technology, and by 1996, three plants were in operation. STAC will also license a fluid-bed plant that is similar to the Hydro-Agri design [130].

Production

The ammonia cost makes up 62 to over 70% of the cash production costs of urea as the natural gas cost varies from $2.00 per million BTU up to $6.00 per million BTU. In general urea cash production costs in dollars per ton are a factor of 1.4–1.6 times the ammonia price in US dollars per ton [135] (see [32, 136, 137] for details about urea production).

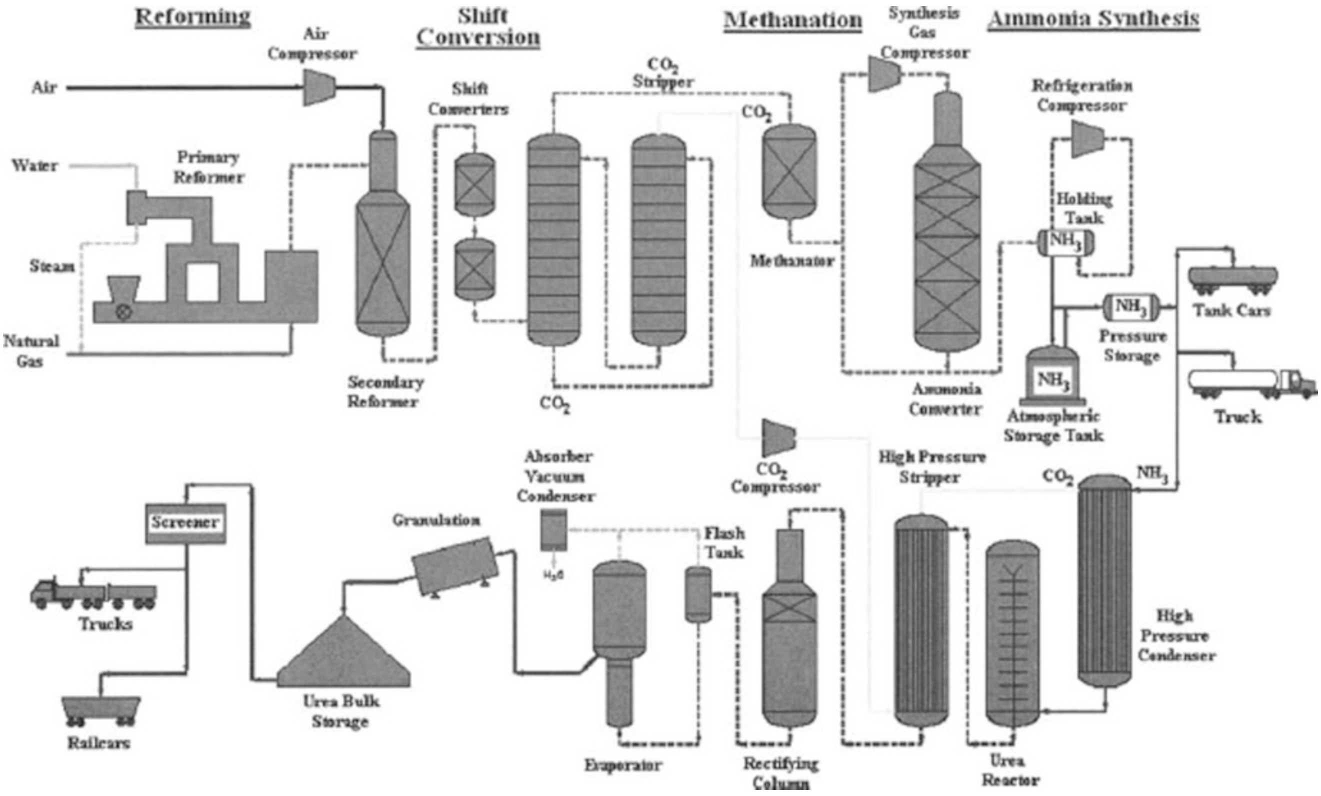


Fig. 22.21 Agrium's Carseland nitrogen operations: ammonia and urea production (Agrium Carseland Nitrogen Operation, Carseland Alberta, Canada, 3/01. Image may not be duplicated or reproduced without the written permission of Agrium)

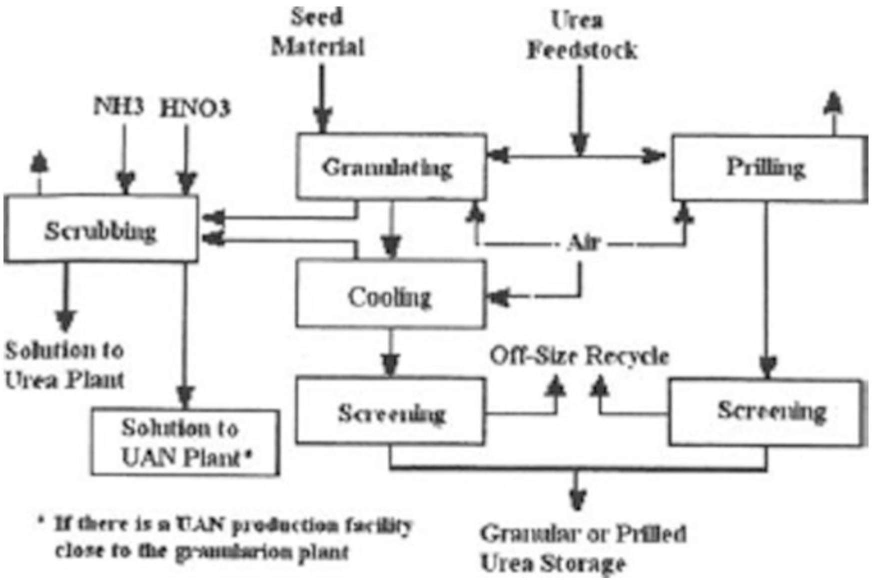


Fig. 22.22 Block diagram for urea granulation and prilling processes [132] (Reproduced by permission of European Fertilizer Manufacturers Association)

Storage and Distribution

Urea may be supplied as a solid or in a solution, and solid urea is classified as granular or prilled products. Prilled products have less desirable qualities than granules and are usually marginally cheaper than granulated product [131].

The majority of urea is designated as fertilizer grade. Technical grade does not contain any additives. Low-biuret grade contains less than 0.3% biuret and is used on citrus crops. Feed grade is fed directly to cattle. It is free of additives and is supplied as microprills with a diameter of about 0.5 mm. Slow-release grades use coatings and additives to increase the amount of nitrogen absorbed by crops. Urea supergranules have diameters up to 15 mm and are used in wetland rice and forest fertilization [131].

The shift from bagged to bulk transport and storage of prilled and granulated urea has called for warehouse designs in which large quantities of urea can be stored in bulk. Caking and subsequent product degradation at unloading are the result of water absorption. To avoid these problems, the warehouse should be airtight and thoroughly insulated. The caking can also be reduced by adding small amounts of formaldehyde (up to 0.6 wt.%) to the urea melt or by adding surfactants to the solid product [131].

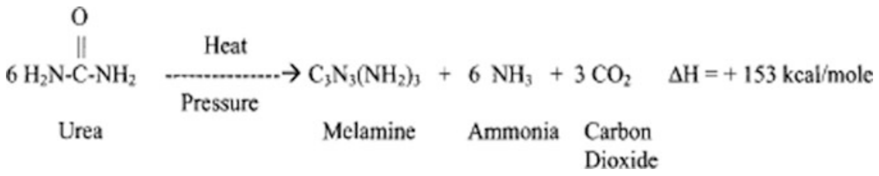
Processes

Melamine can be produced from urea, dicyandiamide, or hydrogen cyanide, but in 2001, all commercial-scale production is based on urea [138].

Melamine is produced from urea by either a high- or a low-pressure process, and either process can consist of one or two stages. The net reaction is [138]

The reactions in the two-stage process are shown in (22.43) and (22.44). The products are quenched with water or an aqueous mother liquor. This stream is then processed by filtration, in centrifuges, or by crystallization [138].

In a typical high-pressure process such as the one licensed by Eurotecnica (see Fig. 22.23 and http://www.eurotecnica.it/melamine.htm), the reaction is carried out in the liquid phase (without a catalyst) at 90–150 bar and 380–450°C.

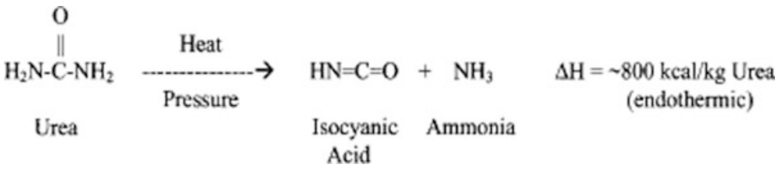


$$\underset{\text{Urea}}{6\,H_2N\text{-}\overset{\overset{O}{\|}}{C}\text{-}NH_2} \xrightarrow[\text{Pressure}]{\text{Heat}} \underset{\text{Melamine}}{C_3N_3(NH_2)_3} + \underset{\text{Ammonia}}{6\ NH_3} + \underset{\text{Carbon Dioxide}}{3\ CO_2} \quad \Delta H = +153\ \text{kcal/mole} \tag{22.42}$$

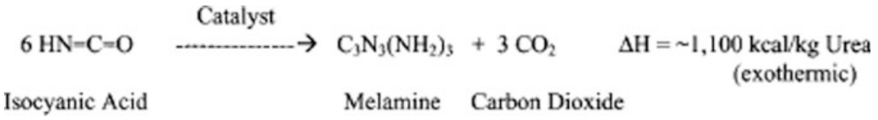
Uses

Urea is used in [135] liquid and solid fertilizers, urea-formaldehyde resins that make adhesives and binders mostly for wood products, livestock feeds, melamine for resins, and NO_x control for boilers and furnaces.

Under these conditions, urea forms cyanuric acid which then reacts with ammonia to form melamine. The first step is the same as (22.43), and the second step is shown in (22.45) and (22.46) [44, 138]:



$$\underset{\text{Urea}}{H_2N\text{-}\overset{\overset{O}{\|}}{C}\text{-}NH_2} \xrightarrow[\text{Pressure}]{\text{Heat}} \underset{\text{Isocyanic Acid}}{HN{=}C{=}O} + \underset{\text{Ammonia}}{NH_3} \quad \Delta H = \sim 800\ \text{kcal/kg Urea (endothermic)} \tag{22.43}$$



$$\underset{\text{Isocyanic Acid}}{6\ HN{=}C{=}O} \xrightarrow{\text{Catalyst}} \underset{\text{Melamine}}{C_3N_3(NH_2)_3} + \underset{\text{Carbon Dioxide}}{3\ CO_2} \quad \Delta H = \sim 1{,}100\ \text{kcal/kg Urea (exothermic)} \tag{22.44}$$

Melamine

Melamine is characterized by its three amino ($-NH_2$) groups and by its 66.6% nitrogen content. Its chemical formula can be written as $C_3H_6N_6$ or $C_3N_3(NH_2)_3$. It was first prepared in 1834, but it was not until 1937 that it was produced commercially.

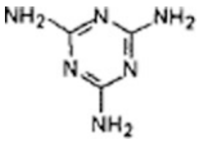
Some of the properties of melamine are listed in Table 22.22 [138].

$$\underset{\text{Isocyanic acid}}{3HN = C = O} \rightarrow \underset{\text{Cyanuric acid}}{C_3N_3(OH)_3} \tag{22.45}$$

$$\underset{\text{Cyanuric acid}}{C_3N_3(OH)_3} + \underset{\text{Ammonia}}{3NH_3} \rightarrow \underset{\text{Melamine}}{C_3N_3(NH_2)_3} + 3H_2O \tag{22.46}$$

In a typical low-pressure process, the reaction is carried out in the vapor phase at 1–10 bar and 350–400°C. This process

Table 22.22 Properties of melamine

Appearance	Fine white crystalline powder
Molecular weight	126.13
Specific density, g/cc	1.573
Melting point	354°C
Boiling point	Decomposes above 300°C and releases ammonia vapors
Bulk density	750 kg/m^3
Toxicity	Mild/low
Stability	Stable
Solubility	Insoluble in most inorganic compounds Very soluble in water
Melamine structure	

uses a catalyst such as modified aluminum oxide or alumino-silicate. Under these conditions, urea forms isocyanic acid as shown in (22.43). On the catalyst, the isocyanic acid is converted to cyanamide or carbodiimide which is subsequently converted to melamine (see (22.47) and (22.48)) [44, 138]:

$$\underset{\text{Isocyanic Acid}}{2\,HN{=}C{=}O} \rightarrow \left(\underset{\text{Carbodiimide}}{HN{=}C{=}NH} \quad \text{or} \quad \underset{\text{Cyanamide}}{H_2N{=}C{\equiv}N}\right) + \underset{\text{Carbon Dioxide}}{CO_2} \tag{22.47}$$

$$\left(\underset{\text{Carbodiimide}}{HN{=}C{=}NH} \quad \text{or} \quad \underset{\text{Cyanamide}}{H_2N{=}C{\equiv}N}\right) \rightarrow \underset{\text{Melamine}}{C_3N_3(NH_2)_3} + \underset{\text{Carbon Dioxide}}{CO_2} \tag{22.48}$$

Carbon dioxide and ammonia are obtained as by-products in both the low- and high-pressure processes and are usually recycled to the urea process to improve process economics. With recycling, about 1.5 lb of urea is used to produce 1 lb of melamine which corresponds to a yield of 95% (compared to a yield of only 49% without recycle) [138].

DSM Melamine started up a new 30,000-tons/year melamine plant in 2004 and reached design capacity in 2006. This plant will use the high-pressure shortened liquid-phase (SLP) process developed by DSM. The SLP process is expected to increase efficiency by 25% and have the same costs as a 100,000-tons/year production facility. The new process requires only three or four processing steps, in contrast to the ten steps in conventional processes [139].

The major melamine process technologies and the suppliers of these technologies are listed in [138].

Production

In 2001, worldwide capacity was 1.1 million tons/year [140]. In 1998, Western Europe accounted for 35% of world capacity, followed by Asia (excluding Japan) with 24%, and both the United States and Japan with 16% each [138]. Demand for melamine is heavily influenced by construction and auto manufacture [138].

Uses

Melamine is supplied in bags, carlots, and truckloads. Some of the melamine uses are surface coatings, laminates, paper treating, molding compounds, textile treating, wood adhesives, and other uses such as ceiling tiles, tire cord, flame retardants, and many other small-volume applications [138, 141].

Aliphatic Amines

Amines are derivatives of ammonia in which the hydrogen atoms in the ammonia have been displaced by monovalent hydrocarbon radicals. Depending on the number of hydrogen atoms displaced, amines are classified as primary (RNH_2), secondary (R_2NH), or tertiary (R_3H). The amines can be categorized further according to the type of hydrocarbon radicals that substitute for the hydrogen atoms: aliphatic (saturated or unsaturated), aromatic, heterocyclic, alicyclic, or any combination of these.

Imines are one group of compounds that are similar to amines. Imines contain an ammonia molecule in which two hydrogen atoms are displaced by bivalent hydrocarbon radicals ($R = NH$). Another group of compounds that are similar to amines is nitriles. In nitriles, all the hydrogen atoms in ammonia are displaced by a trivalent hydrocarbon radical ($RC \equiv H$).

Methylamines

Methylamines are colorless liquids that are volatile at normal atmospheric conditions. They have threshold odor limits of less than 10 ppm, and at low concentrations, they have a fishy smell. At high concentrations, they smell like ammonia. The physical properties are given in Tables 22.23 and 22.24.

Conventional Processes. The reaction of an alcohol with ammonia provides the most common commercial routes to

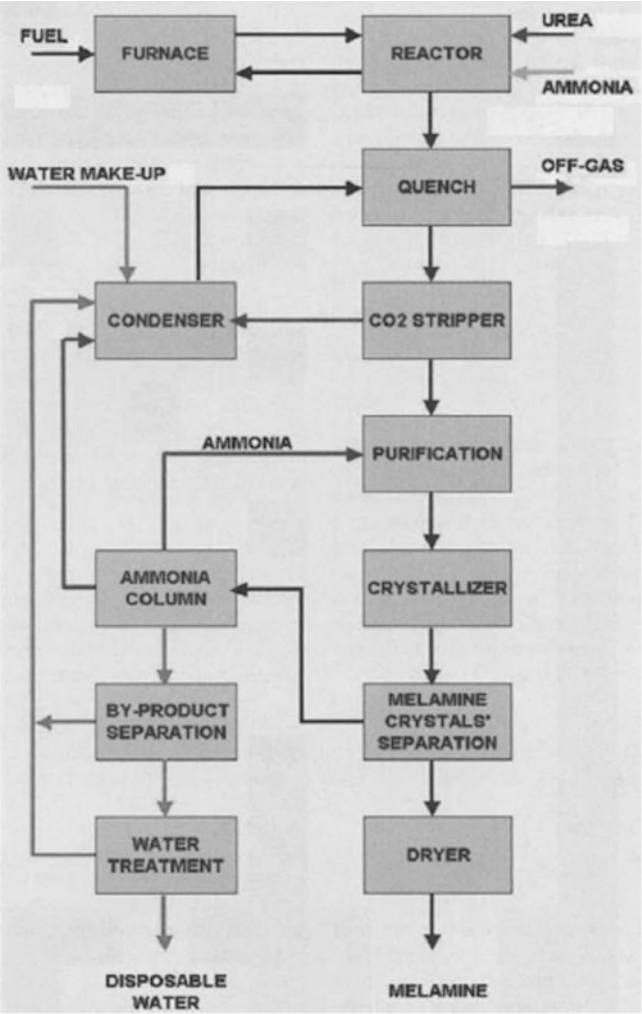


Fig. 22.23 Eurotecnica high-pressure melamine process (Reproduced by permission of Eurotecnica)

alkylamines. Capacities of many plants depend on the product mix of mono/di/tri products as well as the variety of amines (ethyl, propyl, and butyl) [142].

In alcohol amination, methanol and excess ammonia react at 350–500°C and 15–30 bar in the presence of aluminum oxide, silicate, or phosphate catalysts according to the following reactions [44]:

$$NH_3 + CH_3OH \xrightarrow{Cat} CH_3NH_2 + H_2O \quad \Delta H = -5.358 \text{ kcal/mol} \tag{22.49}$$

$$CH_3NH_2 + CH_3OH \xrightarrow{Cat} (CH_3)_2NH + H_2O \quad \Delta H = -9.598 \text{ kcal/mol} \tag{22.50}$$

$$(CH_3)_2NH + CH_3OH \xrightarrow{Cat} (CH_3)_3N + H_2O \quad \Delta H = -14.098 \text{ kcal/mol} \tag{22.51}$$

Table 22.23 Physical properties of anhydrous methylamines

	MMA	DMA	TMA
Chemical formula	CH_3NH_2	$(CH_3)_2NH$	$(CH_3)_3N$
Molecular weight	31.06	45.08	59.11
Freezing point, °C	−93.5	−92.02	−117.3
Boiling point, °C	−6.3	6.9	2.9
Vapor density (1 atm, 25°C), g/L	1.3	1.9	2.5
Specific gravity, air = 1	1.1	1.6	2.0
Liquid density (25°C), g/mL	0.6562	0.6496	0.6270
lb/gal	5.48	5.42	5.23
Liquid viscosity (25°C), cP	–	0.190	0.175
Critical temperature, °C	156.9	164.5	160.1
Critical pressure, atm	73.6	52.4	40.2
Enthalpy of formation, ΔH_f° (25°C), kcal/mol			
Gas	−5.49	−4.41	−5.81
Liquid	−11.3	−10.5	−11.0
Solution (1,000 mol H_2O)	−16.78	−17.3	−18.6
Vapor pressure, psia			
25°C	50	30	32
100°C	400	230	205
Heat capacity, c_p (25°C), cal/°C mol			
Gas	12.7	16.9	21.9
Liquid	–	32.9	32.31
Heat of fusion (m. p.), cal/g	47.20	31.50	26.46
Heat of vaporization (b.p.), cal/g	198.6	140.4	92.7
Autoignition temperature, °C	430	400	190
Flammable limits, vol.%			
Lower	4.9	2.8	2.0
Upper	20.7	14.4	11.6

Source: Du Pont Methylamines Bulletin
MMA monomethylamine; *DMA* dimethylamine; *TMA* trimethylamine

$$NH_3 + (CH_3)_3N \xrightarrow{Cat} CH_3NH_2 + (CH_3)_2NH \quad \Delta H = +8.740 \text{ kcal/mol} \tag{22.52}$$

$$NH_3 + (CH_3)_2NH \xrightarrow{Cat} 2CH_3NH_2 \quad \Delta H = +4.240 \text{ kcal/mol} \tag{22.53}$$

$$CH_3NH_2 + (CH_3)_3N \xrightarrow{Cat} 2(CH_3)_2NH \quad \Delta H = +4.500 \text{ kcal/mol} \tag{22.54}$$

Reactions (22.49), (22.50), and (22.51) are known as the alkylation reactions. They are exothermic and highly irreversible, except for (22.51). Reactions (22.52), (22.53), and

Table 22.24 Physical properties of aqueous methylamines

Properties	40% MMA	40% DMA	60% DMA	25% TMA
Boiling point, °C	48	54	36	43
Freezing point, °C	−38	−37	−74.5	6
Liquid density (25°C)				
g/mL	0.897	0.892	0.829	0.930
lb/gal	7.49	7.44	6.92	7.76
Vapor pressure (25°C)				
psia	5.8	4.2	9.7	6.6
mmHg	300	215	500	340
kPa	40	29	67	45
Flash point, closed cup °C	−12	−18	−52	6

(22.54) are known as disproportionation reactions. They are reversible and are endothermic. The alkylation reactions dictate the rate of consumption of methanol and are somewhat faster than the disproportionation rates that govern the selectivity of the three amines. Guidelines for optimizing the reactions are discussed in [44, 142].

Figure 22.24 shows a typical methylamine process diagram [143]. The raw material and utility use are summarized in Table 22.25 [24].

Although companies can produce a range of C_2 to C_6 amines on a campaign basis in a higher-alkylamine plant, methylamine plants are usually designed and dedicated solely for production of a given chain-length product. The methylamine plants also use a different catalyst from that in a higher-amine plant [142].

Mitsubishi Rayon has operated a process since 1984 in which the equilibrium of methylamine formation is shifted to make more dimethylamine by use of an acid zeolite catalyst. The product stream contains 7 mol.% MMA, 86 mol.% DMA, and 7 mol.% TMA, and the investment as well as the operating costs are lower than the conventional Leonard process that is used by most companies. Because DMA has the highest sales volume, the process might be appealing.

Aldehyde–Amine and Ketamine Hydrogenation Process. Imines can be addition of ammonia to a carbonyl compound, and the imines are hydrogenated to make amines per (22.55) [142]:

$$RCHO + NH_3 \rightarrow RCH = NHH_2RCH_2NH_2 \qquad (22.55)$$

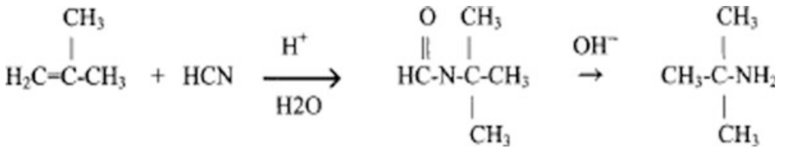
Nitrile Reduction. The choice of catalyst and variations of the temperature, pressure, and concentration of ammonia can control the product mix of amines from hydrogenation of a nitrile. The reaction is shown in (22.56):

$$RC{\equiv}N + H_2 \rightarrow RCH_2NH_2 + (RCH_2)_2NH_2 + (RCH_2)_3N \qquad (22.56)$$

A large excess of ammonia suppresses formation of di- and trialkylamines. Reference [142] describes the process in more detail.

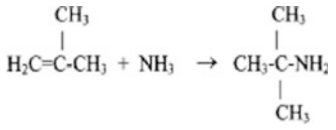
Other Alkyl Amines

Ritter Reaction. Mitsubishi Rayon and Sumitomo Chemical in Japan make *t*-butylamine (22.57). Average yields are about 85% based on isobutylene. Rohm and Haas in the United States make *t*-octylamine from hydrogen cyanide and diisobutylene by a similar route [142].



$$H_2C{=}\overset{\overset{\displaystyle CH_3}{|}}{C}{-}CH_3 + HCN \xrightarrow[H2O]{H^+} H\overset{\overset{\displaystyle O}{\|}}{C}{-}N{-}\underset{\underset{\displaystyle CH_3}{|}}{\overset{\overset{\displaystyle CH_3}{|}}{C}}{-}CH_3 \xrightarrow{OH^-} CH_3{-}\underset{\underset{\displaystyle CH_3}{|}}{\overset{\overset{\displaystyle CH_3}{|}}{C}}{-}NH_2 \qquad (22.57)$$

Direct Amination. Direct amination of olefins is a special case and has been successfully developed for only a few amines. BASF has developed a process for the direct amination of isobutylene to yield *t*-butylamine solutions (see (22.58)) and has operated a commercial plant in Western Europe since 1993 [142].



$$H_2C{=}\overset{\overset{\displaystyle CH_3}{|}}{C}{-}CH_3 + NH_3 \rightarrow CH_3{-}\underset{\underset{\displaystyle CH_3}{|}}{\overset{\overset{\displaystyle CH_3}{|}}{C}}{-}NH_2 \qquad (22.58)$$

Hydrogenation of Anilines. A manufacturing process that is specific for cyclohexylamines is the catalytic hydrogenation of anilines or phenols in the presence of ammonia. Process details are given in [142].

Manufacturing Costs. Manufacturing costs for alkylamines depend on the specific amine and the process. Reference [142] gives the estimated production costs for two different processes.

Production

Methylamines are commonly sold as anhydrous gases or as 25, 40, 50, and 60% aqueous. The other alkylamines are

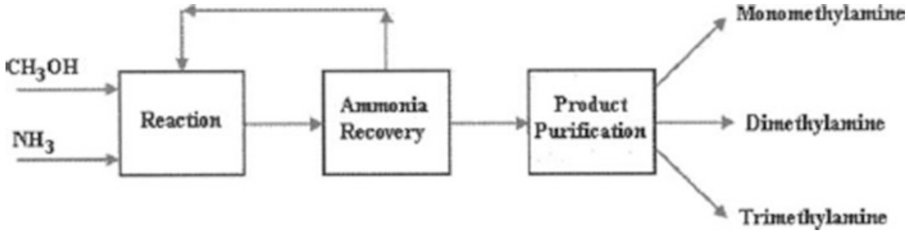


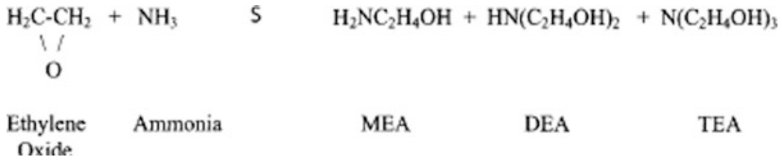
Fig. 22.24 Typical methylamines process

Table 22.25 Methylamine production: raw material and utility use (ton per ton of product leaving the reactor)

	MMA	DMA	TMA
Methanol	1.053	1.451	1.660
Ammonia	0.559	0.385	0.294
Steam	3.0		
Cooling water, m^3	150		
Electricity, kWh	150		

Source: Reproduced by permission of *Hydrocarbon Processing*

usually sold on a 100% basis although some are also marketed as aqueous solutions. In 1997, MMA and TMA prices increased to \$0.73 per lb, and DMA went up to \$0.77 per lb. Since that time, market prices have declined somewhat as methanol prices have receded [142, 144]. From 1995 to 2003, the price of cyclohexylamine has varied from a low of \$0.70 per lb to a high of \$1.35 per lb. High benzene prices have increased the price of aniline and are putting pressure on cyclohexylamine prices [145].



$$H_2C\text{-}CH_2 + NH_3 \quad \text{S} \quad H_2NC_2H_4OH + HN(C_2H_4OH)_2 + N(C_2H_4OH)_3 \tag{22.59}$$

Methylamine growth should remain strong as applications in the largest market segments are growing two to three times quicker than GNP [146]. However, worldwide growth in alkylamines should average a modest 2–3% annually from 2005 to 2009 [142]. From 1998 to 2003, cyclohexylamine production has been constant at 15.5 million lb/year [145].

In the United States, growth is expected to increase at 3.3% per year between 2003 and 2006 [144].

The major world producers of methylamines include BASF, Air Products, DuPont, and UCB. The major producers of C_2 to C_6 alkylamines include Air Products, BASF, and Celanese [142].

The largest consumer of alkylamines are the United States, followed by Western Europe (at ~70% of the United States consumption) and Japan (at ~25% of the United States consumption) [142].

Uses

The three methylamines are important intermediates for the manufacture of solvents, insecticides, herbicides, pharmaceuticals, and detergents. DMA is the most important due to its use in the manufacture of *N*,*N*-dimethyl-formamide and *N*,*N*-dimethylacetamide, which find wide application as solvents [44, 142].

Methyl methacrylate (MMA) ranks second in terms of demand. It is mainly used to make dimethyl urea and *N*-methylpyrrolidone, as well as for methyltaurine, which is employed in CO_2 washes or as a raw material for detergents [44, 142].

TMA plays only a minor role. It is used in the manufacture of choline chloride, quaternary salts, and epichlorohydrin [44, 142].

Some of the other alkyl amines and their uses are

- Ethylamines: Pesticides and catalysts
- Propylamines: Herbicides, detergents, catalysts
- Butylamines: Pesticides, plasticizers, insecticides, catalysts, herbicides
- Cyclohexylamines: Corrosion inhibitors, catalysts, antibiotics, herbicides, cyclamate sweeteners in Asia

Ethanolamines and Secondary Products

The main use of ethanolamines (EOA) is for the manufacture of detergents by reaction with fatty acids. The principal secondary products of EOAs are morpholine, ethylenimine, and ethylenediamine (EDA).

EOAs are corrosive and will attack some plastics and rubbers. They are a moderate fire hazard. MEA and TEA are clear viscous liquids with a mild ammonia-like odor. DEA is crystalline or a viscous liquid. All are soluble in

Table 22.26 Effect of excess ammonia on ethanolamine selectivity [44]

	Selectivity ratios		
Molar ratio (NH_3 to EO)	Mono-	Di-	Tri-ethanolamine
10:1	75	21	4
1:1	12	23	65

Source: Copyright © Wiley-VLN and reprinted by permission of the copyright owner

water and ethanol. Vapor is irritating to the eyes, skin, and respiratory tract, and depression of the central nervous system can occur. They can also be absorbed by the skin in toxic amounts [147].

Ethanolamine Process

Ethylene oxide reacts exothermically with 20–30% aqueous ammonia at 60–150°C and 30–150 bar in a tubular reactor to form the three possible EOAs (monoethanolamine: MEA, diethanolamine: DEA, and triethanolamine: TEA) with high selectivity.

The composition of the reaction product stream can be influenced by temperature and pressure. Higher ratios of excess ammonia lead to a higher MEA content as shown in Table 22.26. The high triethanolamine content that results from an equimolar mixture of reactants indicates that the primary reaction with ammonia is slower than the secondary reactions [44].

A typical EOA flowsheet is shown in Fig. 22.25 [143]. Nippon Shokubai has developed technology that uses a zeolite catalyst that suppresses the formation of TEA and produces more MEA and DEA [147]. Another process flowsheet is described in [148] along with detailed process conditions.

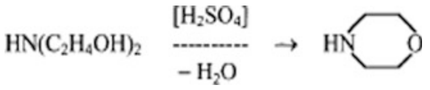
Himtek Engineering in St. Petersburg, Russia, offers EOA technology that operates at 40–70°C and 15–35 atm. The ammonia and part of the MEA are recycled. Three vacuum stills are used to separate the three products. This technology is described in Russian patents No. 2,141,475 and No. 2,063,955 [149].

Production. All EOA plants make a mixture of products. The most important is MEA, which accounts for about half of total production. It is followed by DEA with 30–35% of production. In 2001, Ineos was the largest global producer followed by BASF and Huntsman. EOA demand is growing between 4 and 6% annually with MEA and DEA growth rates substantially exceeding that for TEA. In 2001, EOA supply exceeded demand, and the oversupply situation continued through 2002 with startups of more new plants. About 50% of EOA capacity is in the Americas, more than 30% is in Europe, and the balance is in the Pacific Rim and the Middle East. Table 22.27 summarizes world consumption and growth rates for EOA [35, 147].

Uses. EOA production is consumed in the following areas: detergents (32%), ethyleneamines (20%), gas purification (17%), metal cleaning (10%), herbicide intermediates (10%), textiles (5%), and miscellaneous (6%) [150]. The specific uses of MEA, DEA, and TEA are listed in [151].

Secondary Products of Ethanolamine

Morpholine. Morpholine is a solvent and intermediate for optical brighteners and rubber chemicals. In one process, it is obtained from DEA by dehydration with 70% H_2SO_4 to close the ring:

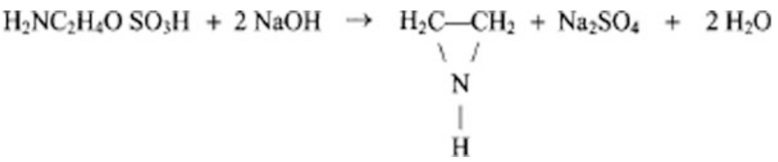


(22.60)

In the United States, a newer production route is used in which diethylene glycol, ammonia, and hydrogen react at 150–400°C and 30–400 bar over a catalyst containing Ni, Cu, Cr, or Co to give morpholine [44].

Ethylenimine. Ethylenimine can be made from MEA via a two-step process that BASF and Hoechst developed [44]:

$$H_2NC_2H_4OH + H_2SO_4 \rightarrow H_2NC_2H_4OSO_3H + H_2O \qquad (22.61)$$



$$H_2NC_2H_4O\,SO_3H + 2\,NaOH \rightarrow H_2C{-}CH_2\,(NH) + Na_2SO_4 + 2\,H_2O \qquad (22.62)$$

Dow makes ethylenimine by reacting 1,2-dichloroethane (EDC) with ammonia in the presence of CaO at about 100°C:

Fig. 22.25 Typical ethanolamine (EOA) process

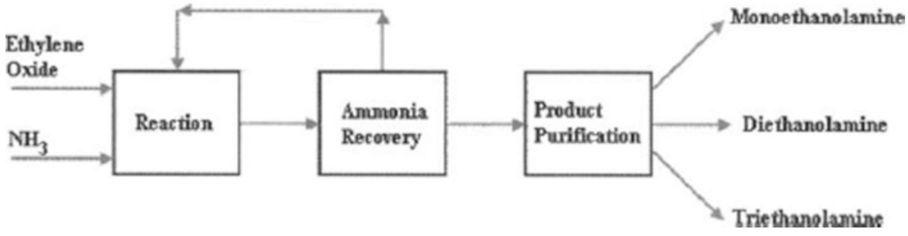
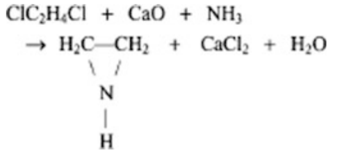


Table 22.27 World consumption of ethanolamines and growth rates [147] consumption (thousands of ton)

Region	1995	2000	2005
Americas	279	358	418
Western Europe	260	329	385
Asia	99	157	197
Other	27	60	90
Totals	*665*	*904*	*1,090*
Growth rates (%)			
Americas	–	5.1	3.1
Western Europe	–	4.8	3.2
Asia	–	9.7	4.6
Other	–	17.3	8.4
Totals	–	*6.3*	*3.8*



$$ClC_2H_4Cl + CaO + NH_3 \rightarrow H_2C{-}CH_2(NH) + CaCl_2 + H_2O \qquad (22.63)$$

Most ethylenimine is converted into polyethylenimine, a processing aid in the paper industry. It also serves as a reaction intermediate when ureas are formed with isocyanates [44].

Ethylenediamine. Commercial EDA is produced by two different routes:

- From EDC and ammonia by Cl/NH_2 substitution
- From MEA and ammonia by OH/NH_2 substitution

The reaction in the first process is shown below, and the process details are given in [44]:

$$\begin{aligned} ClC_2H_4Cl + 2NH_3 \rightarrow\; & NH_2C_2H_4NH_2 \\ & + \text{Higher amines} + NH_4Cl \\ & + \text{Amine clorides} \end{aligned} \qquad (22.64)$$

In 1969, BASF developed a process to make EDA by the reaction of MEA and ammonia [44]:

$$NH_2C_2H_4OH + NH_3 \xrightarrow[H_2]{Ni} NH_2C_2H_4NH_2 + H_2O \qquad (22.65)$$

The properties of EDA and some of the higher amines are given in Table 22.28.

Ethylenediamine Production

In 1999, the EDA capacity in the United States was 390 million lb/year of ethyleneamines. Demand was expected to grow from 310 million lb/year in 1998 to 370 million lb/year in 2003. These estimates include exports that average 70–80 million lb/year and imports that total about 10–12 million lb/year. From 1989 to 1998, growth averaged 4% per year, but it is expected to slow to 3% per year through 2003 [152]. Prices are kept confidential, but the 2001 US EDA price is estimated to be \$0.80–\$0.85 per lb [145].

Ethylenediamine Uses

EDAs are used in the following products: lube oil and fuel additives (20%), chelating agents (16%), wet-strength resins (15%), epoxy curing agents and polyamides (13%), surfactants (10%), oil field chemicals (7%), and miscellaneous (19%).

Hexamethylenetetramine (Hexamine)

Hexamine was first prepared in 1859. It is a white crystalline powder with a slight amine odor. It is soluble in water, alcohol, and chloroform, but it is insoluble in ether. However, the aqueous solutions exhibit inverse solubility, i.e., less hexamine dissolves as the temperature increases. The hydrate, $(CH_2)_6N_4{\cdot}6H_2O$, can be crystallized from the aqueous solution at temperatures below 14°C. Additional properties are listed in Table 22.29.

Hexamine Processes

Hexamine is made by the liquid-phase reaction of ammonia and formaldehyde:

$$\begin{gathered} 4NH_3 + 6HCHO \rightarrow (CH_2)_6N_4 + H_2O \\ \Delta H = -55 \text{ kcal/mol Hexamine} \end{gathered} \qquad (22.66)$$

Process details are given in [153, 154].

Production. Production has declined from 48 million lb in 1995 to 32 million lb in 2001. The decline is mainly due to a

Table 22.28 Properties of ethylenediamine and higher amines

	EDA	DETA	TETA	TEPA	PEHA
Molecular weight	60.11	103.17	146.24	189.31	232.36
Melting point, °C	11–39	−35	−30	−26	–
Boiling point, °C	117	207	277	decomp @ 340°C	
Density, lb/gal	7.67	7.96	8.16	8.33	8.39
Flash point (O.C.), °C	38 102	143	185	185	–
Specific gravity	0.898	–	–	–	–
pH	11.9 (25% solution in water)	–	–	–	–
Vapor density	2.07	–	–	–	–
Nitrogen content, wt.%	46.6	–	–	–	–
Viscosity, mPa s	2.0 (at 20°C)	–	–	–	–
Autoignition temperature, °C	385	–	–	–	–
Physical state	Colorless liquid with amine-like odor	–	–	–	–

EDA Ethylenediamine; *DETA* diethylenetriamine; *TETA* triethylenetetramine; *TEPA* tetraethylene pentamine; *PEHA* penta-ethylenehexamine

Table 22.29 Properties of hexamine/hexamethylenetetramine

Molecular weight	140.19
Sublimation temperature	285–295°C
Flash point	250°C
Density of solid (@ 20°C)	1.33 g/cm^3
Bulk density	700–800 g/L
Particle size	700 μm, maximum
Specific heat	36.5 cal/°C
Heat of formation (@ 25°C)	28.8 kcal/mol
Heat of combustion (@ 25°C)	1,003 kcal/mol
Solubility in water	
20°C	874 g/L
25°C	867 g/L
60°C	844 g/L
pH of 10% aqueous solution	8–9
Vapor pressure (@ 20°C)	0.0035 mbar

decline in the use of phenolic resins that contain hexamine [154]. Hexamine is produced as a granular and free-flowing powder as well as a 42.5% solution. The solution is shipped in tank trucks, railcars, and drums. Solids are packed in bags, fiber drums, and super sacks. Hexamine is sensitive to moisture. Therefore, it should be stored in an atmosphere with a relative humidity below 60% [155].

Uses. Hexamine is used as a stabilizer and surface treatment in numerous applications and industries. Specific applications are described in [156].

Hydrazine

Hydrazine, $NH_2–NH_2$, is the simplest diamine. Anhydrous hydrazine was first prepared in 1894. In 1953, Arch Hydrazine (formerly Olin Chemical) pioneered US production of hydrazine to fuel the first Titan rocket.

Hydrazine is a clear, hygroscopic fuming liquid with the odor of ammonia. It is a mild base that is miscible in polar solvents such as water, ammonia, amines, and alcohols. Additional properties are shown in Table 22.30 [157].

Processes

Five processes are available for the commercial production of hydrazine: the Raschig process, the Raschig/Olin process, the Hoffmann (urea) process, Bayer ketazine process, and the peroxide process from Produits Chimiques Ugine Kuhlmann (of France).

Raschig Process. The Raschig process was discovered in 1907 and then modified into the Olin process. The chemical reactions take place in the liquid phase and involve three steps:

$$NaOH + Cl_2 \rightarrow NaOCl + HCl \tag{22.67}$$

$$NH_3 + NaOCl \rightarrow NH_2Cl + NaOH \tag{22.68}$$

$$NH_2Cl + NH_3 + NaOH \rightarrow NH_2 - NH_2 \cdot H_2O + NaCl \tag{22.69}$$

Hydrazine is produced in the hydrated form with 1 mol of water added. Although a significant fraction of hydrazine is used as the hydrate, numerous applications (such as rocket propulsion) require anhydrous hydrazine. Because of the azeotrope at 68% hydrazine, reactive distillation or extractive distillation must be used to produce pure hydrazine.

Additional process details are given in [158]. The Raschig process is shown in Fig. 22.26.

The Raschig process can also be used to react amines with chloramine to make mono-substituted or unsymmetrical disubstituted hydrazines:

$$NH_2Cl + RNH_2 \rightarrow RNHNH_2 \tag{22.70}$$

Table 22.30 Physical properties of hydrazine

Molecular weight	
Anhydrous	32
Hydrate	50
Boiling point, °C	113.5
Melting point, °C	1.4
Specific gravity, g/mL	
0°C	1.025
15°C	1.014
25°C	1.004
50°C	0.982
Critical temperature, °C	380
Critical pressure, atm	145
Vapor pressure, mmHg	
25°C	14
31°C	20
36°C	100
Viscosity, cP	
5°C	1.2
25°C	0.9
Heat of vaporization, kcal/mol	9.6
Heat of solution (@ 25°C), kcal/mol	−3.9
Heat capacity (@ 25°C), J/mol K	98.87
Heat of combustion, kcal/mol	−146.6
Heat of formation, kcal/mol	
Liquid	12
Gas	23
Flash point, °C	52
Explosive limits in air by vol.%	4.7–100
Azeotrope	
Boiling point (68% hydrazine), °C	120.5

Hydrazine concentration (wt.%) 132						
	100	64	51.2	35.2	22.4	15.4
Melting point, °C	2.0	−51.7	−59.8	−64.6	−26	−14
Boiling point, °C	113.5	120.5	117.2	108	107	103
Density (25°C), g/mL	1.0045	1.0320	1.0281	1.0209	1.0132	1.0083
Viscosity (20°C), μPa s	0.974	1.5	1.44	1.10	1.08	1.04
pH	–	12.75	12.10	–	–	10.5

$$NH_2Cl + R_2NH \rightarrow R_2NNH_2 \qquad (22.71)$$

Raschig/Olin Process. The Raschig/Olin process is used to make anhydrous hydrazine. In this process, (22.67), (22.68), and (22.69) also occur. However, the refining area has a different design from the Raschig process. Additional details are given in [158].

The refining area has five steps: ammonia removal, NaCl concentration and removal, hydrazine–water distillation to reach azeotropic concentration, extractive distillation with aniline to break the azeotrope and distillation to separate hydrazine from aniline. Sometimes 50% caustic replaces aniline in the extractive distillation. If ultra-pure hydrazine is needed, freeze crystallization is used to remove the supernatant fluid. Assays between 99.5 and 99.99% have been achieved.

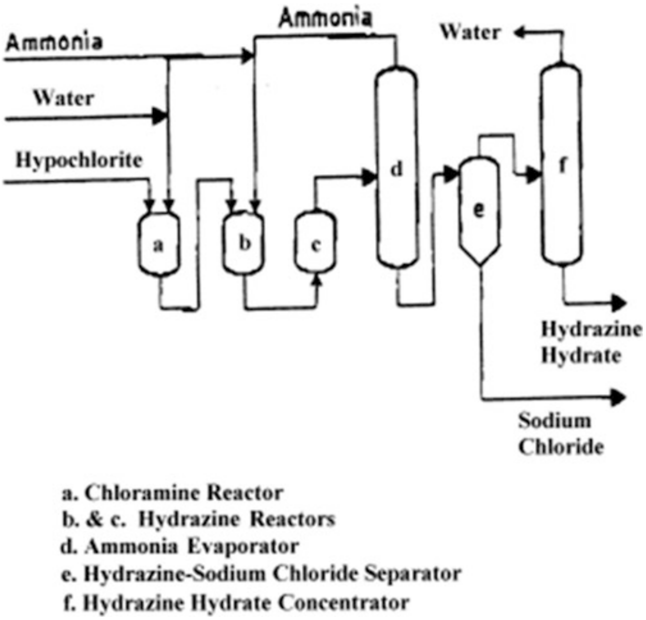


Fig. 22.26 Raschig process for hydrazine production (Courtesy of Wiley-VCH [158])

The overall yield based on chlorine is 65%. The combined yield for (22.67) and (22.68) is about 95%, whereas the yield for (22.69) is 70%. Japanese patent SHO-62-83308 claims to increase hydrazine yield by almost 10% by high-frequency heating immediately after the reactants are mixed in (22.68). A sketch of the Raschig/Olin process is shown in Fig. 22.27 [158].

Hoffman (Urea) Process. The overall reaction for the Hoffman process is

$$\underset{\text{Urea}}{CO(NH_2)_2} + \underset{\text{Hypochlorite}}{NaOCl} + 2NaOH \rightarrow \underset{\text{Hydrazine}}{NH_2 - NH_2} + NaCl + Na_2CO_3 + H_2O \qquad (22.72)$$

Process details are given in [158].

This process was not being operated in 2001 although it has operated commercially in the past. Compared with the standard Raschig process, it was the most economical method for low production levels. However, rapid growth in plant size made it obsolete [158].

Bayer Ketazine Process. The Bayer ketazine process is based on the reaction of chloramine with ammonia in the presence of acetone at pH 12–14. NaOCl, acetone, and a

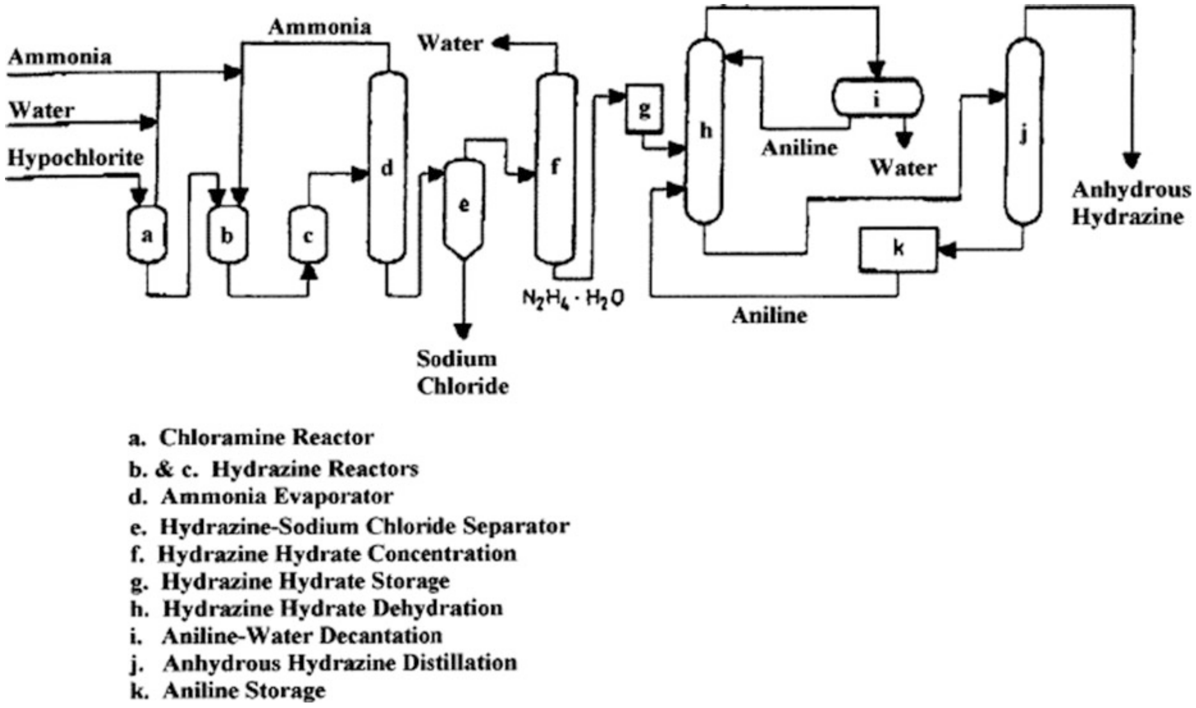


Fig. 22.27 Raschig/Olin process for hydrazine production (Courtesy of Wiley-VCH [158])

20% aqueous solution of ammonia (at a mole ratio of 1:2:20, respectively) are fed to a reactor at 35°C and 200 kPa to make the aqueous dimethyl ketazine solution. Excess ammonia and acetone are removed in a series of columns and recycled to the reactor. The ketazine solution is distilled to make a hydrazine hydrate containing 64% hydrazine [158].

Peroxide Process. This process (see Fig. 22.28) was invented by PCUK and is operated by AtoFina in France. The reaction is carried out in the presence of methyl ethyl ketone (MEK) at atmospheric pressure and 50°C. The molar ratio in the feed of hydrogen peroxide (H_2O_2), MEK, and NH_3 is 1:2:4. Additional process details are given in [158].

Production. Hydrazine is marketed as anhydrous hydrazine, as the monohydrate (64% hydrazine), and as a 54.5% aqueous solution (85% hydrazine monohydrate). The largest US producers of hydrazine based on a 2000 report [159] are Arch Chemical (23 million lb/year), Bayer (17 million lb/year), and Fairmont Chemical (1 million lb/year).

About 29 million lb is sold in the United States, and the manufacturers retain about 7 million lb for internal use. The space industry uses only 5% of all hydrazine produced in the United States [159]. Total capacity in the western world in 1988 was estimated to be about 130 million lb/year [158].

The permissible exposure limit (PEL) for hydrazine hydrate on an 8-h TWA is 0.1 ppm. The ammonia-like

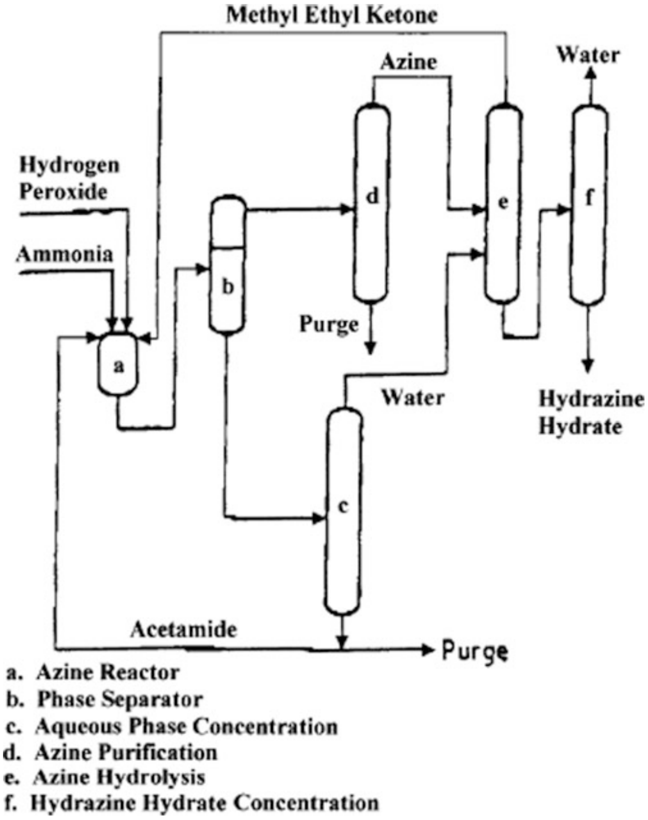


Fig. 22.28 Peroxide process for hydrazine production [158] (Courtesy of Wiley-VCH [158])

odor of hydrazine hydrate normally cannot be detected until the concentration is above 3–5 ppm [157].

Uses. Hydrazine is used directly as an oxygen scavenger and as an energy source. Hydrazine is consumed in the following applications: blowing agents (33%), pesticides (32%), water treatment (18%), and miscellaneous (17%) [160].

Hydrogen Cyanide

HCN is a colorless liquid. However, some iron cyanides are blue (prussian blue), and their presence as a slight contaminant usually gives HCN a bluish tint that led to the German name Blausäure—blue acid. Selected physical properties of hydrogen cyanide are shown in Table 22.31.

Safety

Many people can detect hydrogen cyanide by odor or taste sensation at a concentration of 1 ppm in air, whereas most people can detect 5 ppm. OSHA has set 4.7 ppm as the maximum, average safe exposure limit for a 15-min period. Exposure to 20 ppm of HCN in air causes slight warning symptoms after several hours; 50 ppm causes disturbances within an hour; 100 ppm is dangerous for exposures of 30–60 min; and 300 ppm can be rapidly fatal unless prompt, effective first aid is administered. A small concentration of cyanide always exists in a person's body, and the body has a mechanism to continuously remove small amounts of cyanide [161].

HCN may also present some processing hazards. HCN undergoes an exothermic polymerization at a pH of 5–11 to form solid black compounds. This polymerization can become explosively violent, especially if it is confined. To avoid polymerization, stored HCN should contain less than 1.0 wt.% water; it should be kept cool; and it should be inhibited with sulfuric, phosphoric, or acetic acid [161].

Explosively violent hydrolysis can occur if an excess of a strong acid (H_2SO_4, HNO_3, or HCl) is added to HCN. The reaction is fastest at or near stoichiometric ratios, e.g., 1–2 mol H_2SO_4 per mol HCN, and can cause severe equipment damage if confined [161].

An important environmental issue is the fate of cyanide. HCN, if spilled, evaporates quite readily. That which does not evaporate is soon decomposed or rendered nonhazardous by complexing with iron in the soil, by biological oxidation, or by degradation from photolysis [161]. Several methods for treatment and disposal of cyanide compounds are described in [161, 162].

Table 22.31 Physical properties of hydrogen cyanide

Molecular formula	HC≡N
Molecular weight	27.03
Melting point	−13.24°C
Triple point	−13.32°C
Boiling point, 1 atm	25.70°C
Density, g/mL	
0°C	0.7150
10°C	0.7017
20°C	0.6884
Specific gravity of aqueous solutions (@ 18°C compared to water @ 18°C)	
10.04%	0.9838
20.29%	0.9578
60.23%	0.8290
Vapor specific gravity (@ 31°C, air = 1)	0.947
Vapor pressure, kPa	
−29.5°C	6.697
0.0°C	35.24
27.2°C	107.6
Liquid viscosity (@ 20.2°C)	0.2014
Specific heat, J/mol	
−33.1°C, liquid	58.36
16.0°C, liquid	70.88
27.0°C, gas	36.03
Heat of formation (@ 25°C, 1 atm), ΔH_f, kJ/mol	
Gas @ 25°C	−130.5
Liquid @ 25°C	−105.4
Heat of fusion @ −14°C, kJ/mol	7,100
Latent heat, kJ/mol	23.8
Heat of combustion, net, kJ/mol	642
Conductivity, S/cm	3.3×10^{-6}
Heat of vaporization, kJ/mol	25.2
Heat of polymerization, kJ/mol	42.7
Flash cup, closed cup, °C	−17.8
Explosive limits in air at 100 kPa and 20°C (vol.%)	6–41
Autoignition temperature, °C	538

Manufacture

Six processes are presently of economic significance: the Andrussow process is currently the principal HCN manufacturing process in the world, the BMA process is practiced by two companies and provides high yield and selectivity by using a complex reaction system, the Fluohmic process is of interest in locations where electricity is inexpensive, the formamide process is useful for sites with inexpensive carbon monoxide, the BP (British Petroleum) ACRN process produces HCN as a by-product, and the methanol process.

Andrussow Process. The Andrussow process was patented in Germany by Dr. L. Andrussow in 1933. Its main advantages are low converter investment, low maintenance costs, and

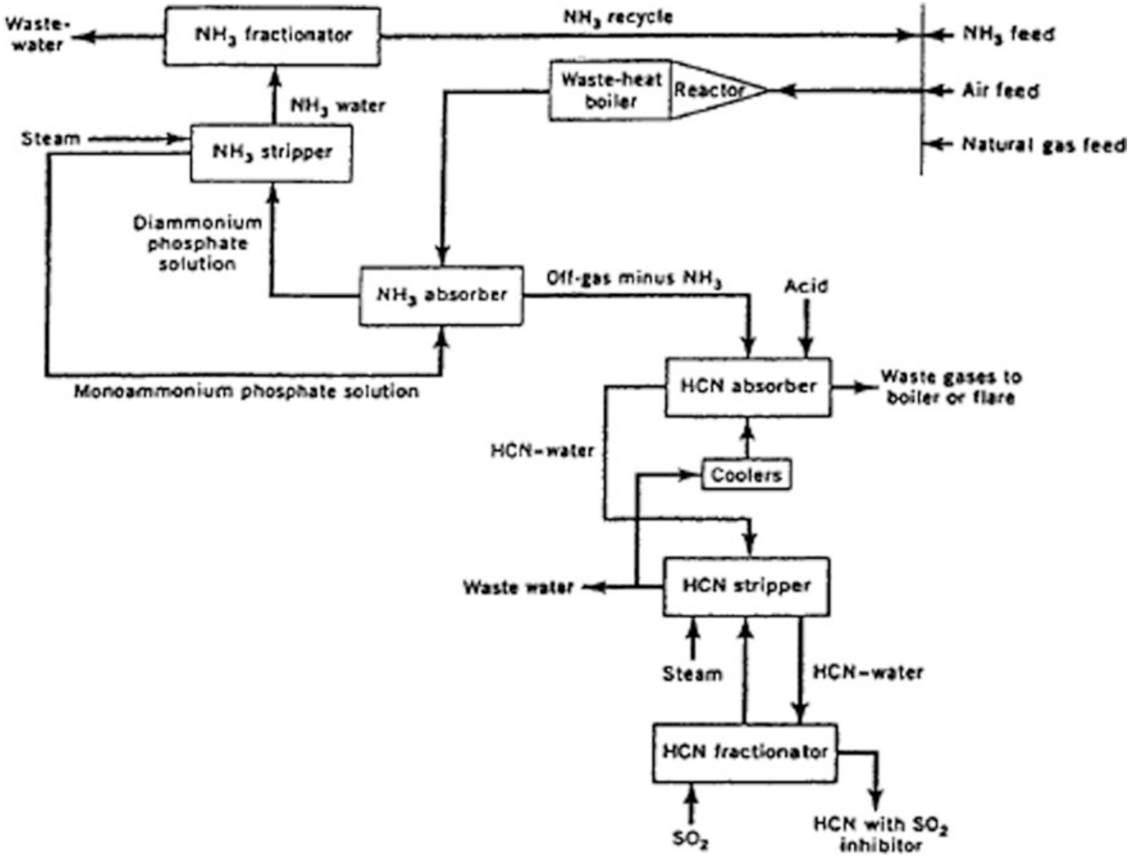


Fig. 22.29 Andrussow HCN process with ammonia removal [161] (Copyright by Wiley and reproduced by permission of copyright owner)

high natural gas yields. The Andrussow process produces HCN by the reaction of ammonia, air, and natural gas at 1,000–1,200°C in the presence of a platinum/rhodium catalyst. The reaction is

$$NH_3 + CH_4 + 1.5O_2[Pt/Rh]1,200°C HCN + 3H_2O \quad \Delta H = -115.2 \text{ kcal/mol HCN} \tag{22.73}$$

More reaction details are given in [161]. The heat of reaction is recovered in the waste heat boiler to generate about 5 lb of steam per lb of HCN produced.

The precious metal catalyst is usually 90% platinum and 10% rhodium in gauze form. The catalyst life can range from 4,000 h to as much as 10,000 h. A catalyst pack is usually replaced because it has been contaminated by minor impurities in the feed gases. Very little catalyst metal is consumed or lost during the life of a catalyst.

The reaction takes place under fuel-rich conditions to maintain a nonflammable feed mixture. Typical feed composition is 13–15% ammonia, 11–13% methane, and 72–76% air on a volumetric basis. Control of feed composition is essential to guard against deflagrations as well as to maximize the yield. The yield from methane is approximately 60% of theoretical.

The reactions in the Andrussow process are more complex than that shown in (22.72) [163]. Most of the heat required for HCN formation is supplied by combustion of methane. This results in an overall reaction that is exothermic [161]. The converter off-gas typically has the following molar composition: HCN = 6–12%, NH_3 = 1.5–3%, H_2 = 7–11%, N_2 = 50–55%, H_2O = 20–25%, CO = 3–4%, CO_2 = 0.3%, and CH_4 = 0.1%.

Two processes can be used to remove NH_3 from the converter off-gases. Figure 22.29 shows a process that recovers unreacted ammonia for recycle to the converter, and the process in Fig. 22.30 uses sulfuric acid to produce ammonium sulfate as a by-product. Ammonium sulfate can create a disposal problem, but the recycle system is capital and energy intensive. Ammonia must be removed from the off-gas before HCN can be recovered because the ammonia promotes polymerization of the HCN [161]. Additional process details are given in [161].

BMA Process. The BMA (Blausäure–Methan–Ammoniak, or HCN–methane–ammonia) process was developed by Degussa around 1949. In the process, ammonia reacts with

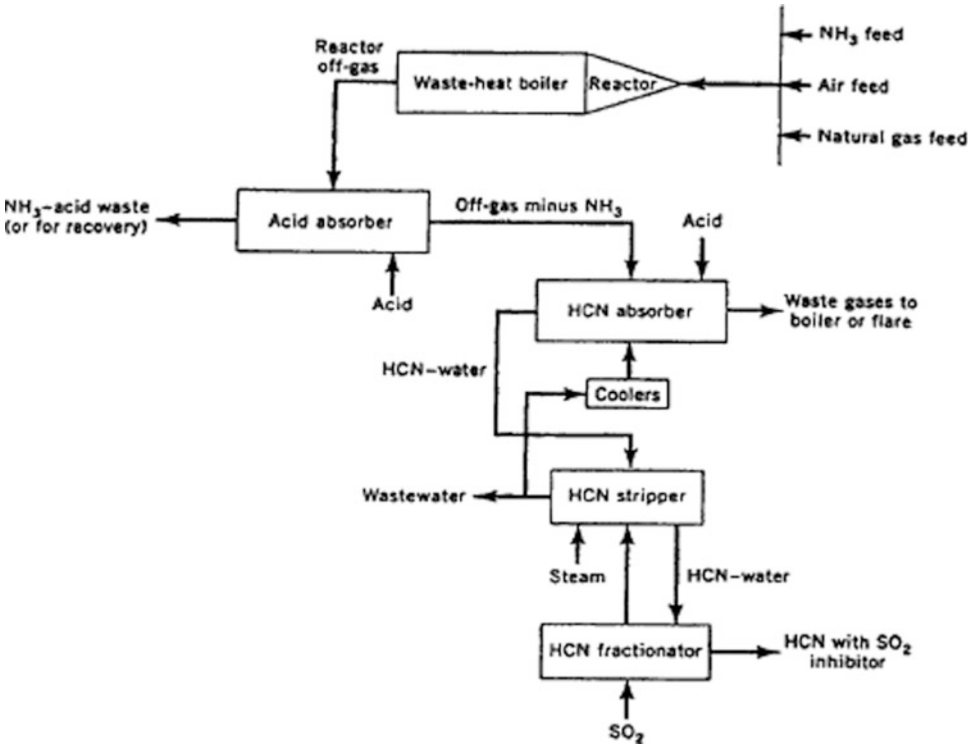


Fig. 22.30 Andrussow HCN process with ammonia removal [161] (Copyright by Wiley and reproduced by permission of copyright owner)

methane in the absence of oxygen in a gas-fired tubular reactor. US patent 5,785, 942 describes the BMA process and improvements that were made during the mid-1990s. The reaction is [44]

$$NH_3 + CH_4 \xrightarrow[\text{[Pt/Ru]}]{1{,}300^\circ C} HCN + 3H_2$$
$$\Delta H = +60 \text{ kcal/mol HCN} \qquad (22.74)$$

In a reactor that is similar to a reformer, the reaction occurs in tubes that are heated externally to supply the endothermic heat of reaction [161]. Sintered corundum (α-Al_2O_3) tubes with an internal layer of platinum/ruthenium catalyst are normally used. To achieve adequate heat transfer, the tubes may be only 3/4 in. in diameter and 6 $^1/2$ ft long. Selectivities of 90–91% for methane and 83–84% for ammonia are reached at 1,200–1,300°C reaction temperatures.

A typical converter is made up of multiple furnaces, each of which contains eight to ten reactors. Each reactor contains 10–30 tubes with a catalyst lining. Catalyst life is approximately 10,000 h [161].

The converter off-gas typically has the following molar composition: $HCN = 20$–23%, $NH_3 = 2$–3%, $H_2 = 70$–72%, $N_2 = 0.5$–1%, and $CH_4 = 1$–2.4%.

The unreacted ammonia is removed from the reactor off-gas by scrubbing with sulfuric acid to make ammonium sulfate. Just as with the Andrussow process, ammonia must be removed from the off-gas before HCN can be recovered because the ammonia promotes polymerization of the HCN. After ammonia is removed from the converter off-gases, the remaining gas stream is processed in a way similar to the Andrussow process [161].

In the BMA process, the gas that leaves the HCN absorber has the following molar composition: $H_2 = 96\%$, N_2 and $O_2 = 1.5\%$, and $CH_4 = 2.5\%$. The vent gas is used either as a chemical feedstock or as fuel gas in the HCN furnaces.

The advantages of the BMA process are high ammonia yields, high natural gas yields, and the useful hydrogen in the vent gas. But the high investment and high maintenance costs for the converter are the disadvantages. The complexity of the reaction system also makes the process less attractive for large-scale production [161]. In 2003, the BMA process was in use at four plants [164].

Fluohmic Process. The Fluohmic process was developed around 1960 by Shawinigan Chemicals of Canada. This process involves the reaction of ammonia with hydrocarbons (usually propane or butane) in an electrically heated, fluidized bed of coke. The reaction is

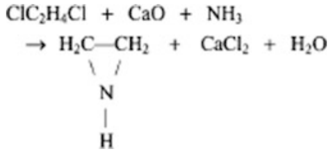
$$3NH_3 + C_3H_8 \xrightarrow{1{,}370^\circ C} 3HCN + 7H_2$$
$$\Delta H = +151 \text{ kcal/mol HCN} \qquad (22.75)$$

The fluidized bed is heated at 1,350–1,650°C by passing an electric current between carbon electrodes immersed in the bed. The performance of the reaction system is reported to be

	Conversion (mol. %)	Yields	
		mol.%	lb/lb HCN
C_3H_8	88	88–90	0.60–0.62
NH_3	86	86–90	0.70–0.73

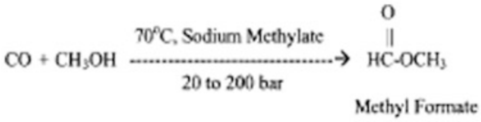
The reactor off-gas typically has the following molar composition: HCN = 25%, NH_3 = 0.25%, H_2 = 72%, and N_2 = 3%. The HCN is separated from the off-gas for refining, and the hydrogen leaves the system as a by-product together with a small amount of nitrogen and unreacted ammonia. The high electrical consumption (~3 kWh/lb HCN) makes the process attractive only in locations where the cost of electricity is extremely low. In 2003, the Fluohmic process was used in Spain, Australia, and South Africa.

Formamide Process. The formamide process was developed by several companies, but in 2001, it was only used by BASF. The process involves the following steps [44]: (1) methyl formate synthesis, (2) formamide synthesis, and (3) HCN formation/dehydration.



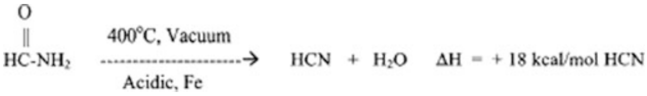
(22.76)

With excess methanol, the conversion to methyl formate is 95% with nearly 100% selectivity.



$$CO + CH_3OH \xrightarrow[\text{20 to 200 bar}]{\text{70°C, Sodium Methylate}} \underset{\text{Methyl Formate}}{\overset{\overset{O}{\|}}{HC}\text{-}OCH_3} \tag{22.77}$$

The conversion of ammonia is 60–85%



$$\overset{\overset{O}{\|}}{HC}\text{-}NH_2 \xrightarrow[\text{Acidic, Fe}]{\text{400°C, Vacuum}} HCN + H_2O \quad \Delta H = +18 \text{ kcal/mol HCN} \tag{22.78}$$

The overall performance of the reaction system is

	Yield	Consumption
	mol.%	lb/lb HCN
CH_3OH	–	0.20–0.30
NH_3	79–90	0.70–0.80
CO	52–94	1.10–2.00

The reactor off-gas contains 60–70% HCN.

BASF has obtained European patent 1,110,913, which describes HCN production from gaseous formamide in a fluidized bed. Elements of the commercialized formamide process are included in a MMA process that Mitsubishi Gas Chemical started up in 1997. The formamide is obtained by reacting hydroxyisobutyramide with methyl formate. Therefore, the external requirements for HCN are reduced to the replacement of yield losses [164]. The process is described in US patent 6,075,162, and the catalyst is described in European patent 1,086,744.

Methanol Process. HCN is produced when methanol reacts with ammonia and oxygen in the presence of an oxide catalyst that contains iron, antimony, phosphorous, and vanadium. The reaction occurs in the vapor phase in a fluidized-bed reactor. The process and the catalyst are described in patents that were issued to Nitto Chemical (now Mitsubishi Rayon) during the late 1990s (European patent 864,532; Japanese patents 10-167,721, 10-251,012, 11-043,323; US patent 5,976,482).

In 2003, Mitsubishi Rayon is believed to operate a commercial-scale, methanol-based HCN process and has offered to license the technology to other companies. The technology might provide a low-cost way to convert an ACRN plant to HCN-only production. However, methanol is a higher-cost source of carbon compared to natural gas, so the methanol process probably has a higher operating cost than the Andrussow process.

ACRN Process. The Standard Oil Company of Ohio (now BP Amoco) commercialized in 1960 a fluidized-bed process in which the catalytic oxidation of a mixture of propylene and ammonia produces ACRN. INEOS purchased the ACRN business from BP in 2005 by-products, from this reaction are HCN and acetonitrile. The yields of HCN depend on the process conditions and on the catalyst system [164]. The reactions are

$$\underset{\text{Propylene}}{CH_2 = CHCH_3} + NH_3 + \tfrac{3}{2}O_2 \rightarrow \underset{\text{Acrylonitrile}}{CH_2 = CHCN} + 3H_2O \tag{22.79}$$

$$\underset{\text{Propylene}}{CH_2 = CHCH_3} + \tfrac{3}{2}NH_3 + \tfrac{3}{2}O_2 \rightarrow \tfrac{3}{2}\underset{\text{Acetonitrile}}{CH_3CN} + 3H_2O \tag{22.80}$$

$$CH_2 = CHCH_3 + 3NH_3 + 3O_3 \rightarrow 3HCN + 6H_2O \tag{22.81}$$

Depending on the type of ACRN process and the operating conditions, about 10–24 lb HCN is obtained per 100 lb of ACRN [44]. US patent 5,840,648 describes a catalyst that permits more HCN production.

During the 1990s, several patents were issued (European patent 878,464, US patent 5,288,473, and US patent 6,204,407) in relation to the addition of methanol to the feed stream of an ACRN plant to increase HCN production.

BP and Asahi are developing processes for the ammoxidation of propane to produce ACRN. This process is believed to yield a lower level of HCN than the optimized oxidation of propylene [164]. In 2007, Asahi started up a propane process in Tongsuh, South Korea.

Other HCN Processes. Some alternate HCN processes that have been studied in the 1990s were

- Production of HCN from acetonitrile (see Japanese patent 10,167,721)
- Production of HCN by use of a corona (German patent 1,054,982)
- Production of HCN by use of microwave heating (US patents 5,393,393; 5,470,541; and 5,529,669)4.

None of these processes are believed to be used in commercial production in 2004.

Production

In 1999, a total of 34 companies were operating 47 HCN production facilities in the United States, Western Europe, and Japan. Capacity in these three regions was about 3.6 billion pounds (1.6 million tons). Direct production accounts for 75% of total capacity, and the balance is coproduct from ACRN production. Worldwide capacity is about 1.8 million tons [164].

Between 1994 and 1999, HCN consumption increased about 2.2% per year. Consumption is expected to increase at around 2.8% per year from 1999 to 2004 [164]. In the United States, production is expected to grow at a rate of 2.4% per year from 1.7 billion pounds in 2003 to 1.87 billion pounds in 2007 [165].

Uses

HCN is usually consumed at its production site, and some consumers without on-site HCN production capability are believed to be developing plans to eliminate HCN shipments. For example, BASF in the United Kingdom produces by-product HCN in an ACRN plant and ships HCN railcars to Ineos. This HCN is used to produce acetone cyanohydrin (ACH) in a MMA plant. To eliminate these shipments, BASF and Ineos plan to (1) install an ACH plant on the BASF plant site and (2) transfer the ACH to Ineos via a 9-km pipeline [166].

ACRN plants are highly dependent on the actions of the acrylic fiber industry, the major outlet for ACRN. As a result, HCN consumers may have to operate at reduced rates if an adequate supply of HCN is not being produced by the ACRN plant. In the early years of the new century, many consumers of HCN from ACRN plants evaluated options that would give them a more consistent supply of HCN. Some ACRN plants were shut down during this time, and market conditions plus company finances will determine whether any of these plants will be restarted.

Some of the uses of HCN are

- Adiponitrile (ADN). ADN is made when HCN reacts with butadiene. It is a key ingredient in production of nylon.
- Acetone cyanohydrin (ACH). ACH is made when HCN reacts with acetone. It is a key ingredient in many MMA plants. MMA is used to make acrylic sheeting and related products.
- Sodium cyanide (NaCN). NaCN is made when HCN reacts with sodium hydroxide (NaOH). NaCN is used in precious metals (gold and silver) extraction, metal plating, and some chemical production processes.
- Methionine. Methionine is an essential amino acid that is used in animal feed. The key ingredients are HCN or NaCN, methyl mercaptan, acrolein, and NH_4HCO_3.
- Cyanuric chloride (CYC). CYC is made when HCN reacts with chlorine and is then rearranged into a ring compound. It is used to make herbicides, pharmaceuticals, and explosives.
- Chelating agents. Chelating agents are made when HCN reacts with formaldehyde and amines to make aminocarboxylic acids, which are then saponified to produce the acid salts. Their major uses are in soaps, water treating agents, and various cleaning agents.
- Other uses include ferrocyanides (for blue jeans), acrylates, lactic acid, pharmaceuticals, and specialty chemicals.

In 2004, HCN was consumed in the following applications: ADN (for nylon 6/6): 50%, ACH (for MMA): 28%, methionine: 7%, sodium cyanide: 6%, cyanuric chloride: 3%, chelating agents: 2%, and miscellaneous (including nitrilotriacetic acid and salts): 4% [165].

Table 22.32 lists the HCN consumption for different products.

Aniline

Aniline was first produced in 1826 by the dry distillation of indigo, the oldest known vat dye. Fritsche also obtained aniline from indigo by heating it with potash (K_2O). Hofmann obtained aniline by reduction of nitrobenzene in 1843. Aniline is a colorless, oily, flammable liquid that is slightly soluble in cold water and infinitely soluble in alcohol and ether. Its physical properties are summarized in Table 22.33.

Processes

The aniline processes that are currently in use include (1) hydrogenation of nitrobenzene, (2) nitrobenzene reduction

Table 22.32 Consumption of hydrogen cyanide by major region, 1999 (thousands of ton)

Product	United States	Western Europe	Japan	Total
Adiponitrile	340	127	–	467
Acetone cyanohydrin	206	152	50	408
Sodium cyanide	72	55	16	143
Methionine and analogues	37	32	4	73
Cyanuric chloride	20	20	1	41
Chelating agents	18	11	–	29
Other	35	13	13	61
Totals	*728*	*410*	*84*	*1,222*

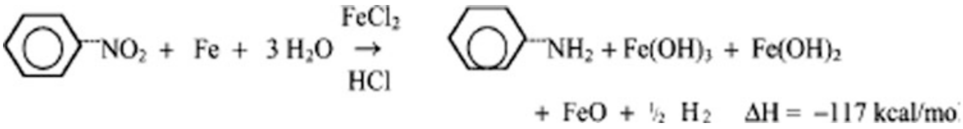
with iron filings, and (3) ammonolysis of phenol. Almost 97% of the nitrobenzene produced in the United States is converted to aniline between the hydrogenation process and the iron filings process [167].

Hydrogenation of Nitrobenzene. The primary aniline production process in the world is the hydrogenation of nitrobenzene. BASF, DuPont, ChemFirst, and Rubicon use this process in the United States. This technology is also used by all Western European aniline producers and all but one Japanese aniline producer.

This process would have been replaced much earlier by more economical reduction methods if it had not been possible to obtain valuable iron oxide pigments from the iron oxide sludge. However, the increasing demand for aniline has far surpassed the market for iron pigments, so this process is no longer preferred [44].

Ammonolysis of Phenol. Aniline can also be produced when phenol is subjected to gas-phase ammonolysis at 200 bar and 425°C.

This is the Halcon/Scientific Design process. The chemistry is

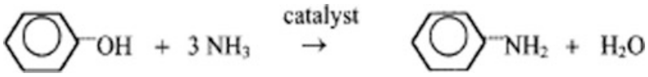


(22.82)

The catalytic gas-phase hydrogenation processes for nitrobenzene can be carried out using a fixed-bed or a fluidized-bed reactor. Details of both processes are given in [44]. A typical process diagram is shown in Fig. 22.31.

Nitrobenzene Reduction with Iron Filings. The older Bechamp method for iron oxide pigment production gives aniline as a coproduct and is operated by Bayer in West Virginia. Nitrobenzene is reduced by reaction with iron filings in the presence of a hydrochloric acid catalyst. The iron is oxidized to the ferrous or ferric state, and the coproduct aniline is separated. The yield is 90–95% of theoretical. The reactions are represented as follows [168]:

$$2\,Fe(OH)_3 + Fe(OH)_2 \rightarrow Fe_3O_4 + H_2O \quad (22.83)$$



(22.85)

The catalysts are $Al_2O_3{\cdot}SiO_2$ (possibly as zeolites) and oxide mixtures of Mg, B, Al, and Ti. These can be combined with additional co-catalysts such as Ce, V, or W. With a large excess of ammonia, the selectivity to aniline is 87–90% at a phenol conversion of 98%. The by-products are diphenylamine and carbazole. This technology is used at one plant in Ohio and at another plant in Japan. The economics of this process are favorable if low-cost phenol is available and high-purity aniline is desired. Capital costs are low because benzene nitration is avoided. A typical process sketch along with a material balance is shown in Fig. 22.32.

(22.84)

Table 22.33 Physical properties of aniline

Property	Value
Molecular weight	93.12
Boiling point, °C	
101.3 kPa (760 mmHg)	184.4
4.4 kPa (33 mmHg)	92
1.2 kPa (9 mmHg)	71
Melting point, °C	−6.15
Density, *d*	
At 20/4°C	1.02173
At 20/20°C	1.022
Viscosity at 20°C, mPa s (= cP)	4.423–4.435
Dissociation constant, *pK*	
At 20°C	4.60
At 60°C	8.88
Enthalpy of dissociation, kJ/mol (kcal/mol)	21.7 (5.19)
Heat of combustion, kJ/mol (kcal/mol)	3,389.72 (810.55)
Specific heat, 20–25°C	0.518
Latent heat of vaporization, J/g (cal/g)	476.3 (113.9)
Flash point (closed-cup), °C	76
Autoignition temperature, °C	615
Flammable limits in air (% by volume)	
LEL	1.3
UEL	11.0
Vapor density (air = 1)	3.22
Solubility (g in 100 g water @ 20°C)	3.5
Odor threshold, ppm	1.1

Uses. MDI (4,4′-methylenebis phenylisocyanate) accounted for almost 85% of the worldwide demand for aniline in 2000. MDI is used primarily to make rigid polyurethane foam and polyurethane elastomers. MDI growth is expected to be 6.0–8.0% per year during the first 10 years of the new century as its use continues to increase in the construction industry (the largest user of rigid polyurethane foam) and the auto industry (the largest user of reaction–injection molding plastics) [168, 169].

Other aniline uses and the percent of worldwide aniline production that goes into these uses are rubber-processing chemicals (9%), dyes and pigments (2%), agricultural chemicals (3%), specialty fibers (1%), and miscellaneous, such as explosives, epoxy curing agents, and pharmaceuticals (1%) [168, 169].

Production. MDI has been the driving force behind the recovery of the aniline business since 1982 when the industry had a capacity utilization rate of less than 50%. By 1996, capacity utilization had approached 95% in some regions [168].

Aniline production is concentrated in the United States, Western Europe, and Japan where aniline capacity is over 1.8 million tons/year [168]. Aniline's global production capacity in 1999 was 2.9 million tons, and demand in 2000 was estimated to be 2.68 million tons/year [170, 171].

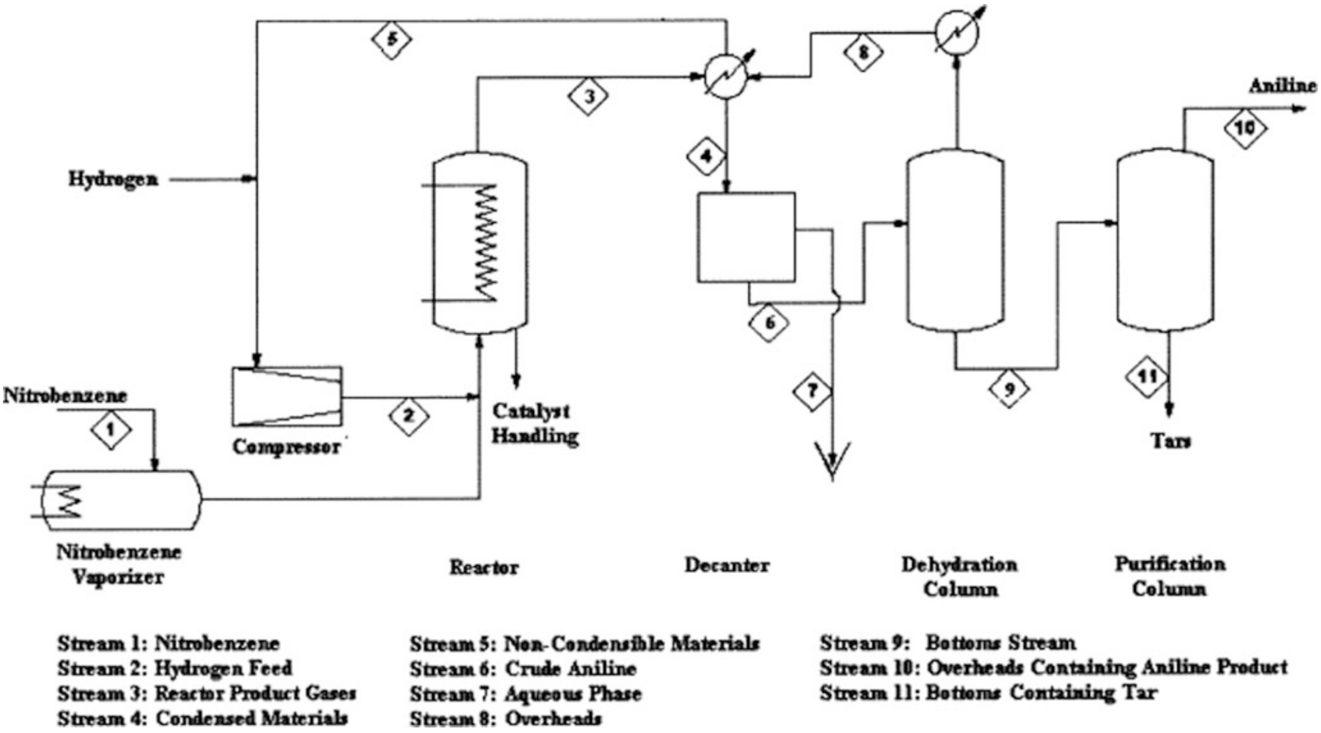
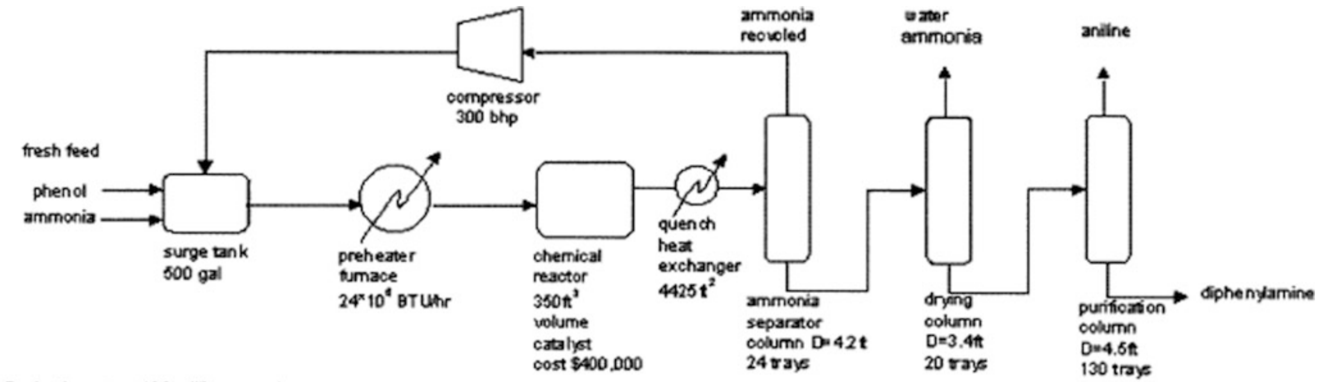


Fig. 22.31 Aniline production process: hydrogenation of nitrobenzene



Production rate = 100 million pounds per year

Utilities:

Fuel: $24*10^6$ BTU/hr

Steam: 600psig - $10.2*10^6$ BTU/hr 150psig - $10*10^6$ BTU/hr

Cooling water: $9*10^6$ BTU/hr for ΔT=20°F

Electricity to drive the motor for the 300 bhp compressor

Preliminary Material Balance on Process

mass flow rates (lb/hr)	fresh feed	ammonia recycle	reactor feed	reactor product	ammonia separator column overhead	bottoms	drying column overhead	bottoms	purification column overhead	bottoms
phenol	13,159	0	13,159	13	0	13	0	13	13	0
ammonia	4,199	43,070	47,269	44,900	43,070	1,830	1,830	0	0	0
aniline	0	10	10	12,890	10	12,880	315	12,565	12,565	0
water	0	11	11	2,528	11	2,517	2,505	12	12	0
diphenylamine	0	0	0	118	0	118	0	118	0	118
total	17,358	43,091	60,449	60,449	43,091	17,358	4,650	12,708	12,590	118
	60449				60449		17358		12708	

			preheater	quench exchanger						
temperature (°F)	100	142	118 670	673 110	112	248	141	289	278	401
pressure (psig)	260	260	260 260	240 230	220	220	6	6	5	5

Fig. 22.32 Aniline production process: ammonolysis of phenol

Aniline pricing tends to track benzene pricing. In general the aniline price will move 1.2 cents/lb for each 10 cents/gal movement in benzene price. Between 1995 and 2001, the aniline sales prices varied from US$ 0.35 per lb to US$ 0.40 per lb [169].

The OSHA PEL for aniline is 5 ppm for an 8-h TWA.

Other Compounds

Several other nitrogen compounds are commercially important, including

- Hexamethylenediamine which is used primarily in the manufacture of nylon
- Dimethylformamide, a versatile solvent for organic and inorganic compounds and an important reaction medium for ionic and nonionic compounds
- Dimethylacetamide, an important industrial solvent for polyacrylonitrile, vinyl resins, cellulose derivatives, styrene polymers, and linear polyesters
- Isocyanates, important materials in the production of foams, resins, and rubbers
- Calcium cyanamide, which is used as a fertilizer, herbicide, insecticide, a steel-making additive, and an ore-processing material. It can also be used to make thiourea, guanidine, and ferrocyanides [172]
- Methyl methacrylate, which is used in plastic glass, resins, and paint [173–175]
- Acrylamide, which is used in water treatment chemicals, paper, and resins [176–178]
- Caprolactam, which is used to make nylon 6 [179–181]
- Acrylonitrile, which is used in many different acrylic fibers [182, 183]

References

1. Mills DL to Maxwell GR, 10 July 2001, BP nitrogen fatality, personal communication
2. Hopkins BS, Smith HR et al (1949) Chemistry and you. Lyons & Carnahan, Chicago

3. Comyns AE (1999) Named processes in chemical technology, 2nd edn. CRC Press, Boca Raton
4. Suzuki M (1998) In: LeVan MD (ed) Adsorption and ion exchange: fundamentals and applications. American Institute of Chemical Engineers, New York, p 120
5. Suzuki M (1998) In: LeVan MD (ed) Adsorption and ion exchange: fundamentals and applications. American Institute of Chemical Engineers, New York, pp 121–122
6. Eur Chem News 53(1391):31 (1989)
7. Chem Mark Rep (1990)
8. Haber F, van Oordt G (1905) Z Anorg Allg Chem 43:111
9. Harding AJ (1959) Ammonia manufacture and uses. Oxford University Press, London
10. Haber LF (1971) The chemical industry 1900–1930. Clarendon, Oxford, 198
11. Vancini CA (1971) Synthesis of ammonia (trans: Pirt L). Macmillan, Basingstoke
12. Jennings JR, Ward SA (1989) In: Twigg MV (ed) Catalyst handbook, 2nd edn. Wolfe Publishing, London, p 384
13. Travis T (1993) Chem Ind (Lond) (15):581
14. Hill WH (1945) In: Lowry HH (ed) Chemistry of coal utilization, vol 2. Wiley, New York, p 1029
15. Cohen JM (1956) The life of Ludwig Mond. Methuen, London, 176
16. Nitrogen (223):25 (1996)
17. Chem Eng Prog 48:468 (1952)
18. Shearon WH Jr, Thompson HL (1954) Modern chemical processes, vol 3. Reinhold, New York, 16
19. Scholvien WR (1931) Chem Met Eng 38(2):82
20. Spitz PH (1988) Petrochemicals: the rise of an industry. Wiley, New York, 84
21. Isalski WH (1989) Separation of gases. Clarendon, Oxford, 155
22. Livingston JG, Pinto A (1983) Chem Eng Prog 79(5):62
23. Chem Eng (Rugby, Eng):21 (1990)
24. Hydrocarbon Process 70(3):132–136 (1991)
25. Chem Eng (NY) 96(7):43 (1989)
26. Chem Eng (Rugby, Eng) (471):21 (1990)
27. Chem Week:13 (1991)
28. Eur Chem News 58(1524):42 (1992)
29. Eur Chem News 60(1592):27 (1993)
30. Oil Gas J 94(47):37 (1996)
31. SRI International (1998) Chemical economics handbook, Nitrogen Industry overview
32. http://www.fertilizer.org/STATSIND. International Fertilizer Industry Association, June 2001. http://www.fertilizer.org
33. http://www.ifdc.org/PDF_Files/FreeFert.statistics.PDF. International Fertilizer Development Center (IFDC), Muscle Shoals, Jan 2004
34. Chemical Marketing Reporter (29 Nov 1999)
35. Chemical Marketing Reporter (18 June 2001)
36. SRI International (2001) Chemical economics handbook, ammonia
37. Ammonia prices hold steady as demand remains firm. Chemical Marketing Reporter (27 Sept 2004)
38. http://www.efma.org/Publications. Production of Ammonia, European Fertilizer Manufacturers Association, July 2001
39. SRI International (1998) Chemical economics handbook, Hydrogen product review
40. Hydrogen, Airco Industrial Gases, Airco, Inc., 1991
41. Alves J (1999) Analysis and design of refinery hydrogen distribution systems. PhD thesis, UMIST, Manchester
42. Hallale N, Liu F (2000) Efficient refinery hydrogen management for clean fuels production. Presented at the AIChE annual meeting, Los Angeles, 2000
43. Hallale N (2001) Burning bright trends in process integration. Chem Eng Prog 97(7):30–41
44. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH Publishers, New York
45. http://www.synetix.com. Hydrogen Plant Products, Imperial Chemical Industries PLC, 14 Aug 2001
46. Foreman JM (1990) Pre-reformer aids syngas units. Hydrocarbon Process 69(12):34B–34D
47. US Patent 3,827,987
48. Nair MPS (2001) Control corrosion factors in ammonia and urea plants. Hydrocarbon Process 80(1):85–93
49. Muradov N (2000) Thermocatalytic CO_2-free production of hydrogen from hydrocarbon fuels. In: Proceedings of the 2000 U.S. DOE hydrogen review, 9 May 2000
50. Simbeck D (2001) Coal—bridge to the hydrogen economy. Presentation at 18th annual international Pittsburgh Coal Conference in Newcastle. http://www.engr.pitt.edu/pcc/2001plenary/simbeckcoaltoh2revised.pdf. Dec 2001
51. Morris P (2000) Clean fuels means hydrogen—where will it come from and what will it cost? European Fuels Week, 10–14 Apr 2000. http://www.airproducts.com. Accessed 12 Sept 2002
52. Buchanan TL, Schoff RL et al (2002) Hydrogen production facilities plant performance and cost comparisons. Final Report, Prepared for United States Department of Energy, National Energy Technology Laboratory, Parsons Infrastructure and Technology Group Inc, Reading, Mar 2002. http://www.netl.doe.gov/coal power/gasification/pubs/pdf/FinalCompReport.pdf
53. Knoblauch K (1978) Chem Eng (NY) 85(25):87
54. http://www.halliburton.com/news/archive/2001. Halliburton KBR signs contract with Liaotong for ammonia plant revamp. Press Release, Dallas, 26 Nov 2001
55. http://www.haldor.dk. Synthesis Gas Technologies, Haldor Topsøe A/S, 13 Aug 2001
56. Fleshman JD (2001) Cost efficient revamps in hydrogen plants. In: Proceedings of NPRA annual meeting, New Orleans, Mar 2001
57. Farnell PW (1999) Synetix's Advanced Gas Heated Reformer. In: AIChE ammonia safety symposium, Seattle, Sept 1999
58. Fleshman JD, McEvoy LJ et al (1999) New hydrogen plant design achieves low cost and high efficiency. In: Proceedings of NPRA annual meeting, San Antonio, 1999
59. Wolff D (2001) Producing hydrogen on-site. Chem Eng Prog 97 (3):51–53
60. American Petroleum Institute. Selected values of properties of hydrocarbons. Research Project 44, National Bureau of Standards
61. Liebner W, Ulber D (2000) MPG—Lurgi Multi Purpose Gasification: application in "gas-gasification". In: 2000 Gasification technologies conference, San Francisco, Oct 2000. http://www.gasification.org/98GTC/GTC00340.pdf
62. http://www.shell.com. Generating energy using gasification, Royal Dutch/Shell Group of Companies, 13 Aug 2001
63. Higman C, Eppinger M (1994) The zero-residue refiner. ACHEMA, Frankfurt, June 1994. http://www.higman.de/gasification/achema94.pdf
64. Hydrocarbon Process 75(4):149 (1996)
65. http://www.mbendi.co.za. Mbendi Information Services (Pty) Ltd., South Africa, July 2001
66. http://www.dakota.com. The coal gasification process, Dakota Gasification Company, 16 July 2001
67. Breckenridge W, Holiday A et al (2000) Use of SELEXOL® process in coke gasification to ammonia project. In: Laurance Reid gas conditioning conference, 2000 conference proceedings, University of Oklahoma, Norman, OK, Mar 2000. http://www.uop.com/gasprocessing/TechPapers/Selexol/cokegasifiAmmonia.pdf
68. Camps JA (2000) Nitrogen fertilizer producers push the boundaries. Jacobs Consultancy. http://www.jacobssirrine.com/aichepush.pdf

69. Rentech buys ammonia plant; Plans conversion to coal. Chem Week 166(33):35 (2004)
70. http://www.hydrocarbononlline.com. Hydrocarbon Online News, Vert Tech LLC, 16 Feb 1999
71. http://www.hydrocarbononlline.com. Hydrocarbon Online News, Vert Tech LLC, 17 July 2001
72. Tebo M (2001) Honda opens first Los Angeles-area hydrogen fuel station. PRNNewswire, American Honda Motor Co., Press Release, 10 July 2001
73. Gronich S (2000) DOE hydrogen program overview. In: Proceedings of the 2000 U.S. DOE hydrogen review, 9 May 2000
74. http://www.itd.anl.gov/highlights/9-2/ceramic.html. Argonne National Laboratory, 13 Aug 2001
75. Carolan MF, Chen CM, Rynders SW (2002) ITM syngas and ITM H2: engineering development of ceramic membrane reactor systems for converting natural gas to hydrogen and synthesis gas for liquid transportation fuels. In: Proceedings of the 2002 U.S. DOE Hydrogen Program Review, NREL/CP-610-32405, Washington, DC, 2002. http://www.eren.doe.gov/hydrogen/pdfs/32405a19.pdf
76. http://www.kvaerner.com. Compact Reformer Technology, Kvaerner process technology, 27 Aug 2001
77. Noble T (2001) Microreactors: smaller can be better. Chem Eng Prog 97(7):10–11
78. Bakemeier H, Huberich T et al (1985) Ullmann's encyclopedia of industrial chemistry, vol A2, 5th edn. Reinhold, New York
79. http://www.hydrogen.org/Knowledge/w-I-energiew-eng3.html. HyWeb: Knowledge—Hydrogen in the Energy Sector, Sept 2002
80. Pressure swing and vacuum swing adsorption systems. Air Products and Chemicals, Inc., 1996–2002. http://www.airproducts.com/membranes
81. Lombard JF (1969) Hydrocarbon Process 48(8):111
82. Lundberg WC (1979) Chem Eng Prog (6):81
83. Carstensen JH et al (1990) New developments in shift catalysts for ammonia. Hydrocarbon Process 3(69):57–62
84. Thompson D (1999) New advances in gold catalysis part II. Gold Bull 32(1):12–19
85. Perry RH, Chilton CH (eds) (1973) *Chemical engineers' handbook*, 5th edn. McGraw-Hill, New York
86. http://www.sud-chemie.com. Purification of hydrocarbon streams. Sud-Chemie AG, 11 Sept 2002
87. Chem Eng Prog 69(2):67–70 (1973)
88. Chem Eng 86(13):88–89 (1979)
89. Hydrocarbon Process 57(4):145–151 (1978)
90. Chem Eng 67(19):166–169 (1960)
91. Amine guard systems in hydrogen production. Presented at the AIChE, 86th national meeting, Houston, 1–5 Apr 1979
92. Amine guard, wet solvent systems. Gas Processing Technologies, UOP LLC, Des Plaines, 2000. http://www.uop.com
93. Bajusz IG, Goodwin JG Jr (1997) Hydrogen and temperature effects on the coverages and activities of surface intermediates during methanation on Ru/SiO2. J Catal 169:157–165
94. Hoost TE (1991) Chemical modification of supported-metal catalysis. PhD dissertation, University of Pittsburgh, Pittsburgh
95. http://www.halliburton.com/KBR/KBRTCH. Kellogg Brown & Root—a Halliburton Company, 22 Aug 2001
96. http://www.hydrocarbononlline.com. Telepneff M (2001) Hyprotech and Synetix integrate catalytic reactor models. Hydrocarbon Online News, Vert Tech LLC, 10 Aug 2001
97. Rhodes AK (1996) New ammonia process, catalyst proven in Canadian plant. Oil Gas J 96(47):37–41
98. http://www.haldor.dk. Ammonia synthesis, Haldor Topsøe A/S, 22 Aug 2001
99. Linde Ammonia Concept (LAC) Plant, Linde AG, Hoellriegelskreuth, Germany, 2001. http://www.linde-process-engineering.com
100. Linde Ammonia Concept (LAC)—Ammonia from hydrocarbons. Hydrocarbon Process 80(3):78
101. Synetix ammonia synthesis catalysts technical manual. Cleveland, 2002. http://www.synetix.com/ammonia/pdfs/216w.pdf
102. Ammonia Casale technologies, Ammonia Casale S.A., Lugano, Switzerland, 2003. http://www.casale.ch
103. Ohri IJ (2001) Energy conservation measures in ammonia plants at IFFCO Kalol and Phulpur Units, IFFCO, India, 2001. http://www.fertilizer.org/ifa/publicat/pdf/tech0039.pdf
104. Filippi E, Pinauda L (2001) Casale offers more ammonia revamp options. Fertilizer Int:27
105. Lippmann D, Deutsch S, Khalil R, Moussa M (1999) The New Abu Qir III ammonia/urea complex—design and start-up experiences. Paper No. S2, 44th annual ammonia safety symposium, Seattle, 27–30 Dec 1999
106. Agarwal P (2002) Ammonia: the next step. http://www.cheresources.com
107. Parkinson G (2002) All aboard, ammonia. Chem Eng, Nov 2002. http://www.che.com/mag/ce3.htm
108. Christensen PV (2001) Design and operation of large capacity ammonia plants. In: 4th conference for development and integration of petrochemical industries in Arab States—Bahrain, Haldor Topsøe, Lyngby, May 2001. http://www.haldortopsoe.com
109. Larsen J, Lippmann D, Hooper CW (2001) A new process for large-capacity ammonia plants. Nitrogen Methanol 253:41–46
110. Davey WLE, Filippi E (2003) MEGAMMONIA—the mega-ammonia process for the new century, Nitrogen 2003, Warsaw, 26 Feb 2003
111. http://www.potashcorp.com/npk_markets/industry_overview/. Industry overview–nitrogen. Potash Corporation, 25 July 2002
112. SRI International (2001) Chemical economics handbook, nitric acid
113. Encarta.msn.com (2001) Ostwald, Wilhelm. Microsoft Corporation
114. http://www.efma.org/Publications. Production of nitric acid. European Fertilizer Manufacturers Association, Sept 2001
115. Sauchelli. Fertilizer nitrogen. ACS monograph series 161
116. Catalytic processes in nitric acid manufacture. In: Fertilizer Society proceedings, London, 1978
117. http://www.enviro-chem.com. Nitric acid plant technology, Monsanto Enviro-Chem, 2001
118. Maurer R, Bartsch U (2001) Enhanced plant design for the production of azeotropic nitric acid. Prepared for presentation at Heraeus nitric acid conference, Johannesburg, 29–30 Nov 2001
119. http://www.chemtics.kvaerner.com. Nitric acid concentration, Kvaerner Pulp & Paper, 2001
120. Kirk-Othmer (1997) Encyclopedia of chemical technology, vol 17. Wiley, New York, pp 80–107
121. Powell R (1969) Nitric acid technology—recent developments. Noyes Development Corp, Park Ridge
122. http://www.efma.org/Publications. Production of ammonium nitrate, European Fertilizer Manufacturers Association, Sept 2001
123. Zapp KH (2000) Ammonium compounds. In: Ullmann's encyclopedia of industrial chemistry. Wiley-VCH, Weinheim. http://www.mrw.interscience.wiley.com/ueic/articles/a02_243
124. http://www.agcentral.com/imcdemo. Production of major nitrogen fertilizers, IMC-Agrico Company, 2001
125. Chem Week 101(2):75 (1967)
126. Erben A (1999) From neutralisation to granulation. Nitrogen Methanol 241:49–52
127. Bovens JL, van Hecke F (1992) The Carnit ammonium nitrate process, Kemira SA, The International Fertiliser Society, Proceeding 320. http://www.fertiliser-society.org/Proceedings/Prc320.HTM
128. Ammonium nitrate, Chemical products synopsis, Mannsville Chemical Products Corp, Adams, Oct 2001

129. http://www.sasol.com/smx. Ammonium nitrate, Sasol SMX, 1998
130. Kirk-Othmer (1996) Encyclopedia of chemical technology, supplement. Wiley, New York, pp 597–621
131. Meessen JH, Petersen H et al (1996) Ullmann's encyclopedia of industrial chemistry, vol A27, 5th edn. Reinhold, New York
132. http://www.efma.org/Publications. Production of urea and urea ammonium nitrate. European Fertilizer Manufacturers Association, Sept 2001
133. http://www.agrium.com. Carseland Nitrogen Operations, Agrium, Canada, 3/01
134. http://www.enviro-chem.com. Urea plant technology, Monsanto Enviro-Chem, 2001
135. SRI International (2001) Chemical economics handbook, urea
136. http://www.potashcorp.com/npk_markets/industry_overview/. Industry overview—nitrogen. Potash Corporation, 2 Nov 2001
137. Urea chemical profile. Chemical Marketing Reporter (2 Dec 2002)
138. SRI International (1999) Chemical economics handbook, melamine
139. DSM expands melamine production in Netherlands. Eur Paint Resin News 38(12):5 (2000)
140. Shan, Cheng-gang, Status of production and demand and market analysis of melamine, Modern Chemical Industry, Author's Abstract, Jan 2002. http://www.xdhg.com.cn/article/2002/abs/e-abs-2002-01.html
141. OMV lifts melamine sales by targeting new markets. Chemical Marketing Reporter (19 Aug 2002)
142. SRI International (2001) Chemical economics handbook, alkylamines
143. http://www.kvaerner.com. Industrial amine process technologies. Kvaerner Process Technology, 27 Aug 2001 (May be under http://www.davyprotech.com as of Nov 2001)
144. Methylamines chemical profile. Chemical Marketing Reporter (4 Aug 2003)
145. Cyclohexylamine chemical profile. Chemical Marketing Reporter (26 July 2004)
146. Chemical Marketing Reporter (24 Apr 2000)
147. Ethanolamines. Eur Chem News 74(1953):16 (2001)
148. http://www.owlnet.rice.edu/~ceng403/ethamsep.html. Garcia CA, Reisdorf M, Palombo L (1997) Separation unit for Ruehl Corp. ethanolamine plant. Rice University, 25 Oct 1997
149. http://www.himtek.ru/english/technology1.shtml. Technology of ethanolamine production. Himtek Engineering, 2000
150. Ethanolamines chemical profile. Chemical Marketing Reporter (11 June 2001)
151. Chemical Marketing Reporter (11 June 2001)
152. Chemical Marketing Reporter (9 Aug 1999)
153. http://www.chemicals-india.com. Hexamine manufacturing process, Kothari Phytochemicals International, 5 Oct 2001
154. Hexamine, Chemical products synopsis, Mannsville Chemical Products Corp, Adams, Oct 2001
155. http://www.wrightcorp.com. Formaldehyde and hexamine, Wright Corporation, 5 Oct 2001
156. http://www.saudiform.com. Hexamethylenetetramine (HMT). Saudi Formaldehyde Chemical Company Ltd., 1999
157. http://www.hydrazinehydrate.com. Hydrazine hydrate: safety and health. Bayer Corporation, 2001
158. Schirmann JP (1996) Ullmann's encyclopedia of industrial chemistry, vol A13. Reinhold, New York, pp 177–190
159. Hypergolic propellant hazard response guide, Cape Canaveral Air Force Station (CCAFS). Draft Volume I, ICF Technology Incorporated, 1 July 1988
160. http://www.hydrazine.com. About Arch Hydrazine. Arch Hydrazine, 2001
161. Kirk-Othmer (1993) Encyclopedia of chemical technology, 4th edn, vol 7. Wiley, New York, pp 753–782
162. Botz M (2001) Cyanide treatment methods. Mining Environ Manage 9(3):28–30
163. Hasenberg D (1984) HCN synthesis on polycrystalline platinum and rhodium. Dissertation, University of Minnesota, Minnesota
164. SRI International (2000) Chemical economics handbook, hydrogen cyanide
165. Hydrogen cyanide chemical profile. Chemical Marketing Reporter (10 Jan 2005)
166. Kvaerner awarded seal alliance contracts by BASF and Ineos. Hydrocarbon Process Int Ed 80(3):41 (2001)
167. http://www.epa.gov/ttn/chief/le/benzene/benz.cfe.pdf. Emissions from major uses of benzene, Section 5E, EPA-454/4-98-011. Locating & Estimating Documents, Technology Transfer Network, United States EPA, pp 5-58–5-61, June 1998
168. SRI International (1997) Chemical economics handbook, aniline
169. Aniline, Chemical products synopsis, Mannsville Chemical Products Corp, Adams, July 2001
170. Production and consumption of aniline in the world. Xian Dai Hya Gong 18(6):34–36 (1998)
171. Markets and economics: aniline. Chem Week 161(43):41 (1999)
172. ntp-server.niehs.nih.gov/htdocs/LT-studies/TR163.html. Bioassay of calcium cyanamide for possible carcinogenicity, 1979
173. Methyl methacrylate. Eur Chem News 77(2029):15 (2002)
174. Methyl methacrylate, Product focus. Chem Week 164(43):32 (2002)
175. The commercialization of a new production process for MMA monomer. Asahi Chemical Industry Co., Ltd. Press Release, 10 Mar 1997. http://www.asahi-kasei.co.jp/asahi/en/news/1996/e970310.html
176. Will R, Toki G (2002) CEH Product Review, Acrylamide, SRI International, Mar 2002
177. Acrylamide, PERP Program—new report alert, Chem Systems. Dec 2002. http://www.chemsystems.com/newsletters
178. Acrylamide chemical profile. Chemical Marketing Reporter (15 Mar 1999)
179. Ritz J, Kieczka H, Moran WC (1996) Caprolactam. In: Ullmann's encyclopedia of industrial chemistry, vol A5. Reinhold, New York, pp 31–50
180. Caprolactam, Chemical products synopsis, Mannsville Chemical Products Corp, Adams, Aug 2001
181. Ring KL (2000) CEH Marketing Research Report, Caprolactam, SRI International, Apr 2000
182. Acrylonitrile. Eur Chem News:16 (2002)
183. Acrylonitrile chemical profile. Chemical Marketing Reporter (19–26 Aug 2001)
184. Abrardo JM, Khurana V (1995) Hydrogen technologies to meet refiners' future needs. Hydrocarbon Process 74(2):43–49
185. Rossini FD and Coworkers (1952) National Bureau of Standards. Circular 500:53
186. Chem Eng News:16 (1979)
187. Hydrocarbon Process 70(3):134 (1991)
188. Terra e Vita (France) 40(22):78–79 (1999)
189. Green DS (2001) Chem Process 74
190. Leseur H, Czernichowski J, Chapelle J (1994) Int J Hydrogen Energy 19(2):139
191. PEP yearbook (1996) Process Economics Program, SRI Consulting
192. http://www.synetix.com. Ammonia plant products, Imperial Chemical Industries PLC, 22 Aug 2001
193. Thompson D (2000) Highlights from recent literature. Gold Bull 33(1):33
194. International critical tables, vol III, pp 304, 305, 309. McGraw-Hill, New York, 1928
195. International critical tables of numerical data, physics, chemistry, and technology, vol IV, p 255. McGraw-Hill, New York, 1928
196. http://www.yet2.com. Nitric acid processes, yet2.com Inc., 2001

197. http://www.che.LSU.edu/courses/4171/computerprojects1.htm. Aniline process flow diagram. Louisiana State University, Sept 2001
198. Aniline chemical profile. Chemical Marketing Reporter (8 Feb 1999)
199. McNulty A (2001) Northern business unit: inert gas lessons learned. Internal Safety Memo, Foinaven Delivery Unit, Northern Business Unit, BP, 29 Mar 2001
200. Bright K (1985) Cyanides in the South African mining industry. Report prepared for DuPont by Business & Marketing Intelligence, June 1985
201. Milmo S (2002) More pressure ahead for Europe's fertilizers. Chemical Market Reporter (5 Aug 2002)
202. Carney MJ (2002) Industrial gases. http://www.uwec.edu/carneymj/Adobe%20slides/Chapter%203%20slides.pdf. Aug 2002
203. http://www.census.gov/econ/www/industry.html. U.S. Census Bureau, Industrial Products Overview, Aug 2002
204. Drnevich FR et al (1999) Production and use of industrial gases for iron and steelmaking, Chapter 5. The AISE Steel Foundation, Pittsburgh. http://www.aise.org/newpubs/ironchap%205.pdf
205. http://www.potashcorp.com/npk_markets/potashcorp_letter/. The PotashCorp letter. Potash Corporation of Saskatchewan Inc., Feb 2002
206. http://www.census.gov/econ/www/industry.html. U.S. Census Bureau, Industrial Statistics, Annual Survey of Manufactures, 11 Feb 2002
207. http://www.epa.gov/ttn/otag. Technology Transfer Network, United States Environmental Protection Agency, 29 Aug 2002
208. Bradford ML, Grover R, Paul P (2002) Controlling NOx emissions—part 2. Chem Eng Prog 4:38–42
209. Chemical commodities. Chemical Market Reporter (22–29 July 2002)
210. http://www.airproducts.com/gases. Gases and equipment, Modes of supply—nitrogen, 21 Aug 2002
211. Hydrogen, Chemical products synopsis, Mannsville Chemical Products Corp, Adams, Aug 2001
212. http://www.ott.doe.gov/facts/archives/fotw205.shtml. Fact #205, Hydrogen cost and worldwide production. Office of Transportation Technologies, 25 Feb 2002
213. http://www.alchemix.net/hyrefining.html. The H2 Technologies economic advantage. H2 Technologies, Alchemix Corporation, Sept 2002
214. Hydrogen chemical profile. Chemical Marketing Reporter (29 Jan 2001). http://www.chemexpo.com/news/profile010129.cfm
215. http://www.chicagobridge.com/howebaker. Hydrogen & synthesis gases. Howe-Bakers Engineers, Ltd., 9 Sept 2002
216. Kot R (2001) Hydrogen attack, detection, assessment and evaluation. In: 10th Asia-Pacific conference on non-destructive testing, Brisbane, Sept 2001. http://www.ndt.net/article/apcndt01/papers/1154/11154.htm
217. http://www.ceteklimited.com. Ceramic coatings for process tubes. Cetek Limited, 18 June 2002
218. Davis RA, Patel NM (2001) Refinery hydrogen management—Part 1 & part 2, Hydrocarbon engineering. Palladian, Public Farnham, England, Aug and Sept 2001
219. http://www.bellona.no/en/energy/hydrogen. Hydrogen technologies. Bellona Report 6:2002, Oct 2002
220. Moulijn JA. Production synthesis gas, Chemical process technology (ST3141) Class notes, 1Q2002. http://www.dct.tudelft.nl/~wwwrace/st3141a/cpt_05.pdf
221. Gross M, Wolff J (2000) Gasification of residue as a source of hydrogen for the refining industry in India. In: 2000 Gasification technologies conference, San Francisco, Oct 2000. http://www.gasification.org/98GTC/GTC00150.pdf
222. Beaudette TM, Bochow C Jr, Slivensky D (2001) Natural gas partial oxidation for chemical processing in Longview, Texas. In: Gasification technologies 2001, San Francisco, Oct 2001. http://www.gasification.org/98GTC/GTC01039.pdf
223. http://www.proquip-corp.com/hydrochem/experience/plant_types.shtml. Hydro-Chem modular hydrogen plants. Hydro-Chem: A Division of Pro-Quip, Holly Springs, Aug 2002
224. Goudarefractories.com/applications_ndustries/petrochem/reforming_and_gasification.asp. GOUDA—applications & industries, Gouda, the Netherlands, Oct 2002
225. Bhandarkar PG (2001) Gasification overview focus—India, Hydrocarbon Asia, pp 46–51, Nov/Dec 2001. http://www.hcasia.safan.com/mag/dec01/t46.pdf
226. Brown JAG, Griffiths J (2002) The gasification of refinery residual oil and subsequent generation of electricity. Institution of Diesel and Gas Turbine Engineers. http://www.jacobsconsultancy.com/idgte%20paper.pdf
227. Tavoulareas ES, Charpentier J-P (1995) Clean coal technologies for developing countries. World Bank Technical Paper No. 286, Energy Series, July 1995. http://www.worldbank.org/html/fpd/em/power/EA/mitigatn/igccsubs.stm
228. Zuideveld PL (2002) Overview of shell gasification projects. Noordwijk, Apr 2002. http://www.icheme.org/literature/conference/gasi/Gasification%20Conf%20Papers/Session%204%20presentationShell.pdf
229. How PEM works, technology introduction. Proton Energy Systems, Wallingford, 2002. http://www.protonenergy.com
230. Praxair starts-up industry's first on-site hydrogen generating system. Praxair, Inc., News releases—1998, Danbury, 12 Mar 1998. http://www.praxair.com/praxair.nsf
231. Brown LC, Besenbruch GE et al (2002) High efficiency generation of hydrogen fuels using nuclear energy, hydrogen and fuel cells annual review, U.S. Department of Energy, May 2002. http://www.eren.doe.gov/hydrogen/pdfs/32405.pdf
232. Brooksbank WJ (2003) HydroMax®: bridge to a hydrogen economy. Alchemix Corporation, Carefree. http://www.alchemix.net
233. Myers DB, Ariff GD et al (2002) Cost and performance comparison of stationary hydrogen fueling appliances. Directed Technologies, Arlington. http://www.eren.doe.gov/hydrogen/pdfs/32405b2.pdf
234. Polymer membrane captures CO2 at high temperatures. Chem Eng Prog 7:27 (2002)
235. Acid gas purification techniques for CO2 (carbon dioxide) benfield process, Air Pollution Control Technology in Japan, Exhaust Gas Treatment Equipment, Kobe Steel, LTD., Tokyo, 2002. nett21.unep.or.jp/CTT_DATA/AIR/AIR_7/html/Air-208.html
236. http://www.linde-process-engineering.com. Rectisol Wash, Linde AG, 2001
237. Selexol Process, Gas Processing, UOP LLC, Des Plaines, 2002. http://www.uop.com
238. Carbon dioxide capture from power stations—absorption technologies. Greenhouse Gas R&D Programme, International Energy Agency, 2002. http://www.ieagreen.org.uk
239. Modernizing ammonia plants with innovative technical solutions. Kellogg Brown & Root, Houston, 2000. http://www.mwkl.co.uk/pdf/modamplt.pdf
240. Korotkikh O, Farrauto R (1999) Selective catalytic oxidation of CO in H2 for fuel cell applications. Engelhard Corporation, Iselin, NJ. http://hwww.unl.edu.ar/cepac/abstract/bfarraut.htm
241. http://www.kvaerner.com. Catalytic Rich Gas (CRG) Technology, Kvaerner process technology, 27 Aug 2001
242. Furukawa SF, Bartoo RK (1997) Improved Benfield process for ammonia plants, UOP technical papers, UOP LLC, Des Plaines, Jan 1997. http://www.uop.com/gasprocessing/techpapers/ImproveBenfield.pdf
243. Christensen PV (2001) Revamping ammonia plants to follow feedstock situation and market demands, Haldor Topsøe Ammonia Publications. Presented at FAI symposium on "improving productivity of ammonia and urea plants" in New Delhi, India, Lyngby, Denmark, Apr 2001. http://www.haldortopsoe.com

244. Qafco Preparing for Higher Exports, Gulf Industry Online.com, http://www.gulfindustryworldwide.com/Articles.asp?Article=1660, 11(5), Al Hilal Publishing & Marketing Group, Sep–Oct 2002
245. Ammonia Properties. The Columbia electronic encyclopedia, 6th edn. Columbia University Press, 2000. http://www.encyclopedia.com/articles/00445Properties.html
246. van den Brink RW (2001) Combined catalytic removal of N2O and NOx from the tail gases of a nitric acid plant. Energy Research Centre of the Netherlands, ECN-RX-01-012, Apr 2001. http://www.ecn.nl/library/reports/2001e/rx01012.html
247. Schwefer M, Maurer R, Groves M (2000) Reduction of nitrous oxide emissions from nitric acid plants. Prepared for presentation at nitrogen 2000, Vienna, Mar 2000
248. Ammonia chemical profile. Chemical Marketing Reporter (25 Nov 2002)
249. Greene R (2003) Turning stranded gas into cold cash. Chem Eng Prog 97(2):13
250. Fritsch S (2000) Steam reformer-based hydrogen plant optimisation. In: Prepared for presentation at HYFORM 2000, Munich, Sept 2000
251. Dall'Aglio L (1998) Fluid drum granulation for ammonium nitrate, Kaltenbach-Thüring, The International Fertiliser Society, Proceeding 415. http://www.fertiliser-society.org/Proceedings/Prc415.HTM
252. Sato L (1990) A new route for granular ammonium nitrate, The International Fertiliser Society, Proceeding 296. http://www.fertiliser-society.org/Proceedings/Prc296.HTM
253. DSM melamine product. DSM Melamine, Sittard, 2002. http://www.dsm.com/dmm/products
254. Bryan KB (1993) Melamine by the melamine chemicals M-II process. PEP Review No. 92-1-1, SRI Consulting. process-economics.com/reviews/92-2-1.htm
255. Melamine process block diagram, Eurotecnica—Melamine, Milan, 2003. http://www.eurotecnica.it/melamine.htm
256. Market survey on melamine industry in China, CCM/Market Reports, Mar 2002. http://www.cnchemicals.com/maindocs/marketreports/maincontent/Melamine.htm
257. Nitroprocess Technologies, Platecxis, Engelhard-CLAL, Paris, 2002. http://www.engelhard-clal.fr
258. Kirk-Othmer (1997) Encyclopedia of chemical technology, vol 17. Wiley, New York, pp 444–467
259. Ammonium phosphates chemical profile. Chemical Marketing Reporter (18 Nov 2002)
260. Hagen K (1990) The future of ammonium sulphate—Co and By-Production, The International Fertiliser Society, Proceeding 295. http://www.fertiliser-society.org/Proceedings/Prc295.HTM
261. Lauriente DH (1997) CEH marketing research report, Ammonium sulfate. SRI International
262. Sulfate A (2000) Chemical products synopsis. Mannsville Chemical Products Corp, Adams
263. Kirk-Othmer (1993) Encyclopedia of chemical technology, 4th edn, vol 10. Wiley, New York, pp 452–453
264. Heilen G, Mercker HJ et al (1996) Aliphatic amines. In: Ullmann's encyclopedia of industrial chemistry, 5th edn, vol A2. Reinhold, New York
265. The Nitto Chemical Methylamines Process, PEP Review No. 89-3-4, SRI Consulting, , Feb 1991. processeconomics.com/reviews/89-3-4.htm
266. Morpholine health and safety guide. IPCS International Programme on Chemical Safety, Health and Safety Guide No. 92, World Health Organization, Geneva, 1995. http://www.inchem.org/documents/hsg/hsg/hsg092.htm
267. Davenport RE (2000) CEH Product Review, Chelating agents. SRI International, Mar 2000
268. Chelating agents, Chemical of the Week, Oct 2001. http://scifun.chem.wisc.edu/chemweek/ChelatingAgents/ChelatingAgents.html
269. Kirk-Othmer (1997) Encyclopedia of chemical technology, vol 13. Wiley, New York, pp 560–606
270. Hydrogen cyanide technology manual. DuPont Chemicals Solution Enterprise, Memphis, May 1994
271. Aniline: supply demand balance and stable Price in China. China Chemical Reporter, 11(25):13–15 (5 Sept 2000)
272. Camera E (1997) CEH data summary, nitrobenzene, SRI International, June 1997
273. Hydrogen data, Air Liquide, Paris, 2001. http://www.airliquide.com/en/business/products/hydrogen
274. Shell sulfinol process. Shell Global Solutions, Shell Oil Products Company, Houston, 2003. http://www.sulfinol.com
275. Aniline chemical profile. Chemical Marketing Reporter (4 Feb 2002)
276. Nitrobenzene chemical profile. Chemical Marketing Reporter (15 Apr 2002)
277. Caprolactam chemical profile. Chemical Marketing Reporter (16 Apr 2001)
278. Chiyoda and DSM Alliance in Caprolactam Projects, Chiyoda Corporation News Release, Yokohama, 26 Apr 2001
279. True closed-loop recycling for nylon. Environmental Building News, BuildingGreen.com, vol 8(9), Brattleboro, Sept 1999
280. Production at evergreen nylon recycling suspended, DSM Corporate Communications, Heerlen, 30 Aug 2001. http://www.dsm.com/newsarchive
281. Schut JH (2002) Big German plant may relieve U.S. Bottleneck in recycling carpet nylon. Plastics Technology Online, New York, May 2002. http://www.plasticstechnology.com/articles
282. Production process, Polyamid 2000, Premnitz, Oct 2000. http://www.polyamid2000.com
283. Adipic acid/adiponitrile, Product Focus. Chem Week 161(9):31 (1999)
284. Stahl WF (1996) CEH product review, Hexamethylenediamine (HMDA), SRI International, Nov 1996
285. Schwarz J, Beloff B, Beaver E (2002) Use sustainability metrics to guide decision-making. Chem Eng Prog 98(7):58–63
286. Acrylonitrile. Asian Chem News:10–13 (2002)
287. Dimethylformamide from supercritical carbon dioxide, Design Projects, University of Notre Dame, South Bend, June 2003. http://www.nd.edu/~enviro
288. Dimethylformamide Producing Companies, China Report: chemical product trends, SRI Consulting, 2003. http://www.cnonline.com/cni_home
289. Dimethylformamide, methylamine and derivatives process technologies, Davy Process Technology, 2003. http://www.davyprotech.com
290. Mitsubishi Rayon reaches agreement on joint venture to manufacture and sell petrochemical products in China, Chemchannels.com, Tokyo, 19 Dec 2002. http://www.chemchannels.com
291. Dimethylformamide, solvents, UCB Chemicals—methylamines & derivatives, 2003. http://www.chemicals.ucbgroup.com
292. Courtney PJ (2002) Light-curing cyanoacrylates. Adhes Age: s8–s11
293. http://www.chemetics.kvaerner.com. Mononitrobenzene, Kvaerner Pulp & Paper, 2003
294. Demand growth of acrylonitrile faster than output increase. China Chemical Reporter 14(30):18 (26 Oct 2003)
295. Acrylonitrile supply short of demand. China Chem Week 464:7 (2003)
296. Norsk Hydro A.S., Oslo. Method for reduction of nitrogen oxide. EP 0 359 286 B1 (1993)
297. Kolmeijer NW (2000) Phosphate fertilizers. In: Ullmann's encyclopedia of industrial chemistry. Wiley-VCH, Weinheim. http://www.mrw.interscience.wiley.com/ueic/articles/a19_421

23

Phosphorus and Phosphates

Glenn A. Gruber

Introduction

Phosphates, compounds of the element phosphorous, are produced from relatively abundant supplies of phosphate rock.

The major use of phosphate is to supply phosphorous, one of the three essential plant foods—nitrogen, phosphorus, and potassium. Phosphate rock extraction from its ore, and its subsequent conversion into fertilizer materials and industrial chemicals, is a relatively mature art. Single superphosphate, a mixture of monocalcium monohydrate and gypsum formed by the reaction of sulfuric acid with phosphate rock, has been used as a fertilizer since the mid-1800s. Phosphoric acid, derived by the treatment of phosphate rock with sulfuric acid so as to produce gypsum in a separable form, was manufactured in many locations by batch and countercurrent decantation methods in the 1920s.

Phosphoric acid produced by the later process is called the "wet process acid" to distinguish it from "furnace phosphoric acid," which is produced by hydrating the phosphorous pentoxide formed by burning elemental phosphorous in air. The "wet" processes utilized for the bulk of today's phosphoric acid and fertilizer production trace their origins to the early 1930s [1].

Higher-purity industrial and food-grade phosphates, until recently, were most often derived from furnace processes. New plants recover purified phosphoric acid suitable for food-grade uses from relatively impure wet process acid, using solvent extraction technology.

Phosphate Rock

Naturally occurring mineral products having sufficient phosphate content to be of commercial value are classified as phosphate rock. The grade or phosphate content of these products has been traditionally reported as percent $Ca_3(PO_4)_2$, which is referred to as bone phosphate of lime (BPL), tricalcium phosphate (TCP), or triphosphate of lime (TPL). Stoichiometric factors relating traditional rock analysis to other commonly used analytical terms are 100% BPL = 45.77% P_2O_5 = 19.96% P.

International trade of phosphate rock is based on dry metric tons, whereas US domestic shipments are reported as short tons, on a dry or as is moisture basis. Specifications for purchased rock may address grade, particle size, moisture content, and chemical impurities such as CaO, MgO, SiO_2, Al_2O_3, Fe_2O_3, F, Cl, Na_2O, and K_2O. The content of organic material and heavy metals is also of importance.

The treatments by which phosphate rock is commonly converted to fertilizers and chemicals are summarized in Fig. 23.1.

Minerals

The most common and widely distributed phosphate minerals are the apatite group, with the general formula $Ca_{10}(PO_4)_6(X)_2$. The apatite is designated as fluorapatite, hydroxyapatite, or chlorapatite, when X = F, OH, or Cl, respectively. The most abundant sedimentary apatite is carbonate fluorapatite (francolite). Relative to pure fluorapatite, francolite is characterized by the substitution of Na and Mg for Ca and of carbonate and fluoride for phosphate. An empirical formula for francolite is given below and the chemical compositions of the end members of the fluorapatite–francolite series, as quantified by McClellan [2], are given in Table 23.1.

G.A. Gruber (✉)
Jacobs Engineering Company, 3149 Winter Lake Road, Lakeland, FL 33803, USA
e-mail: glenn.gruber@jacobs.com

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_23,

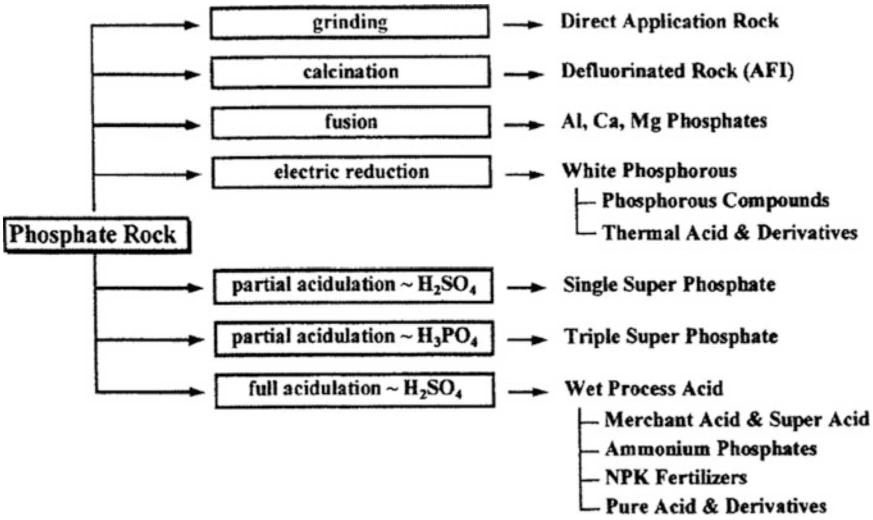


Fig. 23.1 Phosphate rock treatments and end products

Table 23.1 Composition of the fluorapatite–francolite end series

Constituent	Fluorapatite	Francolite
%CaO	55.60	55.40
%P_2O_5	42.20	34.00
%CO_2	–	6.30
%F	3.77	5.04
%Na_2O	–	1.40
%MgO	–	0.70

Source: McClellan [2].

Francolite

$$(\mathrm{Ca}_{10-a-b}\mathrm{Na}_a\mathrm{Mg}_b\mathrm{PO}_4)_{6-x}(\mathrm{CO}_3)_x\mathrm{F}_{0.4x}\mathrm{F}_2$$

Van Kauwenberg [3] has described the mineralogy and alteration of phosphate ores in Florida. Mineralogical composition varies by particle size. Francolite and then quartz are the most abundant minerals for plus 20-mesh particles (pebble), while the reverse is true for particles in the 20–200-mesh fraction (flotation feed). In the −200-mesh size fraction (clay waste) quartz, francolite, wavellite, crandallite, goethite, dolomite, and a variety of clay minerals such as smectite, kaolinite, illite, and palygorskite occur.

Resources and Ores

Naturally occurring phosphates exist, or originated, as accessory minerals in igneous rocks. Prolonged weathering gradually converts the water-insoluble apatite into dissolved compounds that accumulate in the world's oceans. Sedimentary marine deposits (phosphorites) are formed when phosphorous compounds are precipitated by chemical or biological reactions. Bernardi [4] describes secondary enrichment as an important aspect in the formation of sedimentary deposits.

McKelvey [5] reported that the earth's crust contains an average of 0.27% P_2O_5, most of which occurs as apatite species. Sedimentary rocks, which predominate at the earth's surface, host the majority of commercial phosphate deposits discovered to date. Igneous rocks, which make up about 95% of the earth's crust, contain few phosphate deposits of commercial value. Guano deposits formed from the droppings of sea birds or bats are of minor importance, as are guano-related deposits. Northolt [6, 7] describes known phosphate deposits in the world, and estimates that identified phosphate resources in North America total more than 35 billion metric tons.

Resources are typically quantified as in situ tons of phosphatic material, without regard for economic criteria. However, it is preferable to quantify phosphate reserves as tons of phosphate rock recoverable according to specified economic, chemical, and regulatory criteria [8]. The definition of phosphate reserves therefore requires an integrated program of geological exploration, laboratory testing, and classification using applicable criteria.

Significant commercial deposits of sedimentary phosphate ore occur in the United States, the Former Soviet Union, Morocco, China, Jordan, and Tunisia, and lesser deposits are mined in many other countries. Although phosphorite ores generally are classified as having siliceous or carbonate gangue minerals, soluble salts, and organic material are also of concern. The phosphate content of the ores, depending on conditions of deposition and secondary enrichment, ranges from 10% to more than 70% BPL. The recovery of byproducts from phosphorite ores is uncommon; however, uranium has been extracted commercially from phosphoric acid.

The types of igneous rock in which commercial deposits of apatite have been found are nepheline–seyenite and carbonatites. The apatite deposits in Russia's Kola

Fig. 23.2 Loading trucks with phosphate ore in Jordan

Peninsula are associated with nepheline–seyenite. Carbonatite deposits that are mined for their phosphate content include Siilinjarvi in Finland, Jacupiranga and Araxa in Brazil, Phalaborwa in South Africa, and Kapuskasing in Canada. By-product recovery from igneous phosphate ores is common. Nepheline ($NaAlSiO_4$) is recovered from the Russian ore, calcite from the Finnish ore, copper concentrate and baddelyite (ZrO_2) from the South African ore, and barite from Brazilian ore. A minor percentage of the world's phosphate rock production is recovered from guano-related phosphate deposits.

Mining

Phosphate ores are extracted from deep deposits by underground mining methods and from shallow deposits by surface mining methods. Underground mining tends to be more costly and therefore less common for phosphate deposits than surface mining. Because 1 t of phosphate rock has only about 10% of the commercial value of 1 oz of gold, low-cost mining is imperative.

Sedimentary phosphate deposits are exploited by underground mining in China, Mexico, Morocco, and Tunisia. Generally, the flat laying ore is most economically extracted by room-and-pillar mining or long-wall mining. Ore from deep sections of the igneous phosphate deposits in Russia's Kola Peninsula is mined by a block caving technique.

Many shallow deposits have unconsolidated ore covered by unconsolidated overburden. Large electric walking draglines are ideally suited for such deposits, as evidenced by their use at large capacity phosphate mines in the southeastern USA, Morocco, Jordan, Senegal, and Togo. Other deposits have over burden and ore that may be partially or fully consolidated. For these deposits, ripping or drilling and blasting are required to fragment the overburden and ore to the extent that they can be excavated. Power shovels, backhoes, and wheeled loaders are also commonly used for excavation. Figure 23.2 shows a hydraulic shovel loading phosphate ore into a haul truck in Jordan.

Bucketwheel excavators are used for overburden removal at phosphate mines in eastern North Carolina, Senegal, and Togo. One mine in central Florida has used cutter head dredges for both overburden removal and ore excavation.

The method of transporting ore from the mine to the beneficiation plant depends on ore characteristics, mining methods, and local infrastructure. Railroad transport has been practiced in Russia and Iraq. Haul trucks and belt conveyors are commonly used in China, Jordan, Mexico, Morocco, Russia, Syria, Tunisia, and the western United States. Slurry pipelines of 18–22 in. diameter, operating at less than 100 psig, are used exclusively in central Florida, north Florida, and eastern North Carolina. The pipelines may be extended up to 10 miles or more, by installing a series of centrifugal slurry pumps at 4,000 ft intervals.

A typical Florida phosphate mining scheme, utilizing a dragline and slurry pipeline, is illustrated in Fig. 23.3. The dragline first exposes the phosphate ore (matrix) by stripping and casting the overburden into the adjacent mined area. The matrix is then dug by the dragline and placed in a slurry pit, where it is gunned with high-pressure water. Gunning the matrix, as shown in Fig. 23.4, transforms the unconsolidated ore into a slurry which is pumped to the beneficiation plant.

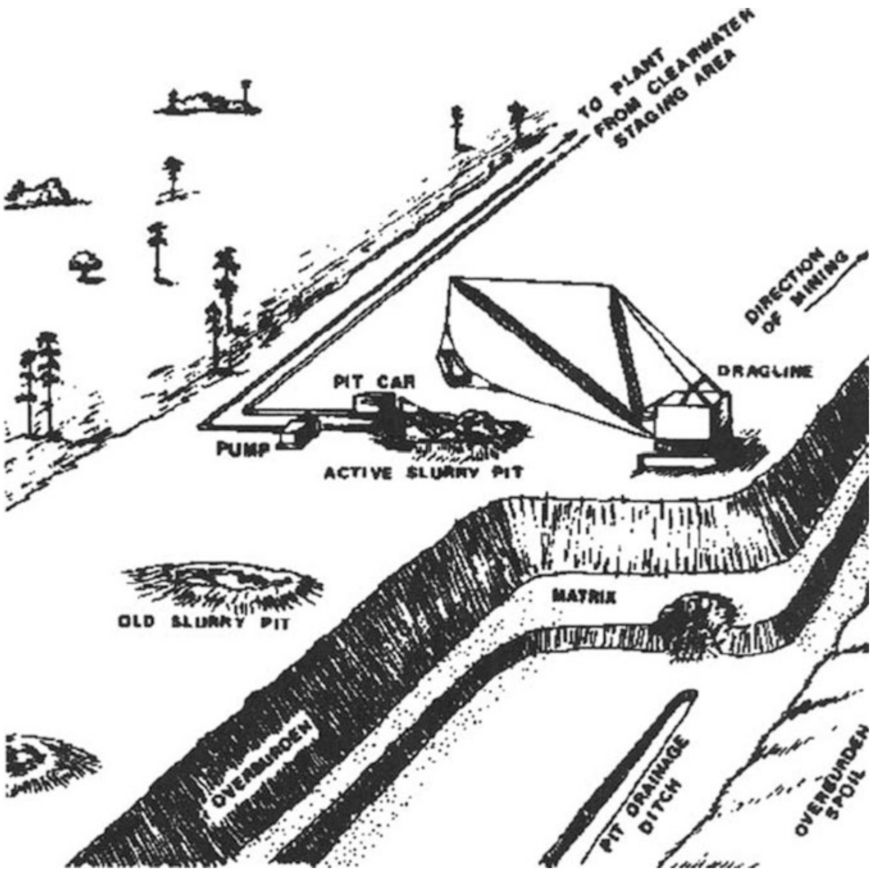


Fig. 23.3 Typical Florida phosphate mine

Fig. 23.4 Gunning phosphate rock in a Florida mine

Beneficiation

Beneficiation, also known as mineral dressing or ore processing, may involve a variety of operations such as size reduction, size separation, mineral separation, dewatering, and thermal processing. Almost all phosphate ores require beneficiation to meet commercial specifications concerning particle size, moisture content, or chemical analyses.

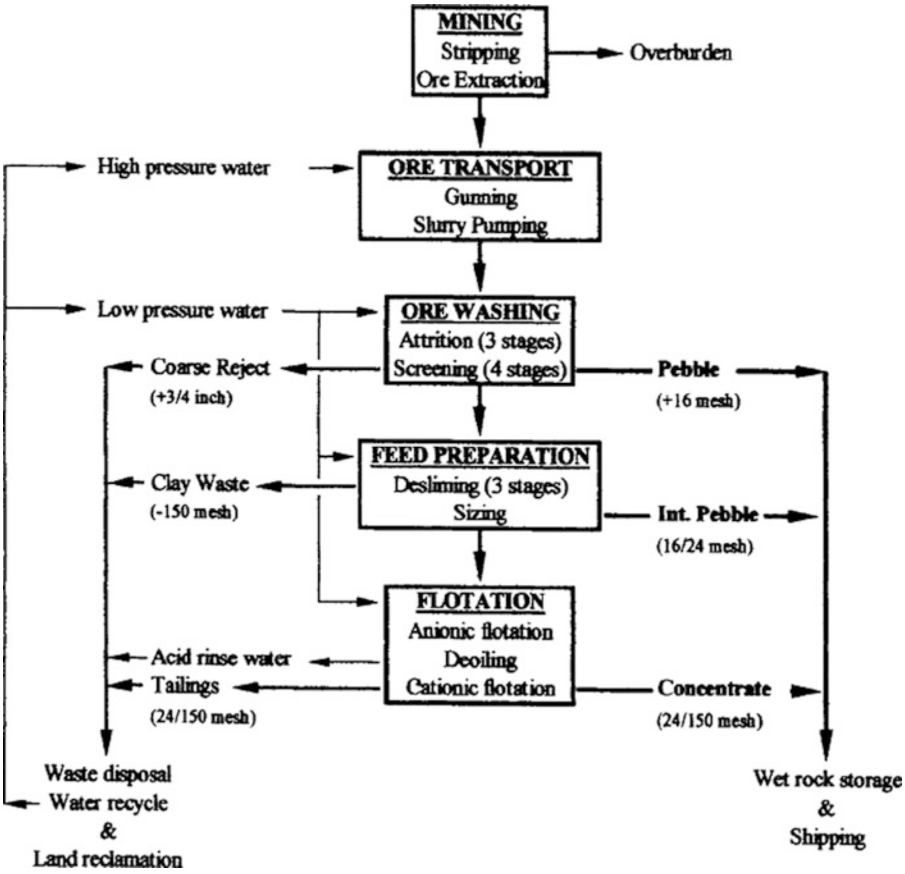


Fig. 23.5 Unconsolidated sedimentary ore beneficiation flow diagram

The usual first beneficiation operation is size reduction, which may be achieved by crushing, grinding, or disaggregating by scrubbing and washing. Particle size reduction liberates mineral species so that they can be separated. Size separation usually follows size reduction. When gangue minerals are more indurated than the phosphate, it is often practical to reject coarse waste material by wet or dry screening. Similarly, when gangue minerals are microcrystalline or softer than the phosphate, fine waste material may be rejected by wet or dry classification. Soluble chlorides, when present, must be removed from phosphate rock by washing with fresh water followed by dewatering.

Gangue minerals frequently have the same particle size as the phosphate mineral grains, and techniques such as heavy media separation, magnetic separation, or froth flotation are required. Heavy media separation is an appropriate process when liberation occurs at 16 mesh or coarser, and the phosphate mineral has a significantly higher density than the gangue (dolomite, calcite, quartz, shale). Low-intensity magnetic separation will remove highly magnetic minerals, such as magnetite [9], from phosphate. High-intensity magnetic separation will remove ankerite [10] and other paramagnetic iron-bearing minerals [11] from phosphate. Froth flotation is the most widely practiced operation for recovering phosphate rock from fines (−20 mesh). Variations of this process are used commercially to separate phosphate from barite, calcite, dolomite, feldspars, nepheline, phlogopite, and quartz. Flotation plants have been constructed and operated in Brazil, Canada, China, Finland, Jordan, Mexico, Russia, Senegal, South Africa, and the United States. Dolomite flotation from phosphate is of increasing interest. In the United States, one commercial plant has a dolomite flotation circuit [12], and other dolomite flotation processes have been demonstrated by pilot plant testing of Florida low-grade pebble [13].

Electrostatic removal of quartz from apatite is technically feasible although it is impractical and costly [14].

A generic scheme for mining and beneficiating central Florida phosphate ore is presented in Fig. 23.5 as a block flow diagram.

Liberation of phosphate from the gangue occurs during ore transport and washing. First, a low-cost product, called pebble, is recovered by screening the ore at about 16 mesh. Secondly, a low-grade product (intermediate pebble) is recovered by sizing the flotation feed at about 24 mesh. Clays are removed from the flotation feed by three or more stages of desliming with hydrocyclones. Finally, a more expensive but higher grade concentrate is obtained by a two-stage flotation process. A rougher phosphate concentrate is recovered by direct flotation with anionic reagents.

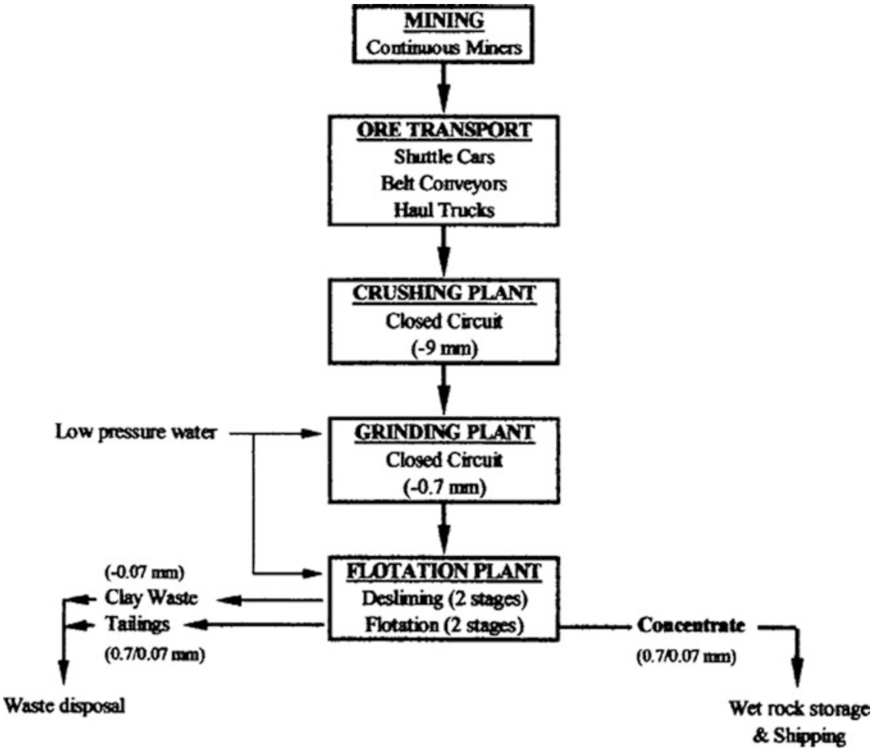


Fig. 23.6 Consolidated sedimentary ore beneficiation flow diagram

After deoiling with sulfuric acid and rinsing with water, the rougher concentrate is conditioned with cationic reagents and subjected to inverse flotation. The phosphate rock product, comprised of pebble, intermediate pebble, and flotation concentrate, dispatched to a chemical plant for conversion to phosphoric acid.

The initial beneficiation steps for consolidated phosphate ores generally differ from those of unconsolidated ores. Figure 23.6 depicts the flow diagram for the San Juan de la Costa phosphate mine in Mexico. The high-grade ore, slightly more than 1 m in thickness, is extracted by room-and-pillar mining. Continuous miners rip ore from the mining face and load shuttle cars, which transfer the ore to feeder-breakers and belt conveyor systems. Outside the mine, ore is loaded into haul trucks and transported to the beneficiation plant. Liberation of the phosphate is accomplished by crushing to 9 mm followed by grinding to 0.7 mm. Following grinding, the ore is deslimed, attrition scrubbed, and deslimed a second time to remove clays and carbonate minerals from the flotation feed. The feed is conditioned with anionic reagents and subjected to rougher and cleaner direct flotation using sea water. The concentrate is washed with fresh water to remove sea salt prior to use in the chemical plant.

Figure 23.7 presents a simplified mid-1980s flow diagram for the Siilinjarvi phosphate beneficiation plant in Finland. The low-grade igneous ore to this plant is carefully blended to avoid changes in plant feed characteristics. Liberation of phosphate is achieved by rod milling followed by closed circuit ball milling.

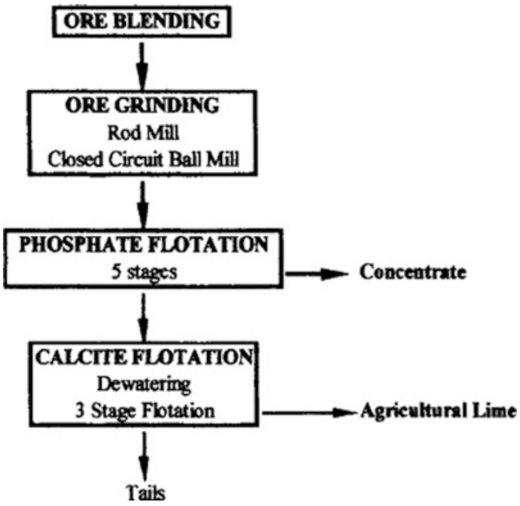


Fig. 23.7 Igneous ore beneficiation flow diagram

The ground ore is conditioned without desliming, at pH 11, using an amphoteric flotation reagent. A phosphate concentrate is recovered by five stages of direct flotation. The tailings from phosphate flotation are dewatered, conditioned with anionic reagents, and subjected to three stages of direct flotation to recover calcite, which is used as agricultural lime.

Table 23.2 Phosphate rock value (United States marketable production)

Crop (year)	Total value (billion $)	Unit value ($/metric ton)
1975	1.107	25
1980	1.148	21
1985	1.236	24
1990	1.075	23
1995	0.947	22
2000	1.060	26
2005	0.999	28

Source: USGS, Mineral industry surveys

Table 23.3 World production of phosphate rock (million metric tons)

	1980	1990	2000	2008
United States	53	46	39	31
Morocco	19	21	22	24
China	11	17	34	61
USSR/Russia	25	37	11	10
Tunisia	5	7	8	8
Jordan	4	6	6	6
Other	22	23	26	36
Total world	139	157	146	176

Source: ifa, World phosphate rock statistics

Production and Value

The dollar value of marketable phosphate rock produced in the United States since 1975 is summarized in Table 23.2. The statistics apply to crop year (July 1 to June 31) and represent current dollars. The unit value of phosphate rock did not keep up with inflation and consequently the production capacity of US phosphate mines that were depleted in the last 30 years has not been replaced on a one-for-one basis. During the 2006–2009 period the international price for exported phosphate rock peaked at about $400 per metric ton and then declined to about $100 per metric ton.

The global demand for food stimulated increased fertilizer usage and consequently increased phosphate rock consumption through 1990. Table 23.3 shows that global phosphate rock production decreased from 1990 to 2000 and then recently exceeded the 1990 level. Although phosphate rock production in the United States declined by about 50% over the last 30 years, the United States remains a major producer. China's commitment to become self sufficient in phosphate rock is evidenced by the 50-million metric ton increase ion production since 1980. China has now replaced the United States as the leading producer of phosphate rock.

Many countries use indigenous phosphate rock as a source of phosphorous for industrial chemicals and fertilizers. Few countries are self-sufficient and supplemental sources of phosphate rock are essential. The cost of imported phosphate rock is markedly influenced by freight, and therefore the low-cost producer is not necessarily the low-cost supplier.

Since 1980, United States went from the major exporter to a net importer of phosphate rock. The phosphate rock is converted in the United States to high analysis of granular fertilizers, some of which are exported. Morocco has been the major exporter of phosphate rock for several years.

Chemical Processing of Phosphate Rock

Phosphate rock is converted into usable chemicals by two methods. In the first, the rock is charged to an electric furnace with silica and coke to produce elemental phosphorus. The phosphorus then is converted into phosphoric acid and other compounds. In the second, the phosphate rock is reacted with sulfuric acid in a medium of phosphoric acid and calcium sulfate crystals to form dilute, impure phosphoric acid, the acid is separated and used to make fertilizers. This is known as the wet process method [15].

The wet process is further divided into two subprocesses based on the type of calcium sulfate crystal produced. The dihydrate process wherein gypsum ($CaSO_4$•2H2O) is produced has been the dominant process, but processes making hemihydrate ($CaSO_4$•1/2H_2O) have become more important over the past decade.

Thermal Process for Phosphorus and Phosphoric Acid

The furnace or thermal process is shown in Fig. 23.8. The approximate reaction is:

$$Ca_3(PO_4)_2 + 6SiO_2 + 10C \rightarrow 6CaSiO_3 + P_4 + 10CO$$

The phosphorus leaves the furnace as a vapor and is condensed by direct contact with water. Phosphoric acid of high purity is made by burning phosphorus with air and hydrating the resulting P_2O_5 with water, according to the reaction:

$$P_4{+}5O_2{+}6H_2O \rightarrow H_3PO_4$$

If even less water is used for hydration, a product known as polyphosphoric acid results. Ordinary phosphoric acid is a solution of the monomer, H_3PO_4, in water, and is called orthophosphoric acid. If a molecule of water is removed between two orthophosphate molecules, the dimer, pyrophosphoric acid, $H_4P_2O_7$, is formed. Similarly, the trimer and higher polymers can be made. Superphosphoric acid is a mixture of orthophosphoric acid and polyphosphoric acid and is now made from wet process acid as described later in this chapter.

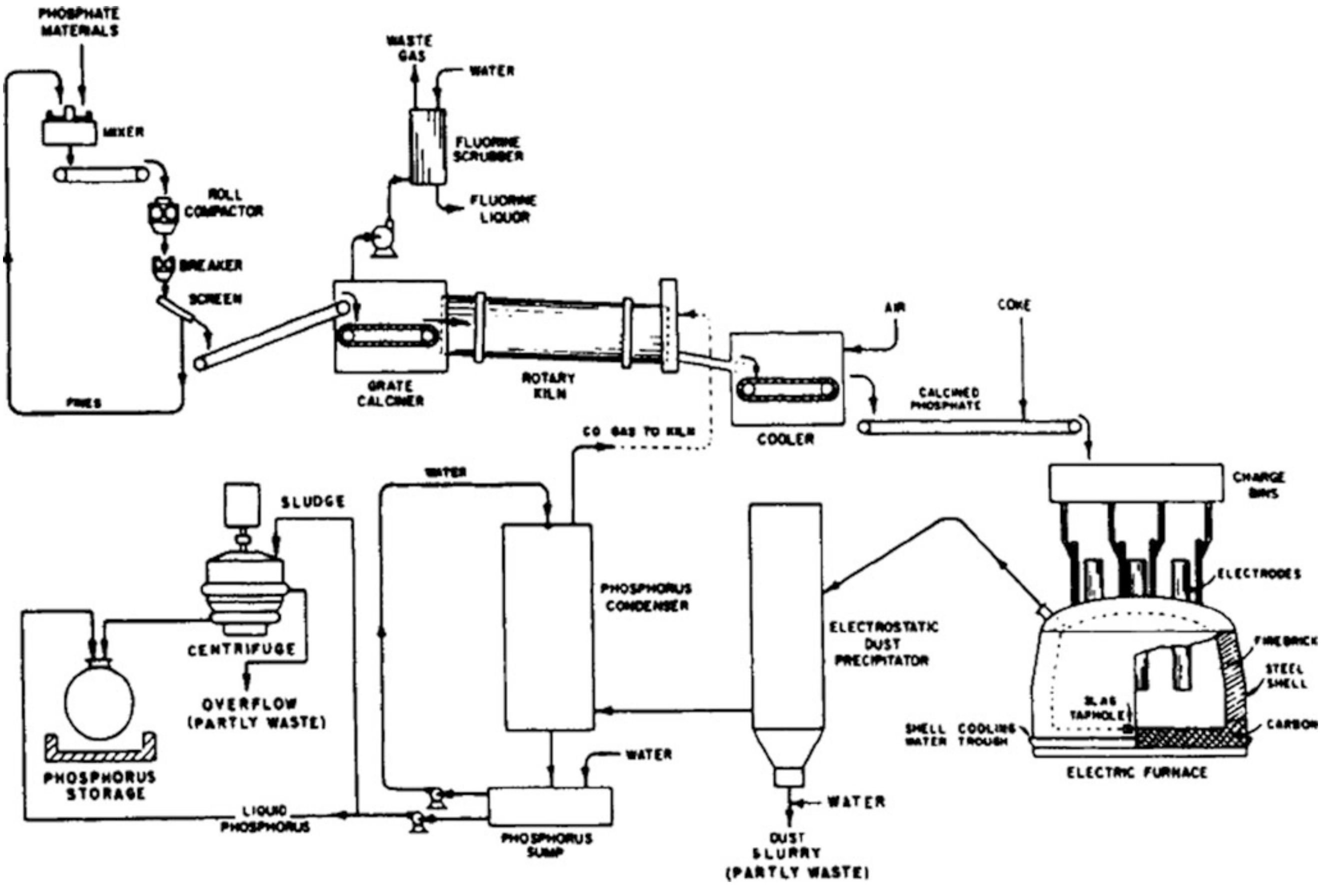


Fig. 23.8 Electric furnace process for production of elemental phosphorus

In 1990, there were about eight plants in operation, some with multiple furnaces, in the United States. By 2000–2001, only one plant remained. New emission standards, high capital and operating costs, and competitive lower-cost wet acid purification technology have spelled doom for most of the furnace plants. A more thorough discussion of electric furnace processing is to be found in the ninth edition of this *Handbook*.

Industrial Phosphates

Furnace phosphorus currently produced in the United States is consumed to make compounds such as phosphorus pentoxide, phosphoric trichloride, and phosphorus pentasulfide, which find use in the preparation of drying agents, plasticizers, oil additives, fire retardants, and insecticides. These products are derived from phosphorus and, therefore, cannot be made from purified wet process acid so there is a continuing demand for a small amount of thermal product. However, for orthophosphoric acid use, the purified wet process acid is cheaper [16]. This has now replaced furnace acid in soft drinks, candy, baked goods, and various other food products. It is also used for pickling metals.

The cheapest and most important salts of purified phosphoric acid are the sodium salts, made by reacting the acid with sodium carbonate or sodium hydroxide. Sodium phosphates may be classified in a general way as (1) orthophosphates, (2) crystalline condensed phosphates, and (3) glassy condensed phosphates.

Three sodium orthophosphates can be prepared, depending on whether one, two, or three hydrogen atoms are replaced by sodium.

Monosodium phosphate is formed in the following reaction:

$$2H_3PO_4 + Na_2CO_3 \rightarrow 2NaH_2PO_4 + H_2O + CO_2$$

Sodium carbonate is also used to make disodium phosphate, Na_2HPO_4, but sodium hydroxide must be used to replace the third hydrogen in trisodium phosphate, Na_3HPO_4. The orthophosphates have a wide range of uses in industry.

Condensed phosphates are made by eliminating water from orthophosphates. The most important compound is sodium tripolyphosphate (STPP), made according to the following reaction:

$$2Na_2HPO_4 + NaH_2PO_4 \xrightarrow{-H_2O} Na_5P_3O_{10}$$

The most important use of STPP is as a builder in detergents. However, legislative restrictions on the use of phosphorus compounds in household detergents have caused a Worldwide flattening of consumption.

Glassy condensed phosphates are represented by sodium hexametaphosphate, in which the O:P ratio is 3:1. There can be considerable variation in the Na_2O to P_2O_5 ratio.

The principal use of the condensed phosphates is to sequester metallic ions in water. They form water-soluble complexes with the metals and prevent metallic compounds from precipitating to cause discoloration, scale, and sludges.

Wet Process Phosphoric Acid

In the wet process, phosphate rock is reacted in a slurry of phosphoric acid and calcium sulfate crystals containing a controlled quantity of sulfuric acid. The simplified reactions for the dihydrate process is as follows:

$$Ca_3(PO_4)_2 + 3H_2SO_4 \rightarrow 3CaSO_4 \cdot 2H_2O + 2H_3PO_4$$

Until 1969, virtually all wet process acid was made at maximum strengths of 26–32% P_2O_5, separating the calcium sulfate as gypsum containing two molecules of water. Since 1969, a substantial number of commercial hemihydrate process plants have been built in various modes. In the hemihydrate process, reaction conditions are higher in temperature and phosphoric acid strength so that the stable solid phase is calcium sulfate with one-half molecule of water.

The hemihydrate process may be a single-stage process, known as HH, in which the hemihydrate solids are the waste product, or the process can have a second step in which the hemihydrate is recrystallized to gypsum, known as hemi-dihydrate (HDH). In the hemihydrate process, acid strengths of 38% P_2O_5 to about 42% P_2O_5 normally are produced although strengths up to 50% can be produced under somewhat more difficult circumstances requiring more filter area.

The neutralization reaction of the above equation is conducted in one or more strongly agitated reaction vessels, whether in a gypsum or in a hemihydrate mode. The system is highly exothermic and the slurry is maintained at 80–85°C for dihydrate processing, 95–100°C for hemihydrate, by evaporative or air cooling. During the reaction of phosphate rock with sulfuric acid, fluorine is evolved and must be scrubbed from the vent gas.

Table 23.4 Annual production of phosphoric acid by region (1,000 mt P_2O_5)

Region	2000	2003	2006	2008
West Europe	1,275	1,117	948	818
Central Europe	540	525	548	503
East Europe & Central Asia	2,425	2,547	2,945	3,066
North America	10,537	10,694	9,571	8,182
Latin America	1,683	1,460	1,599	1,674
Africa	5,274	5,967	5,924	5,296
West Asia	1,742	1,639	1,655	1,364
South Asia	1,014	1,012	1,395	1,138
East Asia	3,202	5,442	9,010	10,172
Oceania	152	438	420	433
Total world	27,843	30,841	34,015	32,646

Source: ifa, World Phosphoric Acid (PA) statistics

Table 23.4 lists the production of phosphoric acid in the recent past, and shows the substantial changes in location of P_2O_5 production over the last decade. The East and West Europe fertilizer industry was in a steep decline by 1999. North Africa and the Middle East are continuing to expand based on their huge reserves of high-grade phosphate. China, with substantial deposits, has now begun to develop these, as evidenced by the significant increase in the production of phosphate rock and phosphoric acid since the year 2000.

Dihydrate Process

The conventional dihydrate process, as operated over the last 70 years, was first demonstrated by the Dorr Company in the Cominco plant at Trail, British Columbia, in 1931 [1]. The principles discovered at that time for making an acid strength of up to 32% P_2O_5 and a highly filterable gypsum crystal are still used today. The salient features of the process included maintaining H_2SO_4 content in the digestion acid of about 2.0–3.0%. Reaction slurry was recycled at a ratio of 10–15:1 relative to the volume of product slurry sent to the filter. The relatively higher recirculation of seed gypsum than had previously been practiced, together with control of the free sulfate, resulted in product acid well above the previous 22% P_2O_5 possible with batch and countercurrent decantation systems.

If the sulfate in the solution is allowed to exceed certain limits, depending on rock reactivity and particle size, the rock becomes coated with gypsum, and the reaction becomes uncontrollable. If the sulfate level is too low, the precipitated gypsum filters poorly and contains excessive quantities of crystallized (solid solution) P_2O_5, leading to excessive P_2O_5 losses. Over the years, the process has evolved to employ increased slurry recirculation and different reactor configurations, In addition, several different filter designs are available.

Table 23.5 Phosphoric acid plants, worldwide dihydrate process, 2009

Process	Number of lines 450 MTPD or larger
Prayon	20–25
Rhone Poulenc	10–15
Nissan H	7
Jacobs/Dorr–Oliver	8
Badger–Isothermal	6

The principal dihydrate processes in use as of 2,000 are shown in Table 23.5. During the decade 1990–2000, many of the smaller plants and even some larger than 450 t per day P_2O_5 plants have been shut down due to environmental or market conditions. This has resulted in more production in Morocco, Jordan, and India, and less in Europe. Most of the world's phosphoric acid is produced by the dihydrate method, but there is likely to be increased production by the hemihydrate method because of advantages in some situations. The dihydrate processes listed have been modified over the years so that many slightly different configurations of the same process may exist.

The unique feature of each process is the reactor system configuration. The reactor or reactors normally provide from 2 to 6 h detention for the gypsum slurry. This is about 0.8–2.5 m^3 of reactor volume per ton of P_2O_5 per day, meaning relatively large vessels totaling 1,500 m^3 to over 2,000 m^3 in size for large plants. The individual reactor systems are described briefly below.

The objective of the reaction system is to produce a highly filterable gypsum crystal that washes well, and that also contains a minimum amount of insoluble P_2O_5. The filter ability of the gypsum slurry depends on the reactor configuration and on the rock source. Highly filterable gypsum slurries are produced from Florida, Togo, and Senegal phosphates. These rocks may provide more than double the P_2O_5 from a given filter when compared with gypsum produced from other sources. However, because phosphate rock is costly to transport, the use of local, less treatable phosphates can be quite economical. Table 23.6 lists the approximate filtration design rates for these groups of commercially available phosphate rock.

The advantages of the dihydrate process vs. the various, newer hemihydrate configurations are as follows.

Table 23.6 Filtration design rates for phosphoric acid produced from various rock sources

Rock source	Tons P_2O_5 produced per m^2 active area
Togo	
Florida	7.5–9.0
Senegal	
Morocco Khouribga	
Western U.S.	4.5–7.0
Kola (USSR)	
North Carolina	
Morocco Safi	
Algeria	2.5–5.0
Tunisia	

1. The water balance permits the use of wet rock slurry feed. This eliminates the cost of drying the rock and the dust nuisance. This is the principal benefit that has deterred most Florida producers from using the hemi routes. Because gypsum has two moles of water of hydration vs. one half mole for hemihydrate, and because product acid strength is 26–28% P_2O_5 normally, the ground phosphate feed slurry can be as low as 65% solids for the dihydrate process. This still leaves enough makeup water for adequate washing of the gypsum. For hemihydrate processes the maximum moisture is generally considered to be about 15–18% in the feed rock. Such a low moisture would call for a dewatering step for ground Florida pebble rock. Producers in Florida have thus far chosen not to go this route because of the difficulty and expense of dewatering.
2. Dihydrate gypsum, in the case of most phosphates, filters at relatively higher P_2O_5 throughput rates than hemihydrate, comparing dihydrate filter performance at 28% P_2O_5 and hemihydrate at 42% P_2O_5.
3. Dihydrate plants are proven at huge outputs, in excess of 2,000 mt per day P_2O_5, thus offering economies of scale.
4. A vast amount of operating data and experience exists on a wide variety of phosphate rocks for the dihydrate mode.
5. Capital costs for the dihydrate system can be less than for the two-stage hemi processes even when evaporation and rock grinding are considered, because of the larger reaction vessels and two stages of filtration in the hemi process.
6. The yield across the filter for the dihy drate process, generally about 96% of the P_2O_5 fed, is about 3–4% above the single-stage hemi processes.
7. Dihydrate process maintenance costs are substantially less than those for hemi processes due to less severe process conditions. The on-stream factor is also higher for the average dihydrate facility.

Major Dihydrate Processes

Prayon process. The Prayon process [17] has evolved from the cascade system first used by the Dorr Company in the 1930s. Instead of round, steel vessels with rubber and brick lining, Prayon developed a multicompartmented reinforced concrete vessel, rectangular in shape, lined with a membrane and carbon brick in the early 1960s. The construction has proved to be exceedingly durable. A similar concrete construction is used by Jacobs and by Rhone Poulenc.

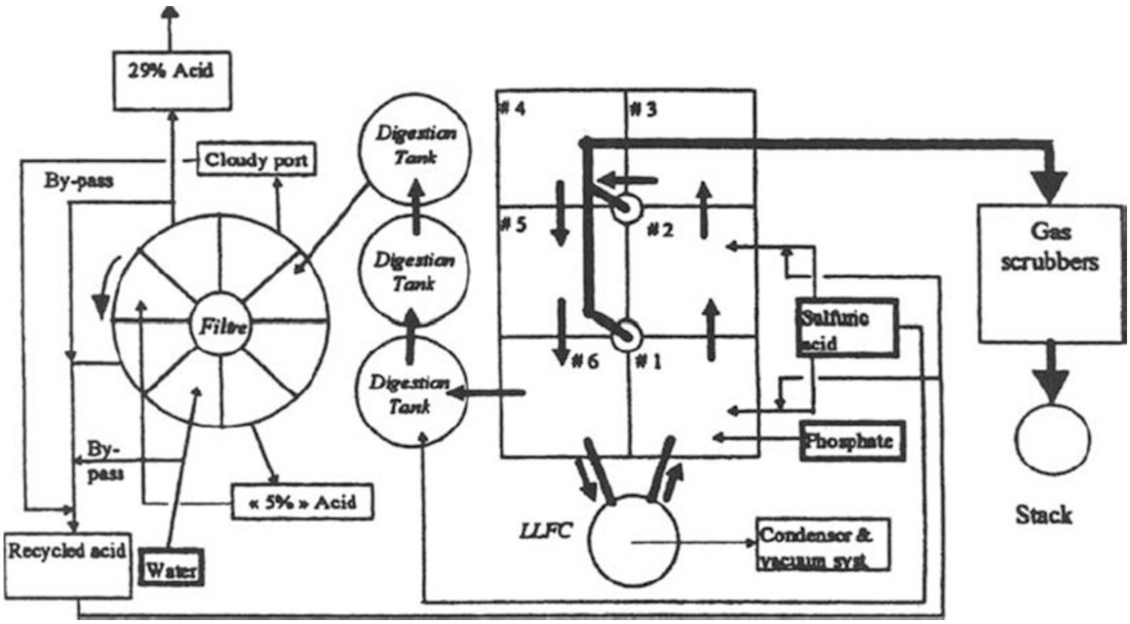


Fig. 23.9 Prayon Mark IV reaction and filtration system

A diagram of the Prayon Mark IV reactor configuration is shown in Fig. 23.9. The Prayon process uses vacuum cooling, a low-level vacuum chamber through which reaction slurry is circulated to maintain a reactor temperature of 80–85°C. There are numerous installations of large Prayon plants in the United States, although the center of process technology is Belgium.

Prayon also offers the Prayon–Central Glass process, a name derived from its Japanese origins. In this process, gypsum is crystallized in a dihydrate mode in the first stage. After separation of most of the P_2O_5 from the slurry by centrifugation or filtration, the gypsum is recrystallized to hemihydrate employing the sulfuric acid ultimately required in the first stage. The recrystallized hemi is quite low in P_2O_5 and impurities suitable for by-product wallboard and plaster, and is readily washed in a second stage of filtration, followed by an agglomeration step that results in a semi-granular by-product. This process has allowed the Engis, Belgium, plant to continue operation by converting all of its gypsum to a saleable product. The Prayon Mark III and Mark IV dihydrate processes are used in Florida and Louisiana in some eight lines operating between 1,350 and 2,000 t per day P_2O_5, and in one revamped plant in Jordan.

Speichim–Rhone Poulenc process. Most of Rhone Poulenc's existing plants are of a single stirred vessel configuration. However, a two-vessel arrangement, the Diplo system, [18] has been offered and retrofitted into several of the previous single reactor plants.

The original Rhone Poulenc Single Tank system is an exceedingly simple plant. It is air-cooled by passing a flow of air over slurry splashers, and does not have a filter feed tank, the slurry simply overflowing the reactor to a vertical pipe on top of the filter feed pump. The newer Diplo system is said to offer a better yield, higher P_2O_5 product acid strength, and a more filterable gypsum by providing a better concentration gradient for gypsum growth.

Rhone Poulenc installations are located in France, Belgium, Morocco, Senegal, Brazil, and China.

Nissan H process. In this process, the rock attack is done under conditions favoring the formation of an unstable hemihydrate. The slurry then is cooled and seeded to recrystallize to gypsum at high sulfate levels, producing a gypsum low in co-precipitated P_2O_5. Yields in the 97–98% range are reported. The process, in operation since the 1960s, has been favored in many instances where the phosphogypsum can be used for wallboard or other building materials. Major installations are located in Japan and Morocco. A large plant at Pernis, Netherlands, has been closed and it appears that Nissan no longer offers phosphoric acid technology.

Jacobs process. Jacobs Engineering acquired the Dorrco process technology in 1974 and has carried on the annular reactor design begun by Dorr in the early 1960s. The reactor configuration is a compromise between a multicompartment system, as used in the earlier Dorr–Oliver cascade system and in the compartmented Prayon reactor, and the true single stirred vessel used by Rhone–Poulenc and Badger.

In the Jacobs system, the annulus of a large concrete tank is fitted with a series of agitators. There is baffling, but there are no walls between the agitators. High slurry recirculation

Fig. 23.10 Oswal phosphoric acid plant during construction—May 1999—Paradeep, Orissa, India

rates are achieved by a combination of back-mixing and slurry pumping. Vacuum cooling normally is used. The Jacobs plants include a 1,500 t per day plant in Tampa, Florida, a 900 tons per day facility at Paradeep, India, and several smaller units. Figure 23.10 shows the Oswal Fertilizers and Chemicals construction site of the world's largest phosphoric acid plant [19–21] 2,650 mt per day P_2O_5 also at Paradeep, Orissa, in mid-1999. Figure 23.11 is a diagram of the reactor itself. Starting in the center foreground of the photo and moving clockwise, the large building, partially roofed, holds five 110 m two-belt filters, next is the wet grinding and screening structure. The concrete reactor tank, having 4,350 m^3 of slurry volume, is shown prior to the installation of the agitators. The three small cylindrical vessels are slurry vacuum coolers. Six evaporators with fluosilicic acid recovery are pictured next. The large tanks are for acid storage and clarification, and finally, on the far right, the cooling tower foundations. Slurry flow is clockwise around the annulus, into the center compartment, then to the long narrow vacuum cooler feed compartment. The cooled slurry flows to the cooler seal compartment opposite the feed compartment. The bulk of the slurry recycles to the annular reactor with the net flow proceeding to the filter feed tank. In principal, the flow pattern is similar to the Prayon reactor system shown in Fig. 23.9. The ground phosphate rock is mixed with recycled cool slurry in the first position of the annulus followed by the addition of sulfuric acid diluted in a mixing tee with weak phosphoric acid from the filter.

Isothermal process. The Badger–Raytheon reactor is a draft tube mixer within a vacuum vessel [22]. All reactants are added to this vessel, which is under vacuum, and cooling and rock digestion are achieved simultaneously.

There are three lines in Florida, two in Mexico, and two in the Western United States. The system has low energy consumption, but lacks flexibility in that it works best with finely ground rock. It has also exhibited some difficulty in handling the high organic Mexican phosphate.

Hemihydrate Processes for Phosphoric Acid

Methods for making higher strength P_2O_5 acid have been known for a long time. The basic hemihydrate–dihydrate process shown in the Hydro Fertilizer flowsheet, Fig. 23.12, is similar to the initial process attempted in 1931 at the Cominco plant at Trail, BC. The hemihydrate—dihydrate process failed there, mostly because of inadequate filters, but the Dorr dihydrate process did emerge successfully.

The first large hemi plant of more modern times was the Kemira hemi-dihydrate two-stage plant at Siilinjarvi, Finland, which started operation in 1969, using the Dorr HYS process. After about 4 years of operation at 250 mt per day P_2O_5, its rated design, the plant was expanded but operated in the dihydrate mode. In the hemihydrate mode, filtration and recrystallization problems had plagued the plant, which operated on Kola rock.

Hemihydrate processes are available for new facilities and also for the retrofitting of existing dihydrate plants. Several conversions to higher-strength acid have been made, where the steam saved in evaporation can replace fuel. Plants that make super-phosphoric acid, or where

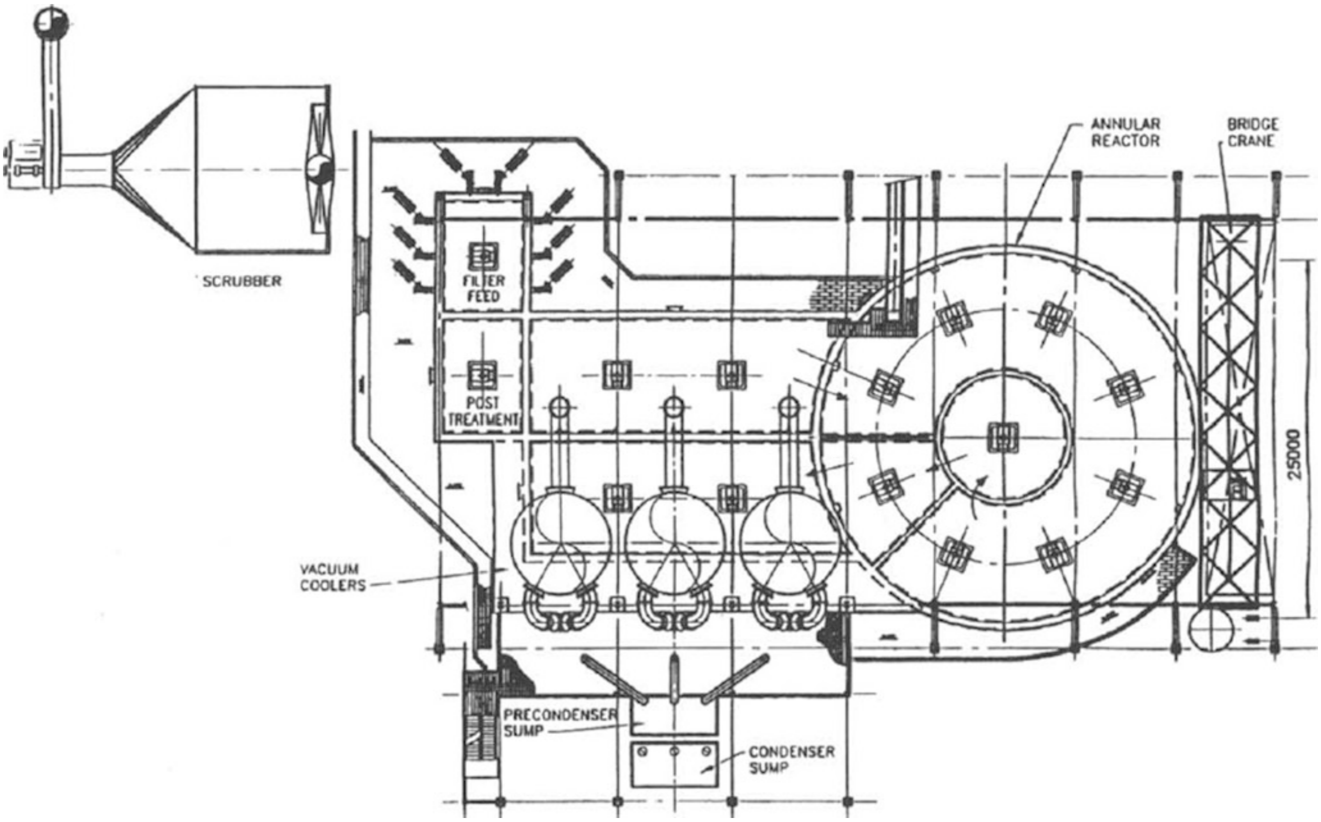


Fig. 23.11 Jacobson reaction system, 2500 tons per day P_2O_5

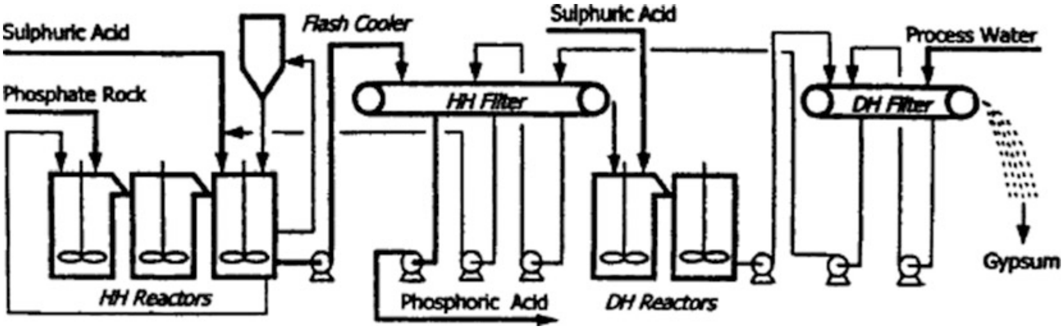


Fig. 23.12 Hydro fertilizer technology hemihydrate process

sulfuric acid plant steam is not available, are likely targets for conversion.

The single-stage hemi process is similar to the front end of the HDH process (see Fig. 23.13). However, in this case, the hemi is sent to waste without deliberate recrystallization after washing. Yields for the single-stage process are generally below the yields of the dihydrate process, whereas yields for the two-stage HDH process are very high; see Table 23.7.

Table 23.8 lists the major hemihydrate plants operating in 2000. Many of the smaller facilities on the ninth edition list have been shut down. However, two major single-stage hemi plants have recently started operation. The Indo-Jordan facility has been particularly successful operating at full capacity from the initial startup, and up to 1,000 t per day P_2O_5 on two 80-m^2 Eimco belt filters.

The WMC plant [23, 24] extends proven hemi capacity to 1,500 t per day P_2O_5, but the plant has had to contend with a highly abrasive, low-grade rock, 23.5% P_2O_5 with over 35% SiO_2. In addition, the PCS (Occidental) single-stage hemi plant, in operation since the early 1980s, has been termed by some as one of the best phosphoric acid plants in the world. Also, the conversion at Arcadian has met the test of time.

Fig. 23.13 Hydro fertilizer technology hemihydrate process

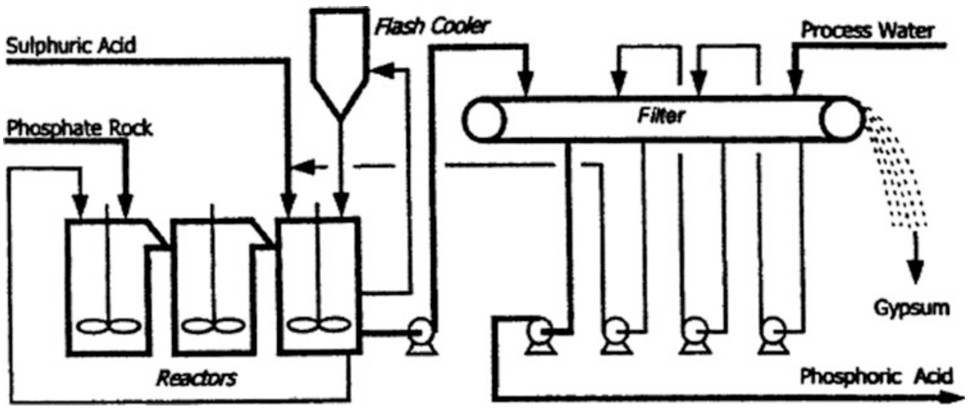


Table 23.7 Range of P_2O_5 yield through filtration as % of P_2O_5

Phosphate acid process	
Hemi-dihydrate	98–99
Conventional dihydrate	95–97
Single-stage hemi	91–94

The advantages of the hemihydrate processes compared with the conventional dihydrate processes are:

1. Energy savings due to higher product acid strength.
2. Higher P_2O_5 recovery and lower H_2SO_4 consumption for the two-stage process.
3. Lower capital cost for the single-stage process.
4. Hemi-hydrate recrystallized gypsum that is relatively pure, 0.2–0.4% total P_2O_5. This makes it more suitable for a cement additive or in wallboard.
5. Less rock grinding required.

Energy savings in phosphoric acid. As has just been noted, the hemi processes provide energy savings due to the reduced steam consumption for evaporation. This steam, even though of low-pressure, 2.0–3.0 kg/cm^2, has significant value for the cogeneration of power. In some cases, the hemi process also saves rock-grinding power. On the other hand, dihydrate processes have the ability to use wet rock slurries and to absorb, in the process, more contaminated water resulting from rainfall collected off the waste gypsum stacks. The latter advantage may be important to maintaining a zero water balance and eliminating costly effluent treatment. The use of wet grinding and slurry feeding eliminates the fuel and electricity consumed in drying the rock.

Another energy-saving option is the use of hot water instead of steam for evaporation. This is an alternative, energy-wise, to the higher-strength hemi processes. Hot water normally is available from the heat of absorption of the sulfuric acid plant. Strengths of 40–42% P_2O_5 can readily be achieved. A few commercial installations exist in Europe. A more expensive alternative is offered in the HRS sulfuric acid process by Monsanto, which converts the absorption heat to low-pressure steam.

Unit Operations

In addition to the reaction step discussed above, there are a number of other unit operations used in producing wet process phosphoric acid.

Calcination. Phosphate rock normally is used as a dry rock or in slurry form. However, in some cases, particularly where the raw phosphate is high in carbonaceous matter or it is desirable to have a clean acid, the rock is calcined. Also, in a few cases, the phosphate rock is calcined, the product slaked, and free lime separated as a beneficiation step. Calcination is energy intensive and produces a less reactive rock and, in some cases, a less filterable gypsum. Therefore, the use of calcination is diminishing, and is being replaced by a wet oxidation step to produce green acid [16]. In separating calcium carbonate, flotation, where it is successful, is favored over calcination because of its lower cost.

Rock grinding. Until 1973, most phosphate was ground dry in roller or ball mills. In that year, Agrico, at South Pierce, Florida, converted one of its dry mills to wet slurry grinding and proved that the plant water balance could manage the rock at a 65–68% solids slurry. Since that time, most US installations have converted to wet grinding.

Relatively fine phosphates, such as Kola, North Florida, Senegal, and Togo, can be processed unground as dry concentrates or as dewatered beneficiated product with 12–18% moisture. In the hemi processes, somewhat coarser feed, -20 mesh, may be tolerated. For dihydrate, it is desirable to feed from 1.5% + 35 mesh (Tyler) to about 8% + 35 mesh.

Filtration. The separation of phosphogypsum or hemihydrate from its mother liquor has always been a difficult operation. The process has been subject to the formation of calcium sulfate, sodium fluosilicate, and other types of scale that clog the cloth and necessitate periodic filter washing. Filter cloth wear is severe, requiring cloth changes as often, in some cases, as 2 or 3 weeks apart.

Table 23.8 Major hemihydrate installations (2000)

Owner	Location	Year in operation	Process	Product acid strength	Rock	Capacity P_2O_5 (million tons per day)
Chinhae	Korea	1990	N-H(C)(NDH)	45	FL	250
PCS (Arcadian)	U.S.	1980	N-H(C)(H)	40	Bou Cra	600
PCS (Occidental)	Florida	1980	Oxy(H)	38	FL	1400
Gresik	Indonesia	1984	Nissan C(HDH)	42	Jordan	550
Nam Hae	Korea	1988	Nissan C(C)(HDH)	42	FL/Jordan	1100
Yong Nam	Korea	1989	Nissan C(C)(HDH)	42	–	400
Copebras	Brazil	1987	Nissan C(C)(HDH)	–	Brazil	450
Coop Chem.	Japan	1987	Nissan C(HDH)	–	–	230
Yunnan	China	1992	N-H(HDH)	45	China	210
Inda-Jordan	Jordan	1998	N-H(H)	42	Jordan	750
WMC	Australia	2000	N-H(H)	42	Queensland	1500

N-H Norsk Hydro, *C* Conversion, *H* Hemihydrate, single-stage, *HDH* Hemihydrate–dihydrate

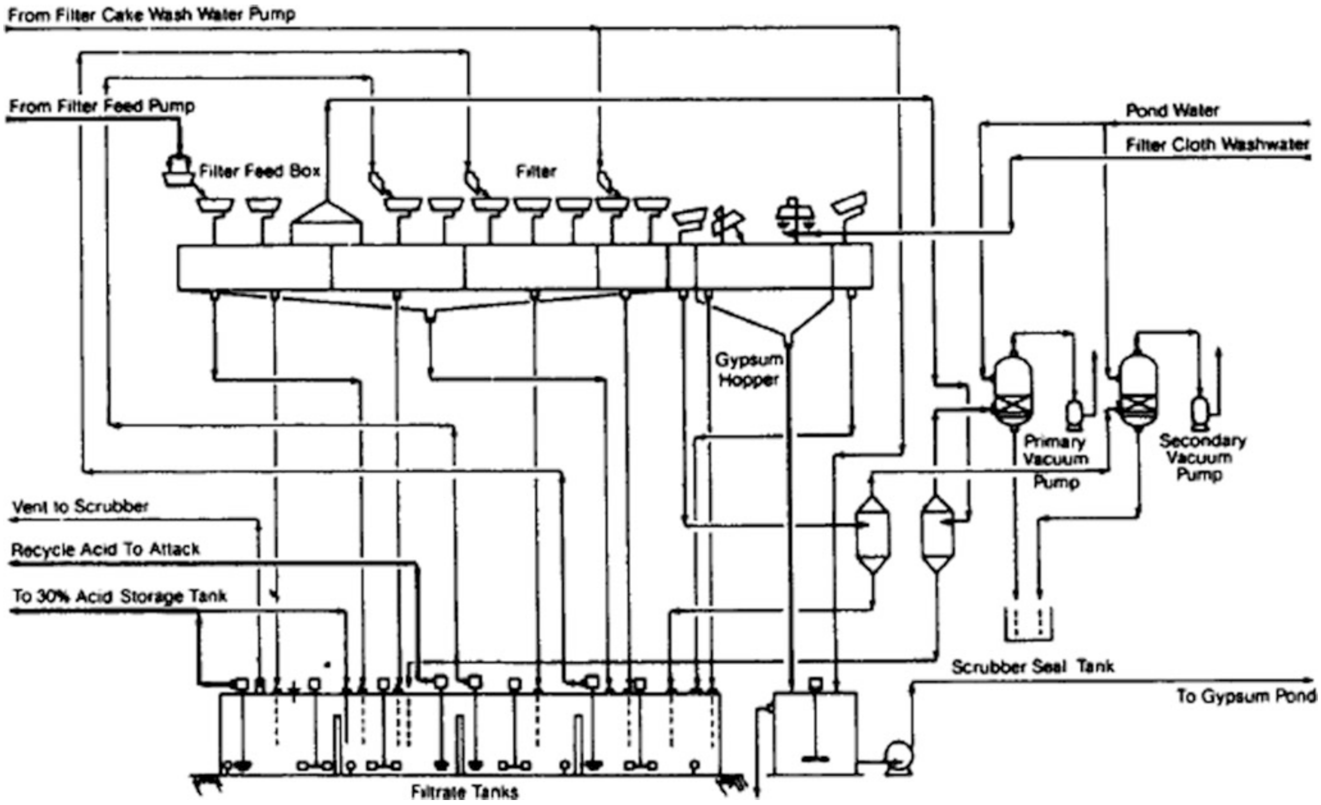


Fig. 23.14 Flow diagram of filtration section of a wet process phosphoric acid plant (Prayon process, courtesy Davy Mckee Corp.)

Three types of filters have predominated over the past 20 years, the most widely used being the Bird–Prayon tilting pan filter shown in Fig. 23.14. The Ucego, a table filter with a peripheral side wall belt that leaves the filter to permit cake sluicing, has been popular worldwide since the late 1960s. In the late 1970s and the 1980s, belt filters became more readily accepted. The belt filter has been used on phosphoric acid since the 1940s and 1950s, but in the past it was plagued by mechanical problems and materials failures.

During the last decade, many successful belt filter installations have been made by Eimco, Filtres Philippe, Delkor, and Gaudfrin. Even in hemi service at Namhae, Indo-Jordan, and WMC, operating at 95°C or above, the belt filters appear to be successful. Because of their long narrow configuration, the belt filter is well suited as a supplemental filter and three have been installed in Florida for this purpose.

The filtration step is a countercurrent washing using two or three washes. Usually the final wash is a contaminated

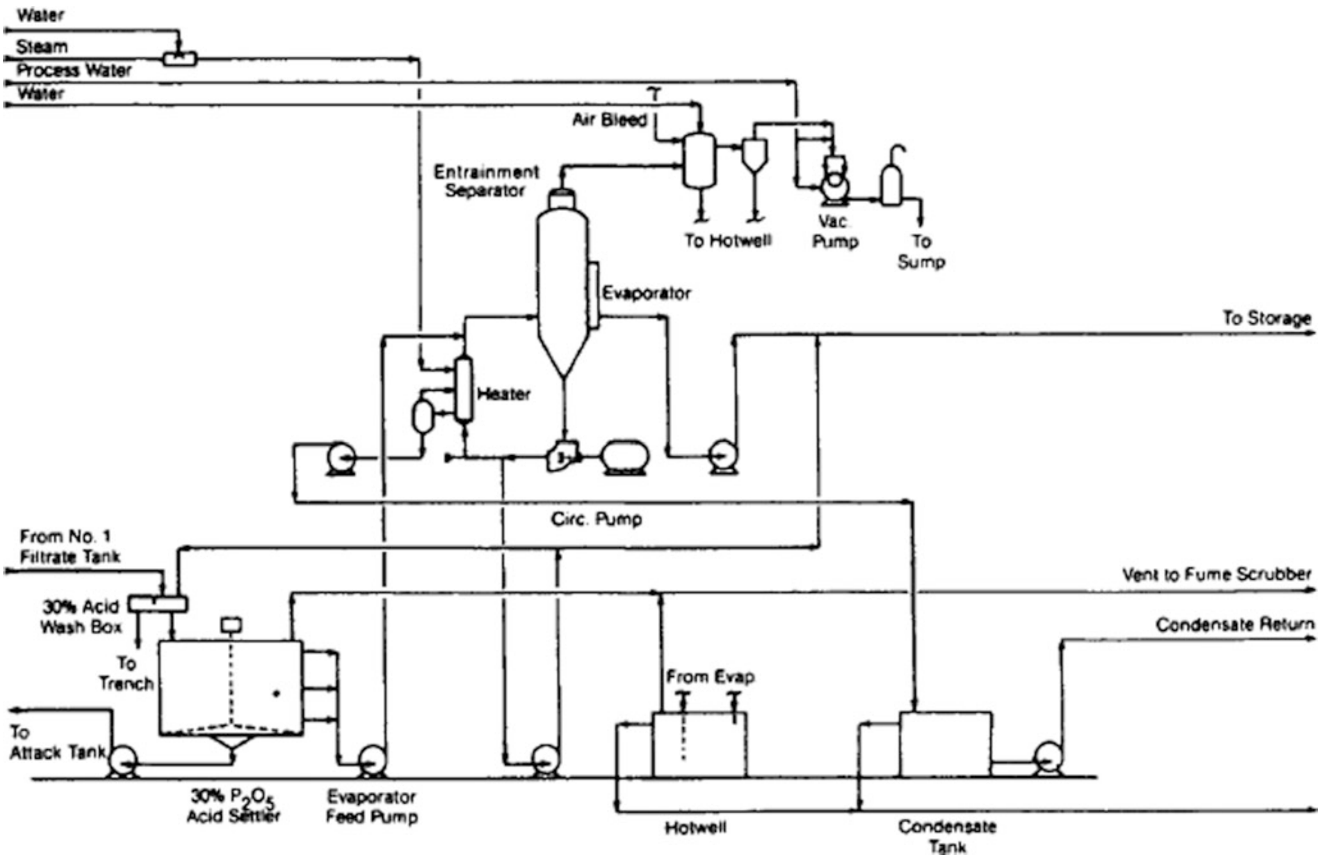


Fig. 23.15 Flow diagram of evaporation section of a wet process phosphoric acid plant (courtesy Davy Mckee Corp)

pond water or a cooling loop water, thus providing for, in most cases, a zero effluent plant.

Sizes of the Bird–Prayon and Ucego filter can be very large, over 200 m^2 of active surface area, allowing rates up to 1,600–2,000 mt per day P_2O_5. Belt filters are, so far, 110 m^2 or less; however, they are relatively inexpensive because little alloy steel is used, and normally two belt filters would be less in first cost than one large tilting pan or table filter.

Evaporation. Phosphoric acid is used for downstream products mostly at 28% P_2O_5, 40% P_2O_5, and 54% P_2O_5. Many plants also make clarified merchant grade acid (MGA), which, at 52–54% P_2O_5, is a world traded product.

Evaporation normally is done under vacuum using forced circulation via an axial flow pump. Vacuum can be accomplished by steam ejectors, vacuum pumps, or with condenser water using an entraining condenser downleg. Heat exchangers normally have been shell and a tube, with graphite tubes in the United States. French practice has been to use carbon block exchangers, but these tend to scale more and are no longer in favor in many areas. Some newer shell and tube exchangers have been Sanicro 28 or Hastelloy G3 or G30. A flow sheet is shown in Fig. 23.15.

Typical analyses of wet process phosphoric acid made from North Carolina calcined concentrate are shown in Table 23.9.

Table 23.9 Typical analyses, wet process phosphoric acid made from North Carolina calcined concentrate

	Weight percentage	
Component	Concentrated acid	Superphosphoric acid
P_2O_5, total	53.0	69.5[a]
Solids	0.2	
Free water	22.0	
Fe_2O_3	1.5	2.0
Al_2O_3	0.7	1.0
F	0.6	0.3
MgO	1.1	1.3
SO_4	2.7	3.7
CaO	0.1	0.2
Sp.g, at 75 °F	1.68	2.0

[a] About 36 % of the total P_2O_5 is present as polyphosphates (courtesy Texasgulf.Inc.)

Superphosphoric Acid

Solution fertilizers have become very popular in the United States. The principal source of P_2O_5 for these fertilizers is wet process Superphosphoric acid containing about 70% P_2O_5 where 35% or more of the P_2O_5 is present in the polyphosphoric form. When this acid is ammoniated and diluted, the iron, aluminum, and magnesium compounds naturally present remain in solution, sequestered by the polyphosphates. Clear solutions result, and there is no clogging of the sprays used for applying the fertilizer. When ordinary phosphoric acid is neutralized with ammonia, heavy sludges form, and the resulting solution is difficult to store and apply.

Superphosphoric acid is made by the additional concentration of clarified phosphoric acid in vacuum evaporators of the falling-film or forced-circulation type. High-pressure steam or Dowtherm vapor is used for heating. Corrosion is a problem, so the equipment is made from high alloy stainless steels. The acid is shipped in special insulated tank cars to the solution fertilizer plants, which are located close to the farm areas they serve. Organic matter contributes to sludge problems in making solution fertilizers; therefore, calcination of the phosphate rock used for making the acid is advantageous. Organics also may be removed by oxidation with nitric acid or ammonium nitrate [16], and several SPA producers have gone to such treatment to improve their product quality.

Wet Process Acid By-Products

Phosphogypsum

About 5 t of gypsum on a dry basis are made for each ton of P_2O_5 produced in a wet process phosphoric plant. This material usually is disposed of as a waste, by impounding it in old mine pits, stacking it in huge piles, or, in some cases, discharging it into very large rivers or river mouths. Phosphogypsum is sold to farmers in California for control of salt buildup in irrigated soils; a small quantity is sold to peanut farmers in the southeastern United States. However, there has been concern about the utilization of gypsum because of its low-level radioactivity in some instances. In addition, it contains fluosilicates and P_2O_5 so that utilization in building products, such as wallboard and blocks, has been limited to Europe and Japan, where natural gypsum is more costly than manufactured gypsum. In these cases, the phosphogypsum generally comes from a hemihydrate process producing a slightly purer form than natural gypsum.

Regeneration of SO_2 from gypsum has been done via the Krupp–OSW process, the most important facility formerly being the Fedmis plant in South Africa, which made about 300 t per day H_2SO_4 and a similar amount of cement. Apparently, because of the high cost of cement and sulfur in this remote location, the plant was not economically viable and shut down in the late 1980s. Because of the present low cost of sulfur and its future abundance as a result of the removal of sulfur from sour natural gas and because of SO_2 removal processes in power generation, interest in sulfur recovery from phosphogypsum has waned.

The Florida Institute of Phosphate Research (FIPR) has studied phosphogypsum utilization at length, but the U.S. Environmental Protection Agency has prohibited its movement from its stacks (piles) because of its low level radiation. The FIPR hopes to show that its use in road building subsurfaces and as an additive to enhance digestion of municipal waste in land fills can be accomplished without harm to the public now or in the future [25, 26].

Fluorine Recovery

Phosphate rock contains about 3.5% fluorine, some of which is recovered as a byproduct in manufacturing wet process phosphoric acid. During acidulation, the fluorine is released as hydrofluoric acid, HF, which reacts with the silica present as an impurity in the rock to form fluosilicic acid, H_2SiF_6. Some of the fluorine is lost with the gypsum as sodium or potassium fluosilicates, and some remains dissolved in the filter acid. When the acid is concentrated, much of the fluorine in the feed is boiled off, appearing as HF and silicon tetrafluoride, SiF_4, in the vapors.

Fluorine is recovered at the evaporator station by scrubbing the vapors leaving the flash chamber. The vapors pass through an entrainment separator to remove fine droplets of phosphoric acid and then into a spray tower where they are scrubbed with a weak solution of fluosilicic acid according to the reaction:

$$2HF + SiF_4 \rightarrow H_2SiF_6$$

Part of the circulating solution is continuously withdrawn as a 20–25% aqueous solution of H_2SiF_6. The solution is shipped in rubber-lined tank cars and is used for fluoridation of drinking water, the preparation of fluosilicates, and production of AlF_3. These salts find use in ceramics, pesticides, wood preservatives, concrete hardeners, and aluminum production.

Uranium Recovery from Wet Process Phosphoric Acid

Uranium recovery was briefly described in the ninth edition. Since 1990, all uranium recovery contracts have expired in the United States and the recovery facilities moth-balled or scrapped. There is no indication that the situation will change in the near future, because the reduced price of uranium no longer makes its recovery from phosphoric acid economical.

Animal Feed Supplements

Calcium phosphates for use in animal and poultry feeds are made from both furnace and wet process phosphoric acids. Dicalcium phosphate, $CaHPO_4$, containing 18.5% P, and

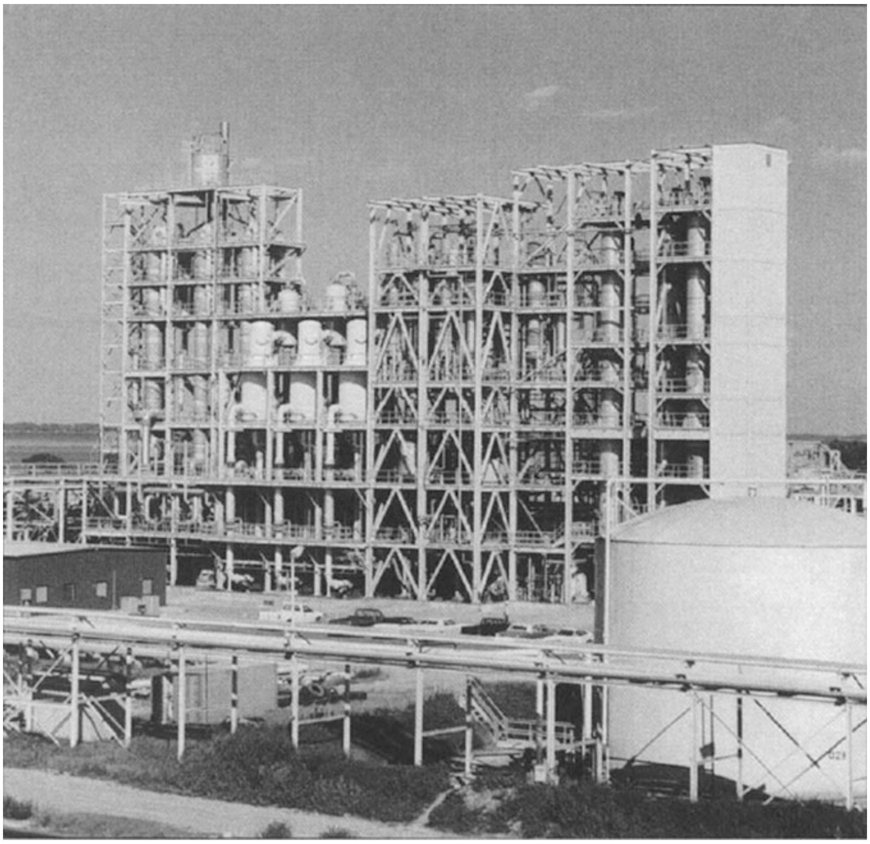


Fig. 23.16 Purifried phosphoric acid plant for P.A. Partnership, Aurora, NC (courtesy P.A. Partnership and Jacobs Engineering)

mono calcium phosphate, Ca $(H_2PO_4)_2 \cdot H_2O$, containing 21.0% P, are made in large tonnages. Both grades are prepared by reacting phosphoric acid with pulverized lime stone in a pug mixer. The limestone must be quite pure, and the phosphoric acid must have a low fluorine content, below 100 P to 1.0°F. If 54% P_2O_5 wet process phosphoric acid is used, it is defluorinated first by adding diatomaceous earth and then sparging the acid with steam. An alternate method is to use wet process superphosphoric acid, which has a low fluorine content. The superphosphoric acid is hydrolyzed by diluting it with water and heating. The pug mixer product is a fine granule, −12 mesh, which is dried and shipped in bulk to feed-mixing plants.

Purified Phosphoric Acid

Currently only one furnace acid plant remains in operation in the United States to supply elemental phosphorus and a few uses that cannot be satisfied by purified wet acid produced from wet process acid.

Wet process phosphoric acid made from calcined rock is preferred feed stock because it is devoid of the soluble organics and sludges present in acid made from sedimentary phosphates. However, calcination is expensive so that some plants are willing to go through laborious clarification to avoid it.

Clear acid is fed to a column or a battery of mixer-settlers and extracted with a solvent such a butyl alcohol or tributyl phosphate. Generally, about three-fourths of the phosphoric acid transfers to the organic phase, leaving the impurities in the raffinate, which is sent to a fertilizer unit to recover its P_2O_5. The yield of cleaned acid can be increased by adding another mineral acid such as sulfuric acid or hydrochloric acid to the extraction step.

After washing, the phosphoric acid is stripped from the solvent with water, and the solvent is returned to the extraction section. The phosphoric acid now is quite dilute and still contains small amounts of impurities. The acid then is concentrated, and the impurities are removed by steam stripping and the addition of reagents and adsorbents followed by filtration. The exact details of the process vary, depending upon the process technology, which is proprietary, the impurities present in the feed acid, and the solvent used. Purified wet phosphoric acid is suitable for both industrial and food-grade use, although food-grade requires another level of purification over industrial or technical grade acid.

There are now several purified acid plants in the United States. Figure 23.16 is a photo of the PCS plant at Aurora, NC. That plant was expanded by adding a third train in 2001 and a fourth train in 2006. The Rhone-Poulenc purification plant at Geismar, LA continues in operation. There are also plants in Morocco, Belgium, Japan, and Israel.

Environmental Aspects

As implied earlier in this chapter, gypsum disposal is a problem that generally has defied an inexpensive solution. As a waste material, it is relatively benign, but P_2O_5, sulfate, fluorine, low-level radioactivity, and other contaminants, including some heavy metals in small quantities, can leach from waste stacks into the nearby groundwater. In Florida, the underlying layers of limestone afford some protection, but the EPA will require future stacks be lined with a membrane to prevent seepage. The "gypsum dilemma" has loomed as an increasing problem over the last decade [25–27].

Although dumping of waste gypsum slurries into the ocean still is practiced in some locations outside the United States, and harmful effects are generally difficult to quantify, there is continuing pressure from environmental groups to cease dumping into rivers and the seas.

Gaseous emission from phosphoric acid plants can be scrubbed with cool contaminated recycle cooling water to relatively low emission levels. However, to minimize cooling tower or cooling pond emission, fluorine recovery often is necessary on those streams that have a significant fluorine content.

Acknowledgement The author acknowledges Dave Leyshon's contributions to the 10th edition of this chapter, which have been retained in this edition.

References

1. Leyshon DW (1990) The origin of the modern dihydrate phosphoric acid process—cominco 1931. A.I.Ch.E, Orlando
2. McClellan GH (1980) Mineralogy of carbonate fluorapatites. J Geol Soc Lond 137:675–681
3. Van Kauwenbergh SJ, Cathcart JB, McClellan GH (1990) Mineralogy and alteration of the phosphate deposits of Florida. US Geol Surv Bull 1914:46
4. Bernardi IP, Hall RB (1980) Comparative analysis of the Central Florida phosphate district to its southern extension. Mining Eng 32 (8):1256–1261
5. McKelvey VE (1967) Phosphate deposits. Geo Surv Bull 1252-D: D1–D21
6. Northolt AJG, Shelton RP, Davidson DF (1990) Phosphate deposits of the world, 1st edn, Phosphate rock resources. Cambridge University Press, Cambridge
7. Northolt AJG, Shelton RP, Davidson DF (1990) Phosphate deposits of the world, 2nd edn, Phosphate rock resources. Cambridge University Press, Cambridge
8. Zellars–Williams Inc (1978) Evaluation of the phosphate deposits of Florida using the minerals availability system, Final report submitted to the U.S. Department of the Interior Bureau of Mines, June 1978
9. Busnardo CA, Olivario RN (1985) "Optimization of the grinding circuit of the Jacupiranga Carbonatite ore in Jacupiranga", Brazil. In: SME/AIME annual meeting, New York, p. 85–98
10. Nofal AM (1993) Egyptian phosphate rocks: important factors affecting thek economic and technical evaluation. In: El-Shall H, Moudgil BM, Wiegel R (eds) Beneficiation of phosphate: theory and practice. Society for Mining, Metallurgy, and Exploration, Littleton
11. Pressacco R (2001) Overview of the Agrium Kapuskasing phosphate operation. CIM Bull 94(1049):55–58
12. Allen MP (1993) The Vernal phosphate rock mill. In: El-Shall H, Moudgil BM, Wiegel R (eds) Beneficiation of phosphate: theory and practice. Society for Mining, Metallurgy, and Exploration, Littleton
13. Glenn A. Gruber, Charles Y. Guan, Michael E. Kelahan (2001) Dolomite flotation-pilot plant studies. Benefication of Phosphates Fundamentals and Technology, Society for Mining, Metallurgy, and Exploration, Inc. ISBN 0-87335-225-4
14. Lawver JE (1985) General principals and types of electrostatic separators. In: Weiss NL (ed) SME mineral processing handbook. Society of Mining Engineers, New York
15. Becker P (1989) Phosphates and phosphoric acid, 2nd edn. Marcel Dekker, New York
16. Leavith, Kranz, Gorman, and Stewart, U.S. Patent 4808,391
17. Theys T, Smith P (June 2002) IMACID, a 1000 ton phos acid unit in Morocco. A.I.Ch.E, Clearwater
18. Satier B, Apostoleris G (1992) Speicbjrn/phone poulenc process. IFA Technical Conference, The Hague, Netherlands, Oct 1992
19. Blythe BM, Leyshon DW, Jaggi TN (1998) Inception of the world's largest phosphoric acid plant. IFA Technical Conference, Marakesh, Morocco, Oct 1998
20. Leyshon DW (1999) Phosphoric acid technology at large, part II. Phosphorus & Potassium, Nov/Dec 1999
21. Davies L (2000) Strength in diversification. Phosphorus & Potassium. May/June 2000
22. Felice C, Martinez J, Hilakos S (1998) Raytheon's isothermal reactor process. A.I.Ch.E, Clearwater
23. Collen D, Duckworth G (2000) Ammonium phosphate plant makes its debut. Phosphorus & Potassium. March/April 2000
24. Gobitt J (2000) A new chapter in hemihydrate technology. A.I.Ch.E, Clearwater
25. Leyshon DW (1996) The gypsum dilemma. Phosphorus & Potassium. March/April 1996
26. Leyshon DW (2001) The gypsum dilemma, new concerns. Phosphorus & Potassium. March/April 2001
27. Wissa AEZ, Fuleihan NF (2000) Phosphogypsum stacks and ground water protection. Phosphorus & Potassium. May/June 2000

Fertilizers and Food Production

24

Amit H. Roy

Introduction

Fertilizers provide plants with the nutrients they need for their growth and development. Plants live, grow, and reproduce by taking up water and nutrients, carbon dioxide from the air, and energy from the sun. Apart from carbon, hydrogen, and oxygen, which collectively make up 90–95% of the dry matter of all plants, other nutrients needed by plants come essentially from the media in which they grow—essentially in the soil. The other nutrients are subdivided into primary nutrients (nitrogen, phosphorus, and potassium) and secondary nutrients (calcium, magnesium, and sulfur). In addition, plants also need other nutrients in much smaller amounts, and they are referred to as micronutrients (boron, chlorine, copper, iron, manganese, molybdenum, and zinc).

To maintain soil fertility and productivity and prevent land degradation, nutrients taken up by crops must be replenished through the application of fertilizers. The use of fertilizer results in many benefits to producers, consumers, and the environment, starting with increased agricultural outputs (mainly food and fiber) to contributing to soil organic matter maintenance, water-holding capacity, biological nitrogen fixation, soil erosion control, other physical and chemical properties, and less extensive land use. These benefits contribute to increased agricultural growth and agribusiness activities, which are catalysts for broadly based economic growth and development in most developed and developing economies; agriculture's links to the nonfarm economy generate considerable employment, income, and growth in the rest of the economy.

A commercial fertilizer is a material that contains at least one of the plant nutrients in chemical form that, when applied to the soil, is soluble in the soil solution phase and assimilable or "available" by plant roots. Most often, this implies chemical forms that are water soluble. However, in the case of phosphorus, solubility in special reagent solutions (citric acid, neutral ammonium citrate, or alkaline ammonium citrate) often is used as a guide for availability to plants. In the case of nitrogen, slow solubility in water may be more desirable from an environmental and efficiency standpoint than easy solubility.

Fertilizer products are customarily designated by a series of numbers separated by dashes. This set of numbers is called "grade" of the fertilizer product. Each of the numbers indicates the amount of a nutrient that the manufacturers guarantee is contained in the fertilizer product. This number includes only the amount of nutrient found by accepted analytical procedures, thereby excluding any nutrient present in a form that is deemed to be unavailable for plant nutrition. The content of each nutrient, expressed as a percentage of total weight, is the guaranteed minimum rather than actual, which is usually slightly higher.

Usually, three numbers are used when giving the grade fertilizer product, and these three numbers always refer in order to the content of the primary nutrients: nitrogen, phosphorus, and potassium. If other nutrients are present, their content can also be indicated in the grade of the fertilizer product; each extra number is followed by the chemical symbol of the nutrient it represents. Many countries indicate the content of phosphorus and potassium not in the elemental form but in the oxide form, P_2O_5 and K_2O. Thus, a fertilizer product with a grade of 12–6–22– 2MgO is guaranteed by the manufacturer to contain: 12% N, 6% P_2O_5, 22% K_2O, and 2% MgO.

The beginning of our dependence on inorganic fertilizer can be traced back to the nineteenth century when Justus von Liebig articulated the theoretical foundations of crop production and when John Bennett Lawes began producing fertilizers containing phosphorus [1]. However, only since the 1960s when global starvation became a real possibility have fertilizers assumed a predominant role in increasing agricultural productivity. Fertilizer was an integral part of

A.H. Roy (✉)
Int'l. Fertilizer Development Ctr., TVA Reservation Road,
PO Box 2040, Muscle Shoals, AL 35661, USA
e-mail: aroy@IFDC.org

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_24, © Springer Science+Business Media New York 2012

technological trinity—seed, water, and fertilizer—responsible for bringing about the "Green Revolution" that helped many densely populated countries, including India, China, and Indonesia, achieve food self-sufficiency in a short span of 20–25 years.

Since the 1960s, global cereal production has more than doubled, increasing from 866 million metric tons (mt) in 1961 to nearly 2,433 million mt in 2010, with developing countries accounting for nearly 60% of the increase (Fig. 24.1). Most of the increased cereal production in South Asia was through higher yields, but increases in sub-Saharan Africa were mostly through area expansion (Fig. 24.2). It is estimated that fertilizers accounted for 56% of the rise in average yields per hectare and about 30% of the total increase in production. Consequently, cereal production closely parallels fertilizer use in developing countries (Fig. 24.3). The only exception is sub-Saharan Africa where per capita food production has decreased since the 1970s. This decline can be attributed to several factors, including low soil fertility, agroclimatic conditions,

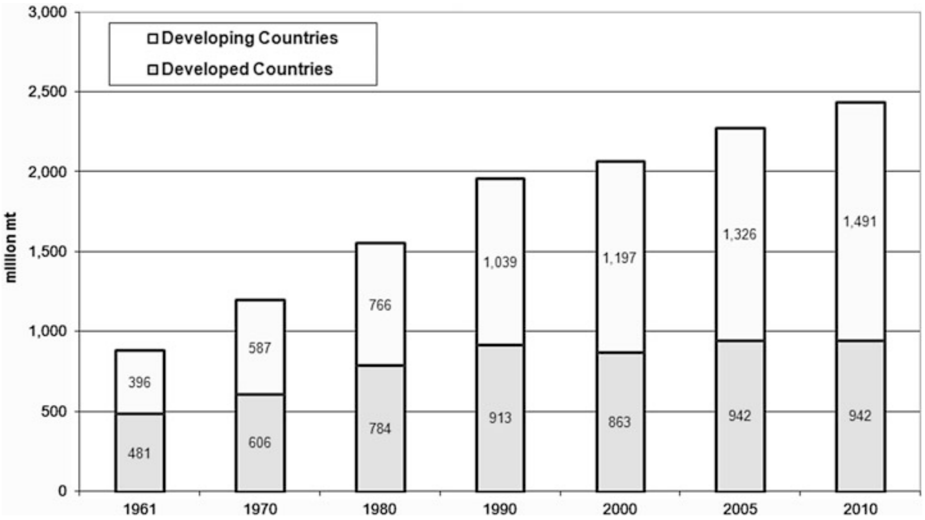
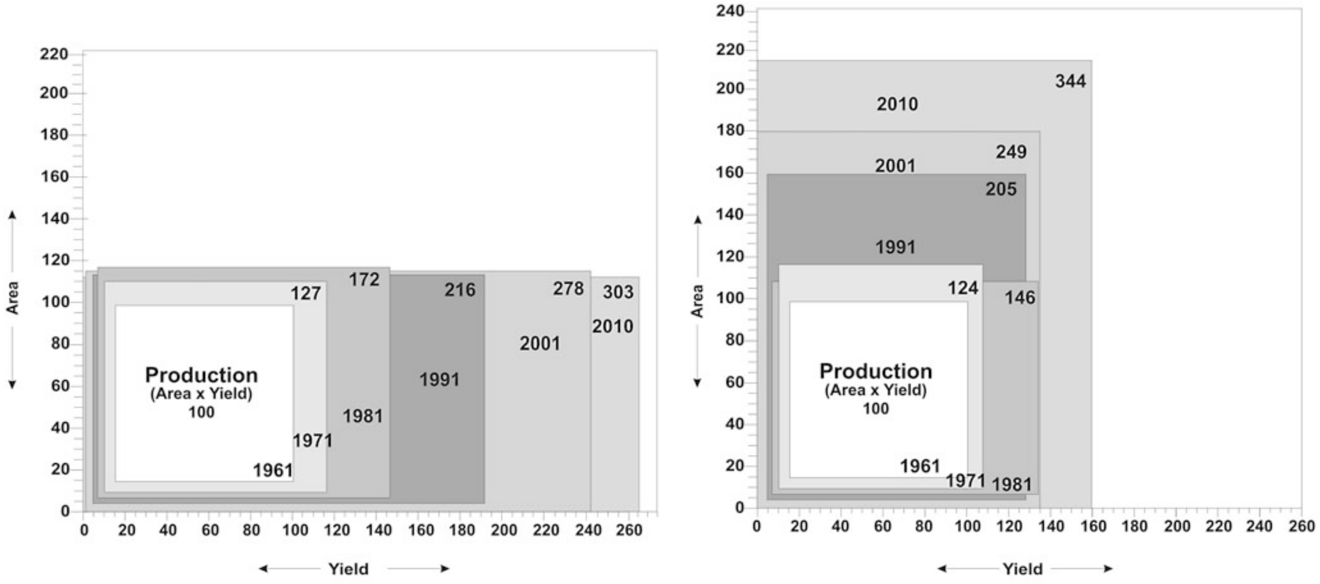


Fig. 24.1 Cereal production in developed and developing countries, 1961–2010



Source: Derived from FAOSTAT data

Cereal Production in South Asia, 1961-2010

Source: Derived from FAOSTAT data.

Cereal Production in Sub-Saharan Africa, 1961-2010

Fig. 24.2 Growth of cereal yield and area in cereal production in South Asia and Sub-Saharan Africa between 1961 and 2010 (1961 = 100 for yield and area)

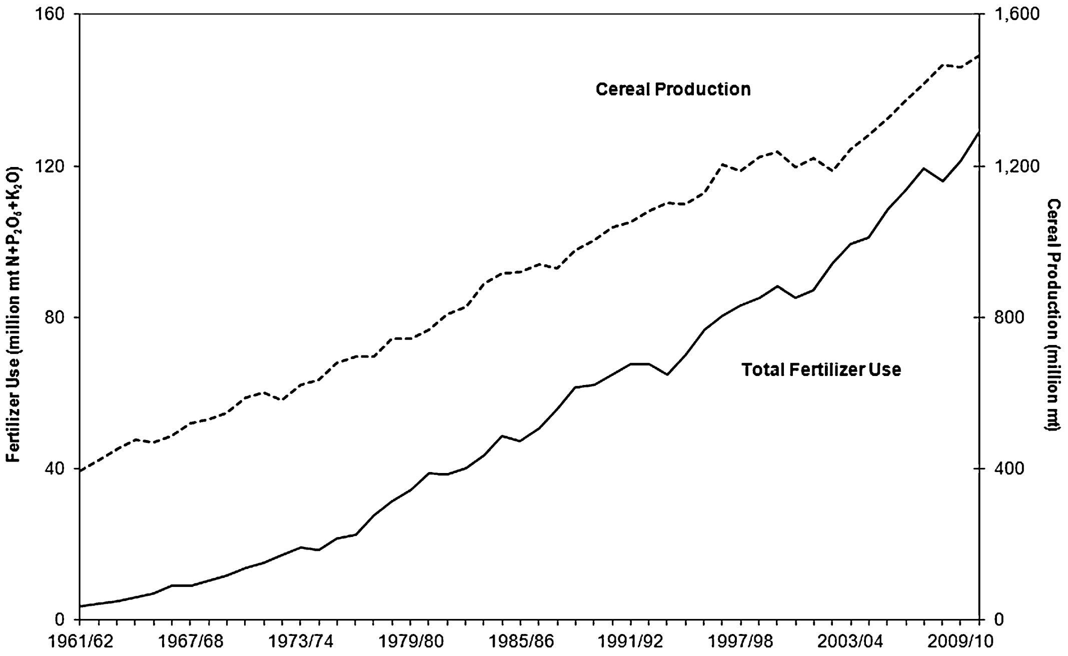


Fig. 24.3 Developing countries: total cereal production and total fertilizer use, 1961/1962–2010/2011

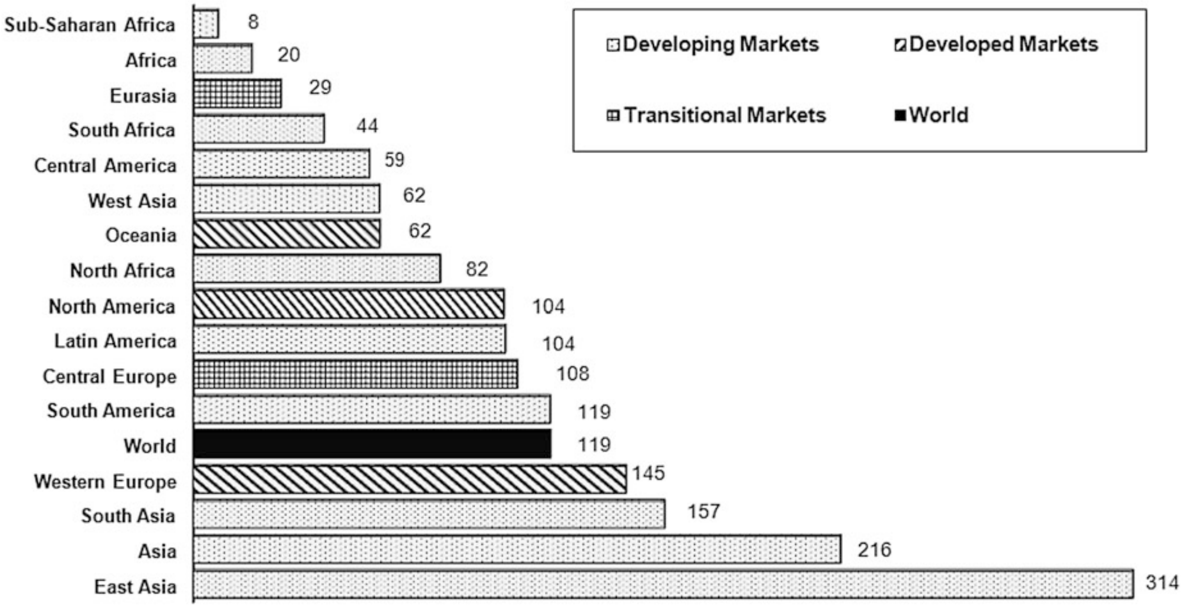


Fig. 24.4 Per hectare fertilizer use by regions and markets, 2010 (kg/ha)

and low fertilizer use. Compared with a world average of 119 kg of nutrients per hectare, sub-Saharan Africa uses only 8 kg of nutrients per hectare, resulting in a significant mining of inherent nutrients from soils that are already low in nutrient status (Figs. 24.4 and 24.5).

Growth in the future demand for food and feed will be influenced by changes in a number of forces, but mainly by growth in population, income levels and economic development, and changes in the food preferences of consumers. Although the United Nations recently adjusted downward its estimates of population growth, still during the next 20 years approximately 75 million people will be added to the world's population each year, increasing it from about 6.8 billion in 2010 to over 9.2 billion by 2050 (Fig. 24.6). Much of this increase will occur in the cities of the developing world, which will account for most of the increased

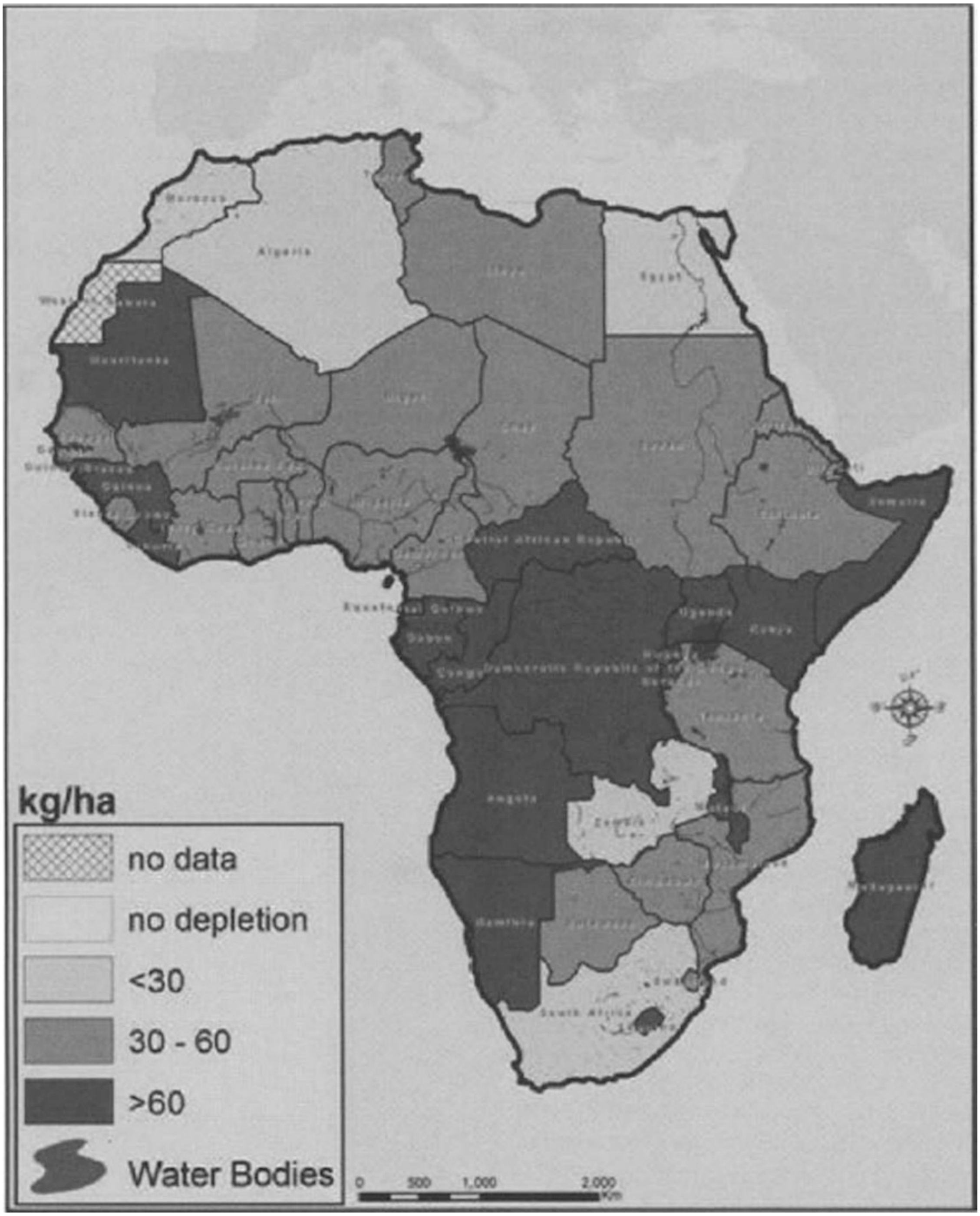


Fig. 24.5 Average nutrient depletion (NPK) in Africa for 2004

demand for food, including meat products. The recent drive toward renewable energy sources, particularly biofuels, will also increase demand. Enthusiasm for biofuels is because of dramatically higher energy prices, geopolitical tensions, and uncertainties about future availability and access to petroleum. Also, farmers must produce more crops from existing farmland if forests and wildlife habitats are to survive. This will require increased crop yields through use of improved seeds and fertilizer. This increased fertilizer use has to be balanced against the environmental and human health concern stemming from intensive fertilizer applications, particularly in industrialized countries.

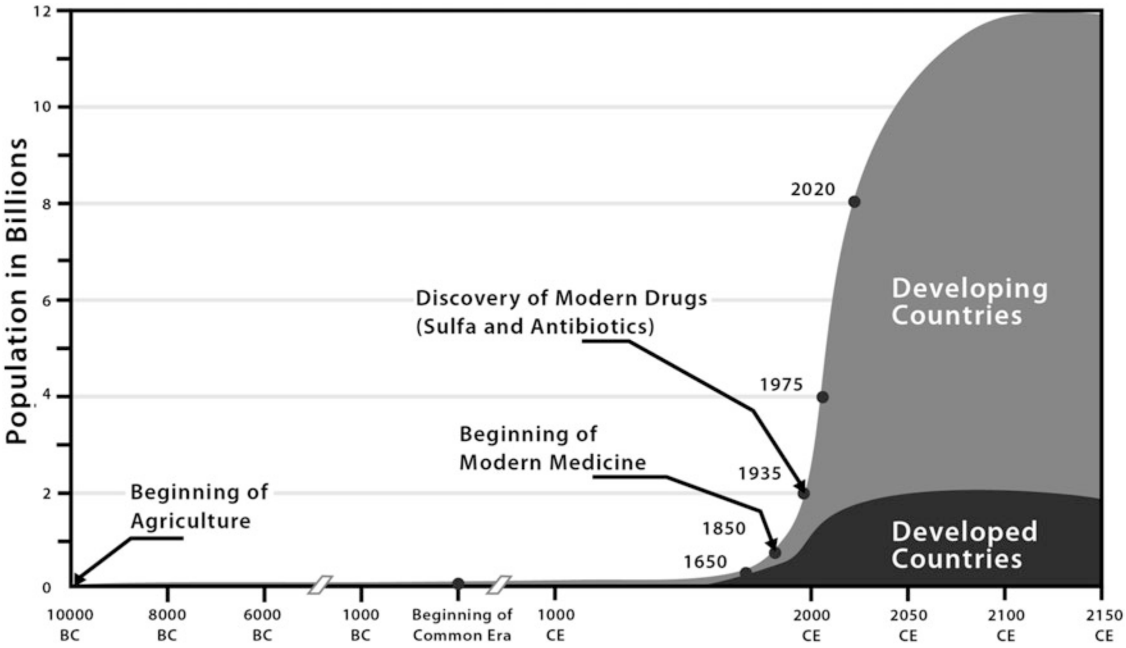


Fig. 24.6 World population projection

Overview of the Fertilizer Industry

Since medieval times, farmers have realized the need to maintain the productivity of soil to achieve improved crop yields. Until the last 200–300 years, however, the approach was highly empirical; only by accident or by trial and error was it found that applications to the soil of various organic wastes or naturally occurring mineral substances such as manure, compost, fish, ashes, saltpeter, and other substances would sometimes increase yields or apparently restore productivity to fields that were considered "worn out."

As more and more chemical elements were identified, scientists became interested in determining the amounts and relative importance of various elements in plants. The German scientist Liebig clarified the value of elements derived from the soil in plant nutrition and stressed the necessity of replacing those elements to maintain soil fertility. He usually is credited with initiating the fertilizer industry. Liebig recognized the value of nitrogen but believed that all plants could get nitrogen from the air, a concept that unfortunately is not true. He envisioned a fertilizer industry with nutrients such as phosphate, lime, magnesia, and potash prepared in chemical factories. In 1840 Liebig published a recommendation that pulverized animal bones be treated with sulfuric acid to make the phosphate more readily available to plants. This practice was accepted, and the production of fertilizers by chemical processing thus began.

Natural organic materials and various chemical by-products represented a large proportion of the total world fertilizer supply until about the middle of the twentieth century; in the later years of the century, however, the dependence shifted almost entirely to synthesized or chemically processed materials. Only by this means has it been possible to keep up with increasing populations, increased farm acreage, and increased plant food needs of new and improved crop varieties. Today, the fertilizer industry utilizes many facets of highly sophisticated chemistry and engineering in the manufacture of fertilizers, and in pollution control, including disposal of waste products. The procuring and handling of raw materials and the distribution and marketing of products also involve the latest technology and many innovations. In addition, the agronomic aspects of fertilizer usage engage the efforts of many individuals and organizations worldwide. The development and introduction of "high-yielding" varieties of seed required more concentrated and efficient fertilizers. Many organizations and private companies worldwide have carried out extensive research and development R&D on improving fertilizer production, developing new products, and increasing efficiency of usage. The most notable among them include the U.S. Tennessee Valley Authority, the International Fertilizer Development Center, and the Rothemstadt Experiment Station.

The fertilizer industry is a large-volume producer, one that falls within the basic industrial structure of a country. The magnitude of this industry is apparent from Fig. 24.7, which shows the yearly global consumption of primary plant nutrients (N, P_2O_5, and K_2O) since 1970/1971. For 2010/2011 the total consumption of N, P_2O_5, and K_2O was

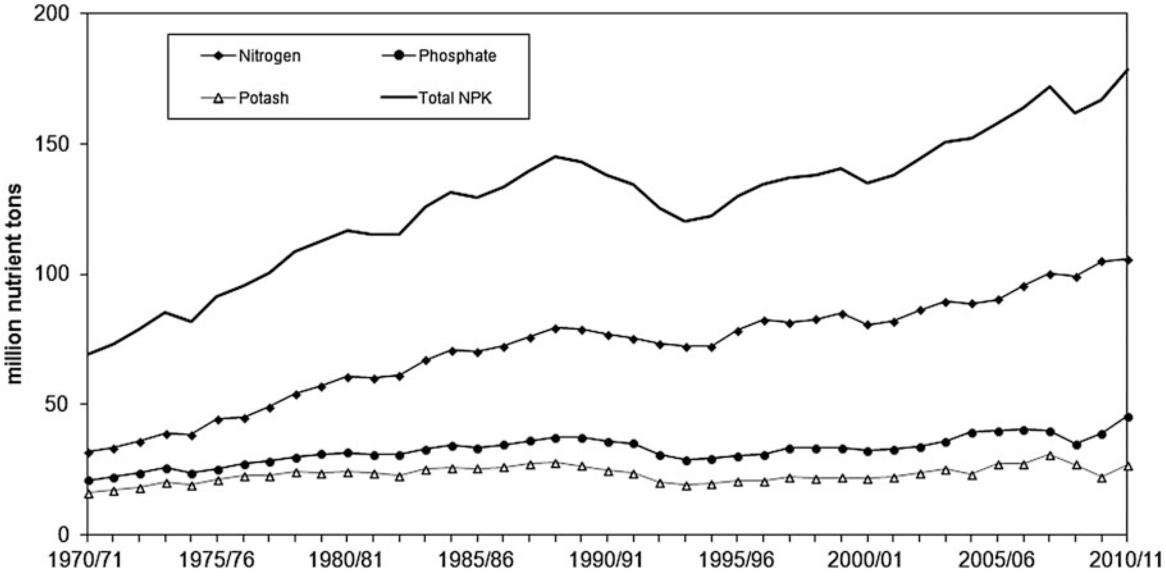


Fig. 24.7 World: nitrogen, phosphate, potash, and total NPK consumption, 1970/1971–2010/2011

178.4 million mt. Of this amount, China, the United States, and India consumed 57.4 million mt, 19.7 million mt, and 28.1 million mt, respectively. Collectively, these three countries accounted for about 50% of the world consumption. Assuming an average nutrient content of all fertilizers for that period of about 40% shows that some 358 million mt of fertilizers was produced and handled by the industry during that year [2]. In general, the industry includes a number of "basic producers" each of which typically concentrates on producing large amounts of single-nutrient or high-nutrient fertilizer products, usually at locations near raw materials. Numerous other components of the industry concentrate on locating close to local farmers and distributing the basic products either in the form received from basic producers or after various secondary-processing operations such as granulation, blending, or conversion to fluids. In the United States, fluid mixtures account for more than 20% of the total fertilizer sales. To a lesser extent, fluid mixtures have been introduced into Europe, North America, and South America. Fluid mixtures are used on high-value crops in some middle eastern countries, particularly Israel, where efficient use of water is as important as the efficient use of plant nutrients.

Raw Materials for Fertilizer Production

The primary raw materials for nitrogen fertilizers are natural gas, naphtha, fuel oil, and coal. The manufacturing of phosphate fertilizers most often requires phosphate rock. Naturally occurring potassium salts form the basis of the production of most potash fertilizers.

Natural gas, naphtha, fuel oil, and sulfur have fairly definable specifications. In contrast, phosphate rock and coal are products that can vary significantly in composition and other characteristics. These variations can affect the processes used to upgrade the "as-mined" ores or the processes for manufacture of fertilizers from beneficiated products. Potash ores also vary greatly in composition depending on origin; however, the end products of mining, beneficiation, and processing generally have relatively constant compositions.

Adequacy of requisite raw materials is the most obvious concern when facing a substantial increase in future demand. Two separate yet intertwined issues in the case of fertilizers are: sufficiency of raw materials and availability of energy to convert them into final products. Potassium is of least concern among the three primary nutrients. Not only is this element abundantly present in the earth's crust, but also it can be found in conveniently concentrated deposits in both deeply buried and near-surface sediments. Potassium deposits in descending order of known reserves are in North America (Canada and the United States), Germany, Russia, Belarus, Brazil, Israel, and Jordan. Even the most conservative reserve base estimates indicate a reserve on the order of 500 years at the level of the late 1990s production [3]. Mining, crushing, and beneficiation of potassium usually amount to about 10 gigajoules (GJ)/t of the nutrient, energy demand comparable with that required in making cement.

Phosphate rock is a finite nonrenewable resource. There is no substitute for phosphorus. Phosphate deposits are known from every continent of the world with the exception of Antarctica. Recently published total world phosphate rock reserves range from 60 billion tons of producible concentrate [4] to 65 billion tons of undifferentiated phosphate rock. Total world phosphate rock resources may range from 290 billion tons to 460 billion tons of phosphate rock at varying P_2O_5 grades [4]. Phosphate deposits are not evenly distributed around the globe. The most abundant known reserves, in descending order of abundance, are located in Morocco, China, the United States, Jordan, and Russia. At the current rates of extraction, these reserves can support the current rate of use for 300–400 years. Of course, in the future based mainly on requirements due to population growth, increased demand for phosphorus-based products will require increased rates of phosphate rock production and the rate of depletion of known reserves may increase. This time horizon can be extended by exploiting by both known, incompletely explored, and yet unknown phosphate deposits at higher costs. Conservation and efficient use of this valuable natural resource and recycling of P has received and will continue to receive more emphasis in the future.

Nitrogen fertilizers via ammonia synthesis account for more than 90% of the world's nitrogen fertilizers. Nitrogen supply for ammonia synthesis is truly inexhaustible since the atmosphere contains 3.8 quadrillion tons of the element. Various feedstocks can be used to obtain hydrogen, and during the last several decades the focus has been to improve the energy efficiency of ammonia synthesis. Natural gas is the preferred feedstock and the best natural gas-based plants now use less than 33 GJ/t N. The global mean, which is affected by more energy-intensive reforming of heavier hydrocarbons (naphtha and fuel oil) and coal, is now between 49 and 55 GJ/t N, roughly half of the level prevailing during the early 1950s.

Even if all of the energy needed to fix the fertilizer nitrogen were to come from natural gas, it would still be less than 7% of the recent annual global consumption of the fuel and less than 2% of all energy derived from fossil fuels [5]. Clearly, there is little reason to be concerned about either the current needs or the future supplies of energy for producing nitrogenous fertilizers via the ammonia synthesis route. Moreover, there is no doubt that higher absolute energy needs for nitrogen fertilizers will be partially offset by improved efficiency of ammonia synthesis and by higher efficiencies of fertilizer use. Because today's low-income countries will experience much faster growth of energy needs in sectors other than the fertilizer industry, the share of global fossil fuel consumption for ammonia synthesis by the middle of the twenty-first century may be only marginally higher than it is today.

Global natural gas resources are considered abundant. Generally, proven reserves of natural gas are considered those quantities that geological and engineering data indicate can be recovered with reasonable certainty in the future from known reservoirs under presently existing operating and economic conditions. Total world proved reserves were 187.49 trillion cubic meters at the end of 2009 [6]. The reserve to production ratio (R/P ratio) was increased at the end of 2009 to 62.8. That is, the amount of proven reserves divided by the production that year indicates the amount of years of reserves remaining at that rate. It should be noted that gar reserves have increased over the last 20 years from 122.4 trillion cubic meters in 1989 to the present figure of 187.49 trillion cubic meters. Similar to many other natural resources, natural gas reserves are a dynamic quantity determined by exploration activity, technology, and economics. In absence of natural gas, naphtha or fuel oil, ammonia synthesis could be accomplished (albeit more costly) by tapping the worlds coal resources or by using a variety of biomass feed stocks.

Nitrogen Fertilizers

Nitrogen is a component of amino acids that make up proteins; chlorophyll (the molecule that captures the sun's energy); enzymes; and the genetic material, nucleic acids. Therefore, this nutrient is required in large amounts by all plants and forms one of three primary nutrients. Although nitrogen is available in abundance (79% by volume) in the earth's atmosphere, only a very limited number of plant varieties, chiefly legumes, can utilize this nitrogen directly from the air. For most plants, including such important cereal crops as rice, wheat, and corn, the nitrogen must be in a chemical form dissolved in the soil solution. Atmospheric phenomena, including lightning discharge, convert nitrogen to nitrogen oxides, which then is absorbed in rain water and enters the soil. However, this contribution is quite small and the resultant nitrogen content of soils is normally very low (less than 0.3%). Continuous cropping without replenishment quickly depletes the soil of this important nutrient.

Until about the early 1950s, the more widely accepted method of supplying nitrogen was through the application of manures/organic waste and the use of crop rotation. Crop rotation entailed growth of a nitrogen-fixing legume crop such as peas, clover, or alfalfa, which then was plowed into the soil to provide nitrogen for a subsequent non-legume crop, for example wheat and corn. This method gives the added bonus of improving soil organic matter status. Such a crop rotation is relatively inefficient with regards to land usage and labor requirement. Modern, high-yield agriculture cannot, under most conditions, be sustained by such a system. Nevertheless, the use of crop rotation is being promoted in many countries of sub-Saharan Africa, particularly landlocked countries where natural soil fertility is very low and the price of fertilizers relatively high.

Natural Organics

Animal and human wastes have long been used as fertilizers, especially in Europe and Asia, particularly China. Even some American Indians are said to have planted a dead fish in each corn hill to increase yields. Such materials contain small percentages of nitrogen and other plant nutrients that are assimilable by plants. Today, the use of raw sewage on crops persists in Asia and Europe, but volumewise is not of great significance. In the United States and most European countries, the use of raw sewage is not considered acceptable, but some municipal sewage plants employ special sterilization, heavy metals removal, and drying processes that yield acceptable fertilizers of low nitrogen content. The volume of such output is small, however. Other organic materials traditionally employed as fertilizers of usable nitrogen content include guano (deposits of accumulated bird droppings), fish meal, and packinghouse wastes including bone meal and dried blood. However, the cumulative importance of all such natural nitrogen sources in modern agriculture is minor. Less than 1% of the total fertilizer nitrogen now used comes from such sources. These products for the most part are not chemically altered prior to use. Processing is mainly physical in nature to improve handling and distribution properties.

Nitrogen Fertilizers from Synthetic Ammonia

The development of a practical ammonia synthesis process in the early years of the twentieth century was a profound scientific achievement of great social significance in view of the subsequent dependence of the world on fertilizer for support of its growing population.

Both the historical background and detailed technical descriptions of modern ammonia synthesis processes are covered in Chap. 22. These processes are all refinements of the original one conceived and developed by Fritz Haber and Carl Bosch in Germany during the period 1904–1913. Basically, a gaseous mixture of nitrogen and hydrogen, in proper proportions to form ammonia, is compressed to very high pressure in the presence of an activated iron catalyst. Ammonia forms and is removed by cooling and condensation. The nitrogen feed is obtained from air by any of several routes that remove the oxygen. Likewise, there are several routes by which hydrogen feed is obtained, most of which involve decomposition of water by reaction with a carbon source such as natural gas, naphtha, or coke. The preparation and the purification of the synthesis gas and the synthesis itself are highly sophisticated modern processes that require great skill and know-how in design, construction, and operation of plants.

There are many processing routes by which synthetic ammonia subsequently finds its way into finished fertilizers. The major routes are outlined graphically in Fig. 24.8 and are discussed in some detail below.

Anhydrous Ammonia. Anhydrous ammonia, without further processing, is an excellent nitrogen fertilizer when properly injected in gaseous form below the surface of many soils. Such direct injection first came into farm use in the early 1940s and rapidly became accepted. Anhydrous ammonia is an important nitrogen fertilizer in the United States, Canada, Denmark, and Mexico, but in some other countries, for various reasons it is a relatively minor nitrogen source [3].

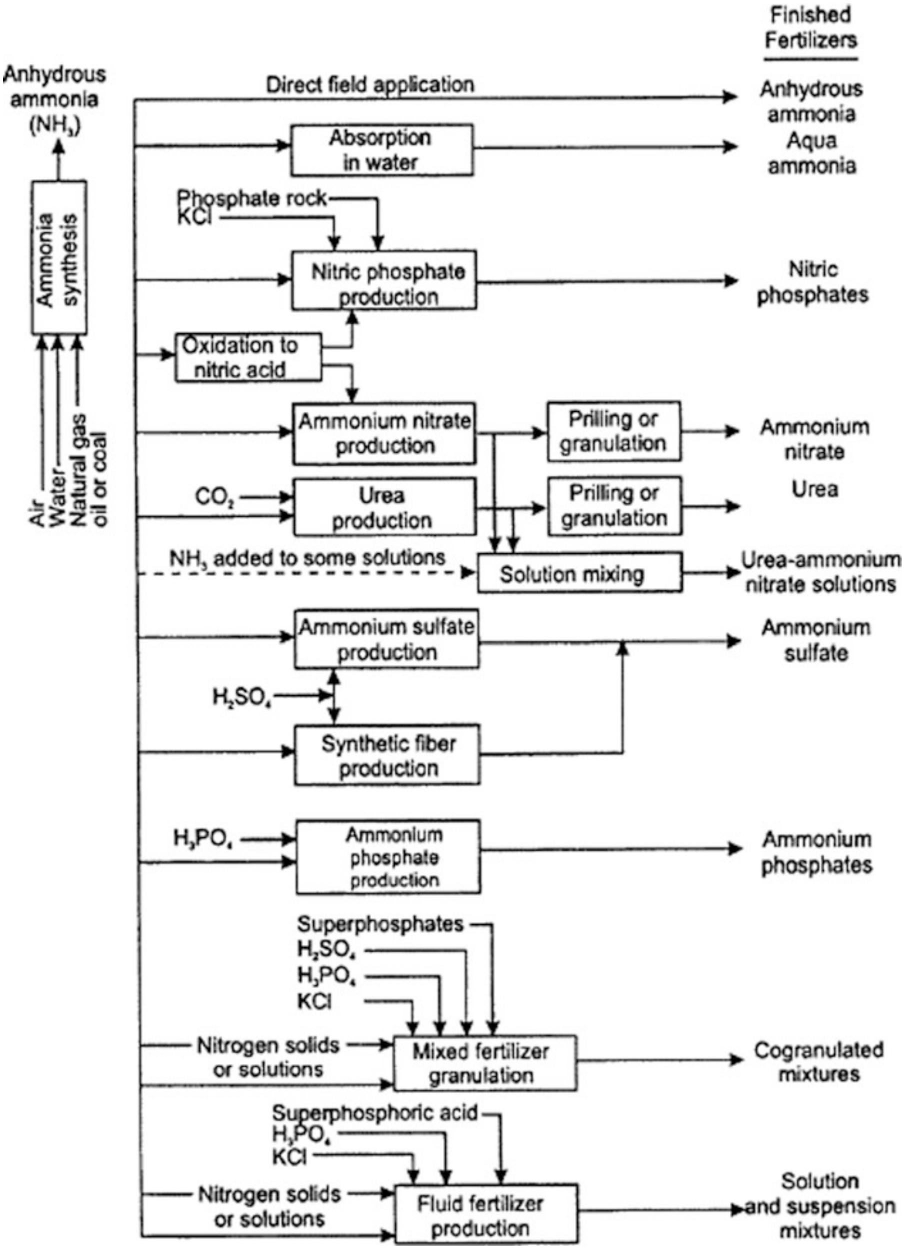
Anhydrous ammonia injection is popular for economic reasons. Ammonia is naturally the cheapest form of fixed nitrogen at the point of production because no further processing is required. Also, the nitrogen content of anhydrous ammonia (82.4% N) is almost twice that of the next most concentrated nitrogen fertilizer (urea: 46% N); this high nitrogen content reduces shipping and handling weight proportionally.

From production to field application, ammonia is handled in liquefied form. Because ammonia is a gas at normal atmospheric pressure and temperatures, its maintenance in liquid form requires either pressurization or refrigeration. Although such handling requires specialized equipment and considerable safety precautions, there are advantages and economies inherent in the handling of a fluid rather than a solid [3].

In order to maintain ammonia in liquid form at atmospheric pressure, it is necessary to refrigerate it to at least −33°C. In storage and handling without refrigeration, rather high pressures must be contained when atmospheric temperatures reach high levels. For example, a pressure of 19 kg/cm^2, which for safety reasons has been set as the upper limit allowable for storage vessels at retail locations, will maintain the liquid form only up to a temperature of 46°C. Safety pop-off valves are set to vent ammonia at higher pressures.

Facilities for handling liquid ammonia from points of production to points of use are highly developed in the United States. A considerable amount of ammonia is moved hundreds of miles through interstate pipelines. Major pipelines run from points of production in Texas, Oklahoma, and Louisiana oil fields all the way to the intensively farmed mid-western area of the United States. Storage at the points of production and at large terminals sometimes is in pressurized vessels (spheres), but more often it is at atmospheric pressure in large, insulated, refrigerated vessels. Refrigeration is effected simply by allowing vaporization of some of the contained ammonia, compressing the resulting vapor, and recondensing it in a water-cooled condenser. Barges with similar self-sustaining refrigeration systems commonly move large amounts of anhydrous ammonia on rivers and on the high seas. Also, specially designed railroad tank cars ranging in capacity from 24 to 73 tons are used widely for ammonia transport. These cars typically are

Fig. 24.8 Major routes of synthetic ammonia into finished fertilizers *(Courtesy TVA)*



nonrefrigerated, instead depending on pressure retention to maintain the liquid form. Large pressurized tank-trucks also are used. At retail depots in farming areas, storage normally is in pressure vessels, often with some compressor capability to recondense vapor and to facilitate transfers. Movement of ammonia to farm fields is via wheeled pressure tanks known as “nurse” tanks. Application equipment, usually tractor-drawn, is equipped with pressure tanks that are filled by transfer from nurse tanks.

Because anhydrous ammonia is gaseous at atmospheric conditions, for fertilizer use it must be injected below the surface of the soil to ensure absorption by the soil solution. Most applicators are equipped with a metering system and a series of soil-penetrating “knives”; the metered ammonia is piped through tubing down the trailing edge of each knife to a single opening at the deepest extremity. A penetration of 15–25 cm is typical. Retention of ammonia in the soil is best when abundant soil moisture is present.

Aqua Ammonia. Ammonia in a simple water solution is popular as a nitrogen fertilizer in some locations, but overall it claims less than 1% of the total nitrogen fertilizer market. This product is made from anhydrous ammonia in units referred to as “converters.” These units simply provide cooling as ammonia and water are combined in metered proportions. Aqua ammonia normally contains 20% nitrogen (24% NH_3). Its chief advantage over anhydrous ammonia is its low vapor pressure; this allows handling in nonpressure equipment, which simplifies design and reduces safety hazards. The chief disadvantage of aqua ammonia as a fertilizer is its relatively low nitrogen content, which limits its shipping range and increases field handling requirements. As with anhydrous ammonia, the preferred application method is knifing below the soil surface to limit volatilization loss; however, a relatively shallow injection depth of only 7–13 cm can be used.

Ammonium Nitrate and Urea. Almost half of the total global fertilizer nitrogen application is as ammonium nitrate or urea, in either solid or solution form, although many countries are restricting the availability of ammonium nitrate because of its potential use by terrorists to make bombs. The history and the technology of ammonium nitrate and urea production from synthetic ammonia are covered in detail in Chap. 29. The production of nitric acid from synthetic ammonia, which is a prerequisite to the production of ammonium nitrate, also is covered completely in Chap. 29. The production plants for all these materials are, like synthetic ammonia plants, highly sophisticated, involving all facets of modern engineering.

The production processes for both ammonium nitrate and urea yield their products first in the form of saturated solutions. Typical ammonium nitrate solution concentration is 75–85% (26–30% N) at 40–75°C, and typical urea solution concentration is about 75% (34% N) at 93°C. To produce the solid forms of these products, the solutions first must be concentrated to extremely low water contents (1–2%) and then fed to prilling or granulation equipment. The resultant solid products that range in size between 2 and 4 mm are highly acceptable to farmers as high-nitrogen fertilizers and to "bulk blenders" as nitrogen sources in the formulation of dry-blend mixed fertilizers, as will be described later. The nitrogen content of solid ammonium nitrate is 34%, and that of urea is 46%.

The concentration and the solidification of ammonium nitrate and urea solutions to form solid fertilizers require considerable expenditure of energy and processing expense; so it is attractive to use the solutions in the formulation of liquid fertilizers, without processing them into solid form. However, neither ammonium nitrate nor urea alone is sufficiently soluble at ambient temperatures to be attractive for direct use as liquid fertilizer. With either compound alone, a solution with a salting-out temperature of 0°C would contain only about 18% nitrogen. Fortunately, however, the joint solubility of the two compounds in water is much more favorable than this; so the liquid nitrogen solutions marketed for fertilizer use almost invariably are joint solutions of ammonium nitrate and urea. Typical commercial solutions contain 28–32% nitrogen, with compositions and characteristics as shown in Table 24.1.

Table 24.1 Physical and chemical characteristics of urea–ammonium nitrate nonpressure nitrogen solutions

Grade, % N	*28*	*30*	*32*
Composition by weight			
Ammonium nitrate, %	40.1	42.2	43.3
Urea, %	30.0	32.7	35.4
Water, %	29.9	25.1	20.3
Specific gravity, 15.6°C	1.283	1.303	1.32
Salt-out temperature, °C	−18	−10	−2

Source: Fertilizer Manual, United Nations Industrial Development Organization and International Fertilizer Development Center, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998

The high-volume use of these solutions is a result of their economy of production and the safety and convenience of their handling and application. The use of nonpressure equipment and surface application (in contrast to knifing) is normal. Nitrogen solutions of other compositions also are prepared commercially, but are used primarily as intermediates in the preparation of solid fertilizers. Such solutions usually consist of various combinations of water, ammonia, ammonium nitrate, and/or urea. Because of the content of free ammonia, most of these solutions require pressurized handling.

Ammonium Sulfate. Although an excellent nitrogen–sulfur fertilizer with good physical properties, ammonium sulfate is now outranked economically by ammonium nitrate and urea. The chief reasons for the unpopularity of ammonium sulfate are its relatively low nitrogen content (21%) and its relatively low solubility in liquid fertilizers. At most U.S. locations, it is uneconomical to produce ammonium sulfate from synthetic ammonia and virgin sulfuric acid. As a result, most of the ammonium sulfate available for fertilizer use is a by-product of other processes. A larger source now is by-product ammonium sulfate solution from the production of caprolactum and acrylonitrile, which are synthetic fiber intermediates. This by-product solution normally is fed to steam-heated evaporator–crystallizers to yield the solid ammonium sulfate. Additional production of by-product ammonium sulfate is obtained from "spent" sulfuric acid, by reaction with synthetic ammonia. In this case, the production of ammonium sulfate is a method of disposing of acid that already has been partially spent in other processes. Some ammonium sulfate is produced by scrubbing ammonia-containing coke-oven effluent gas with sulfuric acid. This production is discussed later.

By-product ammonium sulfate produced by any of the above-mentioned methods is usually in the form of crystals that are 90% or less than 1 mm in size, and thus are much smaller than the size preferred for present-day finished fertilizers. This small size is not undesirable if the material is to be used as feed to a fertilizer granulation plant, but it is undesirable for bulk blending (dry mixing) or direct field application. It is possible to produce suitably large crystals, but only at the expense of greatly reducing crystallizer capacity. Both compaction and granulation processes have been employed with some success on a limited scale to convert the small-crystal material to a larger-size product [7].

Ammonia in Multinutrient Fertilizers. All the ammonia-based fertilizers discussed thus far contain only one of the three major plant nutrients, namely nitrogen. Referring again to Fig. 24.8, one may see that ammonia is the source of fertilizer nitrogen also, either directly or via nitrogen solids or solutions, in multinutrient fertilizers. These are fertilizers that contain two or three of the major plant nutrients—nitrogen, phosphorus, and potassium. Ammonium phosphates, both mono- and di-, are phosphate fertilizer products that also carry nitrogen. (These are discussed later as major suppliers of phosphate.) As suppliers of nitrogen, the ammonium phosphates and other multinutrient solid and fluid fertilizers collectively furnish about 20% of global fertilizer nitrogen. (Production of these fertilizers is also described later.) The source of their nitrogen is, in all cases, ammonia and/or ammonia-based solids or solutions. Nitric phosphates are suppliers of nitrogen through their content of nitric acid and ammonia. They are not presently produced or used to any significant extent in the United States but are produced in Europe and some other locations. (Their production too is discussed briefly later).

Miscellaneous Low-Volume Nitrogen Fertilizers

Mineral Nitrogen. The only nitrogen-containing mineral that is mined commercially for fertilizer use is saltpeter (sodium nitrate, $NaNO_3$). An impure form of sodium nitrate known as "coliche" is found in large, rocklike deposits in a relatively dry desert area of Chile, and these deposits have long been mined to provide natural sodium nitrate for the fertilizer market. The ore is blasted from open mines, sized, and leached with water in vats. Evaporation of the resultant solution yields the commercial crystalline fertilizer product known widely as "Chilean nitrate," "nitrate of soda," or among some farmers simply "sody." Its nitrogen content is 16%. This product first was imported and marketed in the United States in 1830, so it is one of the oldest fertilizers on the American market. Sodium nitrate now, however, is a relatively insignificant component of the world's nitrogen fertilizer market, having been supplanted by synthetic nitrogen compounds.

By-product Ammonia. Coke, an essential ingredient in the smelting of iron ore and in some other metallurgical processes, is produced by the heating of coal in an essentially oxygen-free atmosphere. This heating volatilizes numerous hydrocarbon components of the coal and results in decomposition of some of these compounds. The nitrogen content of coals normally is about 1%, and approximately half of this is volatilized as ammonia gas in the coking process. Recovery of this ammonia has long been effected by passing the coke-oven effluent gas through scrubber–crystallizers containing sulfuric acid. The resultant crystalline by-product, ammonium sulfate, is an excellent fertilizer source of both nitrogen and the secondary plant nutrient sulfur. The nominal nitrogen content of the product is 20%, and the sulfur content is 24%. Prior to the introduction of synthetic ammonia production processes (1913–1925), coke-oven ammonia was a major source of fertilizer nitrogen; however, its production was very small when compared with today's demand for fertilizer nitrogen. Today, ammonia production from coke ovens, although still carried out, is insignificant in volume compared with that from ammonia synthesis processes.

Phosphate Fertilizers

Phosphorus intake is essential to the metabolism of both plants and animals. In the case of plants, the entire uptake is from the soil solution via root absorption. Absorption occurs only when the phosphorus is present in the soil solution in the form of $H_2PO_4^-$, HPO_4^{2-}, or PO_4^{3-} ions [8]. Furthermore, a large proportion of the soil phosphorus is present in chemical forms that are not "available" (assimilable), or are only very slowly available, to crops. Cropping of native soils without phosphate fertilization soon depletes the supply of phosphorus and renders the soil barren. In times before fertilizers became readily available, it was not uncommon to "wear out" a farm by repeated cropping. Crop rotation, which was mentioned earlier as a useful method of converting atmospheric nitrogen to available soil nitrogen, is ineffective as a means of phosphorus fertilization. In fact, the nitrogen-converting legumes generally are voracious consumers of phosphate.

Modern, high-yield agriculture is dependent on regular fertilization with phosphorus compounds that are either immediately soluble in the soil solution or become soluble at a rate sufficient to supply the crop. A factor to contend with in phosphate fertilization is soil "fixation" of phosphorus; that is, the characteristic of many soils to convert at least a portion of applied phosphate to chemical forms that are unavailable to plants. The fixation tendency varies with soil type, and generally is highest in soils of high clay content. The overall result of fixation is that considerably more phosphate must be applied as fertilizer than is removed with crops.

Natural Organic Phosphate Fertilizers

In animal metabolism, phosphorus concentrates in the bones and, in the case of Crustacea, in the shells. Manures and human wastes, previously mentioned as effective but low-volume sources of fertilizer nitrogen, are much less efficient sources of phosphorus than of nitrogen. The phosphorus content of such products is normally only 1.3–3.0%, which is only 25–50% as great as their nitrogen content. Bird droppings, including chicken manure and guano deposits, are somewhat higher in phosphorus content (7% P), but are available in only relatively insignificant amounts overall.

Raw animal bones normally contain 8–10% phosphorus (20–25% P_2O_5); thus they are a relatively rich source. The commercial grinding of bones for fertilizer use began in Europe in the early nineteenth century [9]. The practice persists today in the production of bone meal, but only in very low volume, chiefly to furnish slowly available nutrients for greenhouse and house plants. In about 1830, it was found that pretreatment of bone meal with dilute sulfuric acid greatly enhanced the "availability" of the phosphorus to plants, and the practice became common. The initiation of this practice now is regarded to have had considerable historic significance as the apparent beginning of the chemical fertilizer industry. The supply of bones being very limited, it soon was found (about 1840) that a similar acid treatment of mined phosphate ore rendered the phosphorus "available" as a fertilizer. Thus was born the still important "superphosphate" industry, and the supplanting of the very limited organic phosphorus sources by widely available mineral sources was begun.

Today, organic sources of fertilizer phosphorus account for less than 1% of total worldwide consumption.

Fertilizers from Mineral Phosphates

Essentially all fertilizer phosphorus now is derived from mined ores. (The occurrence, mineral characteristics, mining, and beneficiation of major phosphate ores were described in some detail in Chap. 23.) Worldwide, about 85% of the mined phosphate eventually finds its way into fertilizer [3]. As mentioned earlier, the most conservative estimates indicate a sufficiency for hundreds of years at expected consumption levels. Supply problems of the immediate future will relate chiefly to exhaustion of the better ores, with the result that ores of lower grades and higher impurity contents will have to be processed.

Major routes of mineral phosphate (phosphate rock) into finished fertilizers are outlined in Fig. 24.9 and are discussed below. It is obvious from the figure that although phosphate rock is used directly in several major fertilizer production processes (single superphosphate (SSP), nitric phosphates), most important processes require that the rock phosphate first be converted to phosphoric acid (H_3PO_4). Phosphoric acid production, then, is a very significant component of the phosphate fertilizer industry. The processes used for the production of fertilizer-grade phosphoric acid are known collectively as "wet" processes, and involve, initially, dissolution of the rock in sulfuric acid. (The highly specialized technology of wet-process acid production was discussed in Chap. 23).

Single Superphosphate. In 1842, Lawes in England followed up on the pioneering work of Liebig and received a patent on the use of sulfuric acid in reaction with raw phosphate material to produce "superphosphate" fertilizer. The industry grew slowly, and in 1862 about 150,000 tons of what later would be referred to as "ordinary" (OSP) or "normal" (NSP) or "single" (SSP) superphosphate (16–18% P_2O_5) was produced in England. By 1870, there were 70 SSP plants in the United Kingdom and 7 in the Charleston, South Carolina, area of the United States [10]. In 1888, commercial shipment of phosphate rock from Florida was initiated, and a major mining industry followed in that state, which continues today. As late as 1955, SSP supplied over 60% of the world's phosphate fertilization and was the major phosphate fertilizer in the United States. Rock from Florida or South Carolina was shipped throughout the country to local SSP production facilities. These facilities usually included the capability of blending the SSP with nitrogen and potash materials and bagging the resultant mixtures.

In recent years the production of SSP, to a large extent, has given way to the production of the higher-analysis "triple superphosphate (TSP)" and ammonium phosphates. SSP retains greater importance in some countries. For example, SSP remains popular in Australia, at least in part because its sulfur content is particularly needed there.

SSP is the simplest, as well as the oldest, of manufactured phosphate fertilizers. Its manufacture consists of reacting pulverized phosphate rock with sulfuric acid in comparatively simple equipment, followed by sufficient aging (curing) to complete the reaction and improve the physical condition. The acid rock reaction converts the water-insoluble apatite structure of the rock to soluble monocalcium phosphate. When produced from high-grade rock of 30–32% P_2O_5 content, SSP contains about 20% available P_2O_5; the product made in earlier years from lower-grade rocks contained only 14–18% P_2O_5. SSP can be used for direct application, for bulk blending, or in the production of granular NPK fertilizers. If the product is to be used for direct application or bulk blending, it usually is granulated to improve handling properties.

The main equipment for manufacturing SSP consists of a mixer to bring together sulfuric acid and phosphate rock,

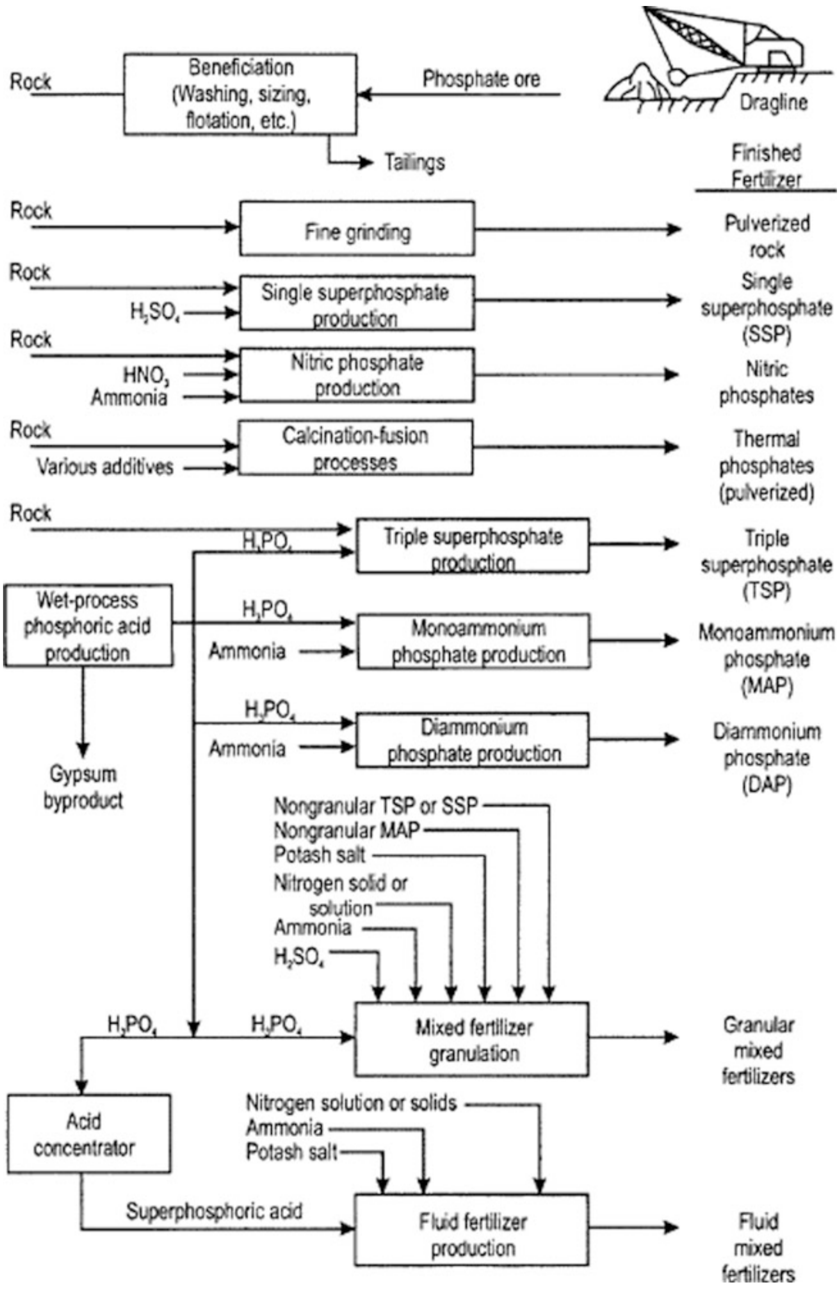


Fig. 24.9 Major routes of mineral phosphate into finished fertilizers *(Courtesy TVA)*

followed by a den, and if granulation is used, some suitable type of equipment for granulation with steam or water. Early plants used batch-type mixers for acid–rock mixing, but these mixers mostly have been replaced by continuous types. Several types of continuous mixers have been employed; of these, the lowest-cost, simplest, and apparently most satisfactory one is the TVA cone mixer, which has no moving parts. Mixing is accomplished by the swirling action of the acid. Short, single-shaft or double-shaft pug mills also are used for continuous mixing. The mixers discharge into a den, the purpose of which is to allow time for the acid–rock reaction to progress and the physical condition to reach a dry, friable state. The dens can be of either the batch or the continuous type. A slat-conveyor continuous den, commonly referred to as a Broadfield den, is used in many continuous systems [11]. A diagram of a typical SSP plant employing a cone mixer and a continuous den (slat-conveyor) is shown in Fig. 24.10. The reaction of phosphate rock with sulfuric acid to produce SSP can be approximated in chemical stoichiometric terms.

However, in general practice the proportioning usually is based on a simpler relationship of about 0.6 kg of sulfuric acid (100% H_2SO_4 basis) per kilogram of phosphate rock (30–32% P_2O_5). The phosphate rock usually is pulverized to

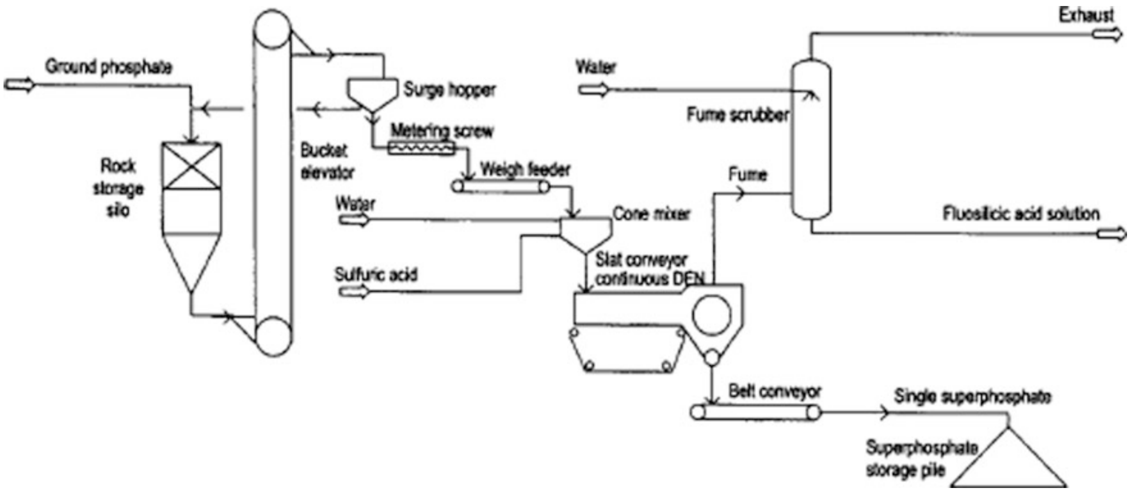


Fig. 24.10 Continuous process for manufacture of single superphosphate *(Source: Fertilizer Manual)*

about 90% smaller than 0.15 mm and 70% smaller than 0.075 mm.

Gases that are released while the superphosphate is solidifying (setting) cause the mass in the den to become porous and friable so that it can be "cut" (disintegrated) and handled readily. SSP made from typical rock will "set" in 40–50 min in a continuous den, whereas the set time in a batch den is 1.5–2 h. The superphosphate usually is held in storage piles (cured) for 4–6 weeks in order to obtain better handling properties and to allow the chemical reactions to continue. The usual grade of SSP made from Florida rock is 20% available P_2O_5. A typical analysis is shown below:

Analysis, % by Weight

P_2O_5								
Total	Available	H_2O Soluble	CaO	Free Acid	SO_4	F	R_2O_3	MgO
20.2	19.8	18.0	28.1	3.7	29.7	1.6	1.6	0.15

Because of the low phosphate analysis of SSP (20% P_2O_5), economics favor shipping the rock (32% P_2O_5) to local plants and producing the superphosphate there. A majority of the producers use the locally produced SSP in formulations for granular NPK fertilizers made in an ammoniation–granulation process. The maximum amount of SSP possible is used in the NPK formulations because usually it is the lowest-cost form of P_2O_5; however, because of the relatively low P_2O_5 content of SSP, not much can be used in high-analysis NPK formulations. SSP will readily absorb about 6 kg of ammonia per 20 kg of P_2O_5 during the ammoniation–granulation process. Production of granular SSP can be carried out by starting with either cured or ex-den SSP. The SSP is fed to a rotary drum or pan granulator along with a water spray and/or steam. The granulator is followed by a dryer and screening (sizing) system. Fines and crushed oversize material are fed back to the granulator.

Triple Superphosphate. TSP is made by acidulation of phosphate rock with phosphoric acid, using equipment and processes similar to those for SSP. As in SSP production, the acid–rock reaction converts the water-insoluble apatite structure of the rock to water-soluble monocalcium phosphate. TSP, with its relatively high P_2O_5 content of 45–46%, did not appear on the scene in any appreciable quantity until wet-process phosphoric acid (see Chap. 23) was produced commercially. TVA carried out demonstration-scale production of TSP in the late 1930s by using electric-furnace phosphoric acid. Extensive agronomic testing and market development, through the use of large tonnages of this TSP by farmers and fertilizer manufacturers in TVA demonstration programs, led to the rapid acceptance of TSP. Producers of phosphate rock in the United States and other countries moved into production of wet-process acid and TSP. Logistics favored production of the higher-analysis TSP (46% P_2O_5 vs. 32% P_2O_5 in phosphate rock) near the source of the phosphate rock, and shipment of the TSP to distribution facilities or blending plants near the markets. Today, most producers of TSP are located adjacent to phosphate mining operations in Florida, Morocco, and Tunisia. Since its introduction, TSP has steadily displaced SSP in the marketplace.

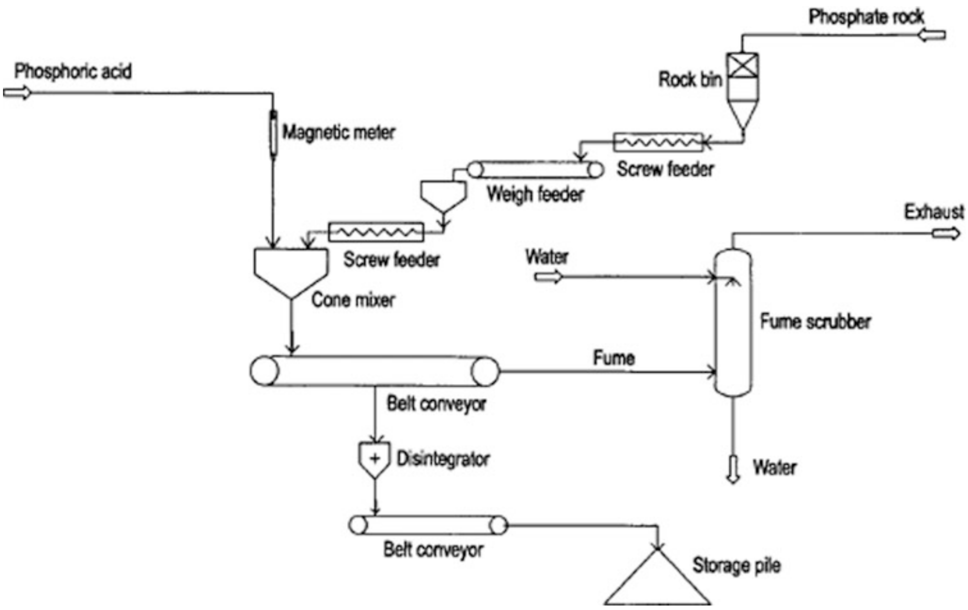


Fig. 24.11 Continuous system for manufacture of nongranular triple superphosphate *(Source: Fertilizer Manual)*

The cone mixer has been used almost universally in the production of nongranular TSP. Because the "set time" for TSP is only 14–20 min, as compared with 40–50 min for SSP, a simple, cupped conveyor belt usually is used to hold the acidulate until the TSP solidifies; this is instead of the slat-type den used for SSP. With a belt about 1.5 m wide and 30 m long, the production rate usually is 40–50 tons/h. A flow diagram for TSP production by the cone mixer-"wet-belt" system is shown in Fig. 24.11.

Proportioning for TSP typically is 2.4–2.5 kg of P_2O_5 from acid for each kg of P_2O_5 from rock. Nongranular TSP usually is cured 4–6 weeks prior to shipment or used at the site. Typical chemical analysis of TSP made with Florida rock is shown below:

Analysis, % by Weight

P_2O_5									
Total	Available	H_2O Soluble	Free Acid	CaO	R_2O_3	MgO	F	H_2O
46.9	46.3	42.0	3.4	19.3	3.1	0.5	2.7	4.5

The use of TSP in granular NPK fertilizer formulations, together with or in place of SSP, has allowed the production of higher-analysis grades of granular NPK fertilizers, such as 13–13–13 instead of 10–10–10. The TSP can readily be ammoniated using 3.5 kg of ammonia per 20 kg of P_2O_5 during the granulation process.

TSP is produced in large quantities in granular form for use in direct application and for use in bulk blends. In some processes cured or ex-den TSP, like SSP, is granulated in a rotary drum or pan granulator using steam and water to promote granulation [3]. A slurry-type granulation process, as outlined in Fig. 24.12, is now the process of choice of major producers of TSP. Pulverized phosphate rock is treated with wet-process phosphoric acid in a two-stage reaction system, and the slurry is sprayed into a pug mill or rotary drum where it is layered on recycled undersize and crushed oversize at a ratio of 1 kg/5–12 kg of recycle. Product granules (45% P_2O_5) are quite spherical and dense. The lower grade of the granular product, as compared with that of cured nongranular TSP, results from the need to decrease the acidulation ratio from the usual 2.45 kg of acid P_2O_5 per kilogram of rock P_2O_5 to about 2.25 in order to control the free-acid content and to minimize stickiness. For production rates higher than about 25 tons/h, a rotary drum usually is used instead of a pug mill.

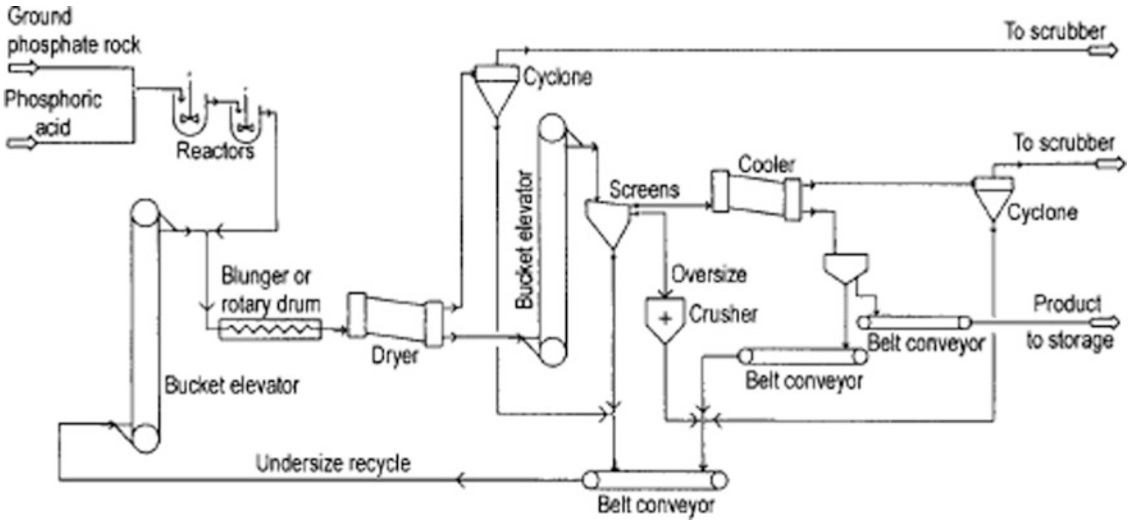


Fig. 24.12 Slurry process for the manufacture of granular triple superphosphate *(Source: Fertilizer Manual)*

Nitric Phosphate. Fertilizers that are referred to as nitric phosphate or nitrophosphate are produced by acidulation of phosphate rock with nitric acid or with mixtures of nitric and sulfuric or phosphoric acids. The primary advantage of nitric phosphate processes is that no sulfur or less sulfur is required as compared with superphosphates or ammonium phosphates; this is particularly important during a shortage of sulfur, or in locations where sulfur must be shipped long distances. A variety of processes and equipment have been used in Europe since the late 1930s [3, 12]. Also there are a number of plants in Central and South America and in Asia. The production of nitric phosphates is complex. Simple substitution of nitric acid in a superphosphate-type acid–rock reaction is not feasible because (1) decomposition of the nitric acid would occur and cause noxious fumes and loss of nitrogen and (2) the product would be extremely hygroscopic and unstable.

One process modification that overcomes these difficulties uses only nitric acid for acidulation; the extraction slurry is cooled to crystallize calcium nitrate, which then is removed by centrifugation. This process is referred to as the Odda process. The calcium nitrate by-product is either sold as a fertilizer or converted to coproduct ammonium nitrate. The extraction slurry, with its lowered calcium content, then is ammoniated and granulated. Various process modifications utilize a rotary drum, pug mill, or spray drum (Spherodizer) for granulation of the slurry. NPK nitric phosphate grades also have been produced by a prilling process in a few European plants. In earlier years, a main disadvantage of nitric phosphate processes was the low water solubility of phosphate in the products, as the main phosphatic constituent was dicalcium phosphate. The use of supplemental phosphoric acid, or "deep cooling" by refrigeration to remove a higher proportion of calcium nitrate in the Odda-type processes, now allows water solubility of 60% or higher. The remaining 40% of the phosphate, although water-insoluble, is almost completely citrate-soluble, and this is "available" as plant nutrient. Popular grades of nitric phosphates include 14–14–14, 22–11–11, 20–20–0, and 16–13–0.

Diammonium Phosphate (DAP). Ammonium phosphates did not come on the fertilizer scene in significant quantities until the early 1960s; however, they have rapidly become the leading form of phosphate fertilizer in the world. Almost all new phosphate fertilizer complexes built in recent years, and those planned, are for the production of ammonium phosphate as the major product. In 1960, world production of ammonium phosphates was equivalent to only 0.5 million mt of P_2O_5. By 1990 it was up to 15.4 million mt and by 1999 it has reached 16.7 million mt. Ammonium phosphates now provide two-thirds of total world phosphate fertilization.

Prior to the late 1940s, DAP $(NH_4)_2HPO_4$ was considered to be too unstable for use as a fertilizer. However, at about that time, exploratory work by TVA indicated otherwise. The first demonstration-scale DAP made by TVA for fertilizer use was a monocrystal product made by ammoniation of pure electric-furnace acid in a vacuum crystallizer [13]. Demonstration programs showed this material to be stable and to be an excellent high-analysis fertilizer. Small amounts of similar product were produced by others using electric-furnace acid and by-product ammonia. However, general acceptance of DAP by the fertilizer industry was not possible until procedures could be developed to produce it from wet-process phosphoric acid instead of the relatively expensive electric-furnace acid. This occurred in 1960–1961 with the development and patenting of the "TVA process" for production of granular 18–46–0 grade DAP from wet-

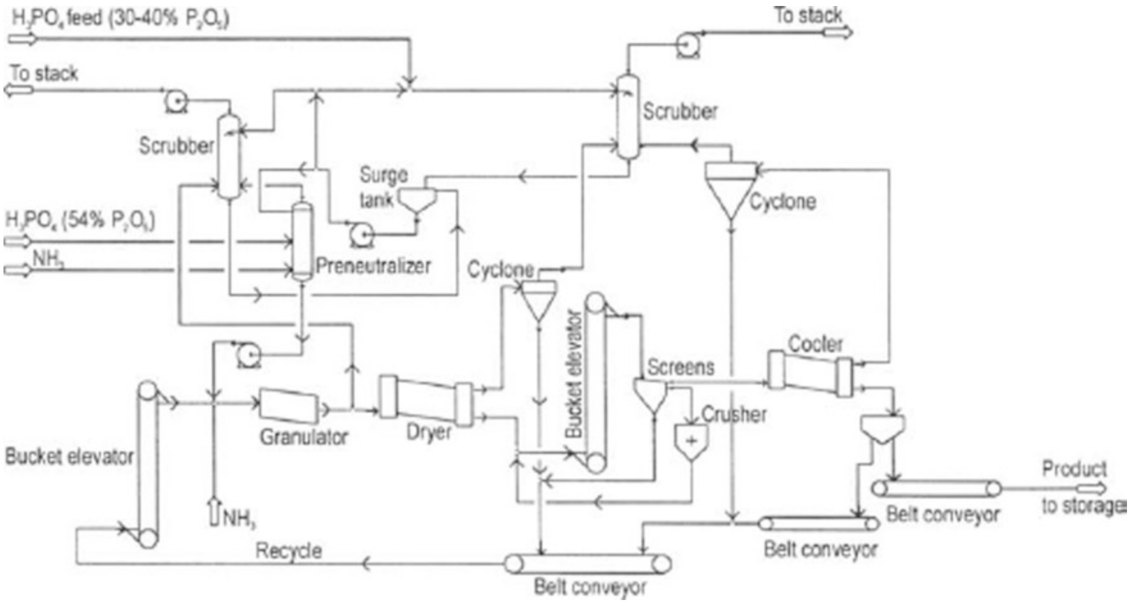


Fig. 24.13 TVA process for production of granular diammonium phosphate *(Source; Fertilizer Manual)*

process acid [14, 15]. This relatively simple process was rapidly adopted by the industry, and remains the basic process for production of this major phosphate fertilizer. Many granular DAP plants have production capacities of about 50 tons/h, with some going as high as 70–100 tons/h.

The flow diagram of a typical granular DAP production unit of the TVA type that has become standard in the industries is shown in Fig. 24.13. Wet-process phosphoric acid of about 40% P_2O_5 content (often a mixture of 54% P_2O_5 acid and acid from the scrubbing circuit of 28–30% P_2O_5 content) is fed to a preneutralizer vessel. Anhydrous ammonia is sparged into this preneutralizer through open-end pipes that project through the walls. A key feature of the process is that this neutralization of the acid is carried to a closely controlled NH_3:H_3PO_4 mole ratio of about 1.4. Ammoniation to this degree provides maximum solubility of ammonium phosphate, as is shown by the solubility curve of Fig. 24.14. The heat of the ammoniation reaction evaporates considerable water, with the result that the water content of the slurry is reduced to only 16–20%. The slurry still is fluid and pumpable because of the high solubility at a 1.4 mole ratio.

The preneutralizer slurry is pumped at a controlled rate to a rotary-drum, ammoniator–granulator; there it is distributed onto a rolling bed of recycled solids. The most commonly used metering system for the preneutralized slurry is a variable-speed centrifugal pump with automatic control from a magnetic flowmeter. Some plants have had success with a magnetic flowmeter and automatic control valve of a special ball type.

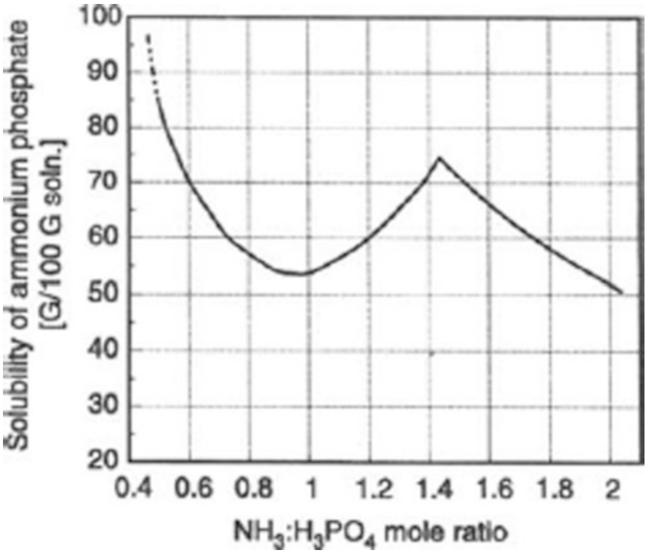


Fig. 24.14 Effect of NH_3: H_3PO_4 mole ratio on solubility of ammonium phosphate at 75 °C *(Source: Fertilizer Manual)*

Ammonia is sparged beneath the bed in the rotary drum to ammoniate the slurry further to near the theoretical DAP ratio of 2.0; the usual finishing NH_3:H_3PO_4 mole ratio is 1.85–1.94. Metallic phosphate impurities such as iron, aluminum, and calcium phosphates account for most of the deviation from theoretical. This ammoniation in the granulator drastically decreases ammonium phosphate solubility (Fig. 24.14), and thus promotes solidification and granulation. Ammonia absorption is not 100% complete in either the preneutralizer or the granulator. However, the unabsorbed ammonia is recovered in a scrubbing circuit containing phosphoric acid of about 30% P_2O_5 content. Solid recycle to the granulator drum consists of undersize from screening of dried product plus some crushed, dry product and oversize. These combined materials are recycled at a rate of 5–7 kg per kilogram of product, and variation of this recycle rate is the primary method of controlling granulation efficiency. Product discharged from the granulator is dried with moderate heat to a 180–190°F product temperature. Higher drying temperatures must be avoided because of a relatively low melting point for DAP. Most plants screen hot and cool only the product fraction because the material is not too sticky for recycle without cooling. Rotary coolers or fluidized-bed-type coolers are used. The product, with a moisture content of 1.5–2%, does not require a conditioner. It has excellent storage and handling properties in bags or in bulk. It has been shown that the iron and aluminum phosphate impurities in the product are present in a gel-like structure that favorably hardens the granules and increases their resistance to moisture [16]. The critical relative humidity of DAP made from wet-process acid is about 70%; thus it is relatively nonhygroscopic.

Construction materials other than mild steel are required only for the acid lines, the preneutralizer, the slurry handling system, and the scrubbing circuits. Type 316 L stainless steel or rubber- and brick-lined mild steel is used for the preneutralizer. Type 316 L stainless is used also for the slurry pumps and piping. Fiberglass-reinforced polyester plastic and high-density polyvinyl chloride are sometimes used for wet-process acid pipes and for scrubbers. Scrubbers also may be constructed of rubber-lined mild steel.

Monoammonium Phosphate (MAP). “DAP,” and the grade designation “18–46–0” have become extremely familiar terms in the world fertilizer industry. However, substantial interest has developed also in fertilizer use of MAP, $NH_4H_2PO_4$. Agronomically, MAP is favored where soils are mainly alkaline, for example in Canada, Egypt, and Pakistan. From an economic standpoint, the relatively high P_2O_5 contents of typical MAP grades (11–52–0 to 10–54–0) make them attractive where the primary interest is in producing and shipping phosphate; the MAPs provide a higher P_2O_5 payload than does DAP.

TVA developed two comparatively minor modifications of the granular DAP process to allow production of granular MAP [17]. In one method, the acid is ammoniated to an NH_3:H_3PO_4, mole ratio of only about 0.6 in the preneutralizer and then to about 1.0 in the granulator drum. The 0.6 ratio is, like the 1.4 used in DAP production, a high-solubility point in the ammonium phosphate system (Fig. 24.14); the subsequent ammoniation to 1.0 in the granulator decreases solubility and thus promotes granulation. In the other procedure, which has been preferred by the industry, acid in the preneutralizer is ammoniated to the high-solubility NH_3:H_3PO_4 ratio of about 1.4, as in DAP production; then additional wet-process acid is distributed onto the bed in the granulator to adjust back to the low-solubility MAP mole ratio of about 1.0. With either modification, the remainder of the process is the same as for DAP, except that, because of the higher temperature stability of MAP, a higher drying temperature can be used to increase the production rate for MAP.

Starting in about 1968, simple processes were developed for the production of nongranular (sometimes called powdered) MAP. Chief developers were Fisons and Scottish Agricultural Industries (SAI) in the United Kingdom, Swift in the United States, and Nissan in Japan. These processes involve variations in the simple reaction of wet-process acid with ammonia followed by spray-type drying with air. On the basis of this work, a number of comparatively low-cost units for production of nongranular MAP have been built commercially, including plants in the United Kingdom, the Netherlands, Japan, Australia, Spain, the United States, Brazil, and Iran. This intermediate usually is shipped to other plants, where it then is used in the production of NPK solid or fluid fertilizers. Thus far, however, the nongranular MAP has not attained the popularity that was predicted for it in the mid-1970s.

Miscellaneous Low-Volume Phosphate Fertilizers

Phosphate Rock for Direct Application. Application of finely pulverized raw phosphate rock directly to the soil has been utilized almost from the beginning of fertilization practice. Although such direct use eliminates the need for most processing, there are major drawbacks. The ores are of low phosphorus content in comparison with most manufactured fertilizers, which increases shipping cost. Also, the very fine grinds that are required to encourage solubility are difficult to handle and to apply. The major drawback, however, is agronomic. The agronomic effectiveness of raw rock depends upon the origin and the chemical and mineralogical nature of the particular rock, the pH and other characteristics of the soil, and the crop produced. Researchers at TVA and IFDC have characterized phosphate rocks from a large number of sources as to their reactivity [3, 18]. In general, phosphate rock of sedimentary origin is more reactive than igneous source. Phosphate rock from North Carolina (U.S.), Sechura (Peru), and Gafsa (Tunisia) are at the top, and are about equal in reactivity and suitability for direct application. Other phosphate rocks that are reasonably reactive and are marketed for this purpose include some types from Morocco, Israel, Algeria, and a few other locations. In 2002, less than 1.0 million mt of phosphate rock was used for direct application.

Defluorinated Phosphate Rock. There is substantial production of defluorinated phosphate rock for fertilizer use in Japan (about 100,000 mt/year). Ground, high-grade rock is mixed with small proportions of sodium carbonate or sulfate and wet-process acid. The mixture is calcined at a temperature of 1,350°C in an oil-fired rotary kiln 45.0 m in length and 2.7 m in diameter. The product contains 38–42% P_2O_5 of which more than 90% is soluble in neutral ammonium nitrate solution and is an effective fertilizer on acid soils. During the production of defluorinated phosphate rock, substantially all fluorine is driven off. Sodium bifluoride ($NaHF_2$) is recovered as a by-product. A similar product is made in the United States, but it is mainly used for animal feed supplement.

Calcined Aluminum Phosphate. Ores consisting mainly of hydrated aluminum phosphate minerals are useful for direct application after thermal treatment (calcination) at a controlled temperature. Calcination removes the combined water and destroys the crystallinity of the phosphate minerals, thereby improving the reactivity of the phosphate [3].

At present, only two aluminum phosphate ores are known to be used for direct application after thermal treatment. One such ore is found near Thies, in Senegal. The “as-mined” ore contains about 29% P_2O_5 on a dry basis; after calcination it contains over 34% P_2O_5. The calcined product is ground to about 95% smaller than 0.15 mm and marketed for direct application under the trade name “Phosphal.” About 75% of the P_2O_5 is soluble in alkaline ammonium citrate, and the product is effective as superphosphate in calcareous, neutral, and acid soils. The product is used in France, Senegal, and some nearby countries.

A similar ore occurs in the upper stratum (“C-grade”) of the phosphate deposit on Christmas Island. The C-grade ore contains about 25% P_2O_5 dry basis. Calcination is carried out in a fluidized bed, operated in the range of 400–600°C. The maximum solubility in alkaline ammonium citrate solution is about 70%.

Partially Acidulated Phosphate Rock (PAPR). Not all phosphate rock is suitable for direct application and may require the addition of a more soluble form of phosphate. An alternative is partial acidulation to render its P_2O_5 more available. The PAPR process depends on treating ground phosphate rock with only a portion of the stoichiometric value of acid, for example 50% PAPR [19]. The amount of water-soluble phosphate in PAPR varies according to the degree of acidulation. The crop response to the product is often quite similar to fully acidulated products such as SSP or TSP.

Sulfuric and phosphoric acids are the most common acids used for partial acidulation. Possibilities exist for using three processes: a conventional run-of-pile process followed by granulation, IFDC-developed singlestep acidulation and granulation, and a slurry granulation TSP-type process.

Enriched Superphosphate. This product, which is also referred to as “double” superphosphate, is essentially a mixture of SSP and TSP, usually made by acidulation of phosphate rock with a mixture of sulfuric and phosphoric acids. The desired concentration of mixed acid can be obtained by mixing concentrated sulfuric acid (93 or 98% H_2SO_4) with dilute phosphoric acid (30%) thus avoiding the need for concentrating the latter. Production processes and equipment are about the same as for SSP [20].

Theoretically, any grade between SSP and TSP can be produced, but the usual range is 25–35% P_2O_5. The product is useful in sulfur-deficient areas where SSP would supply more sulfur than necessary. Only a small amount of enriched superphosphate is produced and used in Australia.

Basic Slag. This is also known as Thomas slag and is a by-product of the steel industry. Iron made from high-phosphorus ore is converted to steel in a Thomas converter by oxidation in contact with a basic (high CaO) slag. Sometimes phosphate rock is deliberately added to the blast furnace charge to increase the phosphorus content of the iron thereby increasing the P_2O_5 content of the slag. A typical range of composition of high-grade basic slag is:

The calcium magnesium phosphate (CMP) is found to be more effective than superphosphate when used on acid soils. The product has a liming value equivalent to 0.5–0.7 ton of calcium carbonate per ton of material. The magnesium oxide and soil-soluble silica is available to growing plants.

Rhenania Phosphate. Rhenania phosphate is another thermally produced phosphate fertilizer. It is made by calcining a mixture of phosphate rock, sodium carbonate, and silica in

While the present market for basic slag is quite small, it is primarily produced and used in Western Europe where high-phosphorus iron ore occurs.

Fused Calcium Magnesium Phosphate. In a process developed by TVA, a mixture of phosphate rock and olivine or serpentine (magnesium silicate) is fused in an electric furnace [21]. The molten product is quenched with water and used in a finely divided state as a fertilizer. The product, a CMP glass, contains about 20% P_2O_5 and 15% MgO. Over 90% of the product is soluble in citric acid.

The theoretical compositions of some magnesium-containing minerals that can be used to produce CMP are:

Olivine	$(Mg,Fe)_2SiO_4$
Serpentine	$Mg_3H_4Si_2O_9$
Garnierite	$(Mg,Ni)H_2SiO_4$
Magnesite	$MgCO_3$

Although small compared with soluble phosphates, CMP is used in Japan, China, Korea, Taiwan, and South Africa.

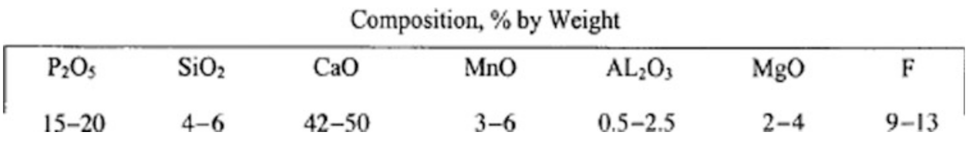
The reason for the decline in production and use is the relatively high (850 kW h/ton of product) energy requirements.

A rotary kiln at 1,250°C [3]. Enough sodium carbonate is used to form the compound $CaNaPO_4$ and enough silica to form Ca_2SiO_4 with excess calcium. Typical charge proportions are one part sodium carbonate to three parts of phosphate rock and enough silica to raise the SiO_2 content of the product to about 10%. The overall reaction in producing Rhenania phosphate is assumed to be:

$$\begin{aligned} &Ca_{10}F_2(PO_4)_6 + 4Na_2CO_3 + 2SiO_2 \\ &\quad \rightarrow 6CaNaPO_4 + 2Ca_2SiO_4 + 2NaF + 4CO_2 \end{aligned}$$

Any grade of phosphate rock can be used, but since the grade of the product is determined by the grade of the rock, a high grade is preferred.

A typical product contains 28–30% P_2O_5, nearly all soluble in neutral or alkaline ammonium citrate solution. The product is applied in a powdered form or granulated with potash salts. Some grades are produced containing magnesium or boron, which are added during granulation as kieserite or borax, respectively.



Composition, % by Weight

P_2O_5	SiO_2	CaO	MnO	AL_2O_3	MgO	F
15–20	4–6	42–50	3–6	0.5–2.5	2–4	9–13

Potassium Salts

The element potassium has been recognized to be beneficial to plant growth since J.R. Glauber.

In the Netherlands first proposed, in the middle of the seventeenth century, that saltpeter (KNO_3) was the "principle" of vegetation. This salt, derived from the leaching of coral soils, was thus the first chemical compound intentionally applied to crops to satisfy their nutrient needs. Justus von Leibig had established the essentiality of the element for plant growth in his pioneering work published in 1840. The potassium or potash chemical industry dates from the year 1861, following the discovery by German chemists of a process for recovery of KCl (muriate of potash) from rubbish salts, the residues remaining from the extraction of sodium chloride (NaCl).

The term "potash" generally refers to potassium chloride, but it is also used to designate other potassium compounds that may be qualified by the words muriate of potash, sulfate of potash, etc.

The functions of potassium in the plant are manifold. This element serves to activate or catalyze a host of enzyme actions, to facilitate the transport of nutrients and assimilates in the xylem and phloem, to maintain the structural integrity of the plant cell, to regulate turgor pressure, to mediate the fixation of nitrogen in leguminous plant species, and to protect plants to some degree from certain plant diseases.

Potassium is a relatively abundant element in the earth's crust, ranking seventh in concentration. It is widely distributed geographically and is commonly found in association with sodium compounds. The feldspars, muscovite (white mica), granite, and gneiss are rich sources; but because they are siliceous and refractory, they are difficult and costly to convert to forms suitable for use as fertilizers. The common potassium minerals are listed in Table 24.2.

The name potash derives from an early production method in which potassium carbonate, leached from wood ashes, was crystallized by evaporating the leachate in large iron pots. The salt potassium chloride (muriate of potash or

Table 24.2 Common K minerals

Mineral	Formula	K content (g/kg)	K_2O content (g/kg)
Sylvite	KCl	524.4	631.7
Carnalite	$KCl{\cdot}MgCl_2{\cdot}6H_2O$	140.7	169.5
Kainite	$KCl{\cdot}MgSO_4{\cdot}3H_2O$	157.1	189.2
Langbeinite	$K_2SO_4{\cdot}2MgSO_4$	188.5	227.0
Leonite	$K_2SO_4{\cdot}MgSO_4{\cdot}4H_2O$	213.3	256.9
Schoenite	$K_2SO_4{\cdot}MgSO_4{\cdot}6H_2O$	194.2	233.9
Polyhalite	$K_2SO_4{\cdot}MgSO_4{\cdot}2CaSO_4{\cdot}2H_2O$	129.7	156.2

Source: Stewart, J. A., in *Potassium in Agriculture*, R. D. Munson (Ed.), p. 83, ASA-CSSA-SSSA, 1985

KCl) is now the major source of the element (95%); other important salts are potassium sulfate (sulfate of potash, K_2SO_4), potassium magnesium sulfates of varying K/Mg ratios, and potassium nitrate (KNO_3).

The potash industry is based on very large deposits of water-soluble potassium minerals resulting from the evaporation of shallow seas or natural brine lakes over a geological time span. These evaporates normally are located at depth in the earth's mantle. Ironically, the best deposits are found in areas quite remote from the more productive agricultural regions, which thus are the areas most in need of this element. Typically, the deposits are to be found in horizontal tabular bodies or beds and may occur at depths up to 2,100 m or more. The beds may be only a few centimeters to a few meters thick, but commercial production is limited to strata that are at least 1 m thick. As with hard-rock mining, the potash-bearing ores are extracted or harvested with continuous mining machines and brought to the surface through vertical shafts. Where the ores occur below a depth of about 1,100 m, or where the beds exhibit geological anomalies (e.g., folding), the potash is dissolved in a brine solution and pumped to the surface for recovery using solution-mining techniques.

Potassium Minerals

The ore zone or stratum typically contains potassium or potassium–magnesium minerals along with halite (sodium chloride). Muriate of potash is refined from sylvinite ore, a mechanical mixture of potassium chloride (KCl) and sodium chloride (NaCl). Because the latter salt is injurious to most crop plants, the KCl (sylvite) must be separated from the NaCl (halite).

For coarse-grained ores, physical methods employing froth flotation and/or heavy-media methods are employed to achieve separation. The heavy medium may consist of, for example, pulverized magnetite in a brine solution. The specific gravity of the medium is adjusted so that it falls between that of KCl (1.99 g/cm^3) and that of NaCl (2.17 g/cm^3).

When the crushed ore is placed in this medium, the potassium values are floated off, and the contaminating sodium values sink and are drawn off and rejected from the bottom of the flotation vessel. Entrapped magnetite from both the product and waste streams is easily recovered magnetically and returned to the flotation vessel.

In froth flotation the crushed ore, which has been scrubbed to remove contaminating clay particles, first is treated with an aliphatic amine acetate and a froth-promoting alcohol. The water-repellant amine acetate is attracted preferentially to the sylvite particles. In a subsequent step, the ore or pulp is maintained in suspension, and the coated sylvite particles are attracted to and entrained on the rising air bubbles generated by the flotation agitator. They rise to the surface and are collected by hydraulic and mechanical means. The potassium-depleted pulp is then discarded. The sylvite or muriate is collected, centrifuged to remove the hydraulic medium, dried, screened, into various particle size ranges, and sent to storage.

Potassium–Magnesium Minerals

Other important sources of potassium for fertilizer use are the double salts kainite ($KCl{\cdot}MgSO_4{\cdot}3H_2O$) and langbeinite ($K_2SO{\cdot}2MgSO_4$). The former is recovered from potash mines in Germany and Sicily; the latter is recovered and processed from mines in New Mexico (U.S.). These minerals supply important quantities of both magnesium and sulfur as well as potassium, and all are essential nutrients for plant growth. Kainite has a theoretical composition of 15.99% K (19.26% K_2O), 9.94% Mg, and 13.11% S, and langbeinite contains 18.85% K (22.7% K_2O), 11.71% Mg, and 23.18% S.

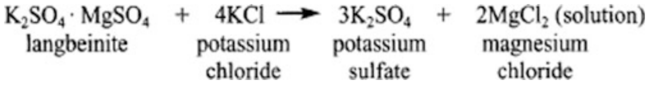
In the processing of langbeinite, the contaminating chlorides (NaCl, KCl) are removed by extracting the crushed ore with water. Centrifugation, drying, and screening follow to prepare the product in various size grades for the market. Where the mixed ore contains recoverable quantities as KCl as well as langbeinite, froth flotation and heavy-media methods may be employed to recover both potassium-containing salts.

Potassium Sulfate

Sulfate of potash (K_2SO_4), unlike the earlier-discussed potash salts, does not occur as natural deposits. It can be recovered by fractional crystallization from such natural brines as those of the Great Salt Lake in Utah and Searles Lake in California. Here separation and recovery are achieved by solar evaporation in shallow ponds. These processes can be utilized only where a suitable brine source is available, and where solar evaporation rates are high.

A much more important source of K_2SO_4 is the product resulting from either a controlled decomposition reaction of a complex sulfate salt, reaction of the salt with KCl, or both. Thus, kainite ($KCl{\cdot}MgSO_4{\cdot}3H_2O$) may be reacted with water in a two-step reaction, first to form schoenite ($K_2SO_4{\cdot}MgSO_4{\cdot}6H_2O$), with the schoenite then reacted with KCl to form K_2SO_4.

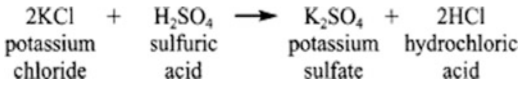
Langbeinite likewise is reacted with KCl to form sulfate of potash, according to the following reaction:



$$\underset{\text{langbeinite}}{K_2SO_4 \cdot MgSO_4} + \underset{\text{potassium chloride}}{4KCl} \longrightarrow \underset{\text{potassium sulfate}}{3K_2SO_4} + \underset{\text{magnesium chloride}}{2MgCl_2\ (\text{solution})}$$

Approximately 50% of world K_2SO_4 production is derived from the reaction of KCl with either kainite or langbeinite.

For those producers who do not have access to such suitable sulfate salts to serve as the sulfate donor, the so-called Mannheim process, in which sulfuric acid is reacted with KCl, may be utilized. The following reaction unfortunately generates the corrosive substance hydrochloric acid as a coproduct:



$$\underset{\text{potassium chloride}}{2KCl} + \underset{\text{sulfuric acid}}{H_2SO_4} \longrightarrow \underset{\text{potassium sulfate}}{K_2SO_4} + \underset{\text{hydrochloric acid}}{2HCl}$$

Potassium sulfate, though relatively costly to produce, is a preferred source of potassium for some vegetable and fruit crops and tobacco—crops that are sensitive to the chloride forms.

Potassium Nitrate

The third most widely used potassium salt in agriculture is potassium nitrate (KNO_3). An ore containing sodium nitrate, potassium nitrate, some chlorides, and sulfates is mined in Chile. The ore is leached and part of the sodium nitrate is crystallized and removed.

The salt is produced in small amounts commercially in the United States, Israel, and Norway by the reaction of KCl with nitric acid. Though the salt is an attractive fertilizer material containing the two essential nutrients, nitrogen and potassium, in plant-available form, it is expensive to produce and thus enjoys only a limited market.

Potassium Phosphates

Potassium phosphates are excellent fertilizers, and their very high analysis is an advantage that has stimulated much research in an effort to find an economical production process. However, no process has been developed that is economical enough to result in widespread production; therefore, present use is limited to special purposes for which the high cost can be justified.

At present, most of the potassium phosphates used in fertilizers are produced from potassium hydroxide or carbonate and phosphoric acid and are used in liquids for foliar application or other specialty uses.

Some of the potassium phosphates are:

	Approximate grade	
Monopotassium phosphate	KH_2PO_4	0–52–35
Dipotassium phosphate	K_2HPO_4	0–40–54
Tetrapotassium pyrophosphate	$K_4P_2O_7$	0–43–57
Potassium metaphosphate	KPO_3	0–60–40

In addition, a potassium polyphosphate solution of 0–26–27 grade has been produced from superphosphoric acid and potassium hydroxide; it contains a mixture of ortho, pyro, and higher polyphosphates.

Another process that received attention involves the use of concentrated sulfuric acid to treat potassium chloride to produce potassium bisulfate: [22, 23]

$$KCL + H_2SO_4 \rightarrow KHSO_4 + HCL$$

The hydrogen chloride is recovered as a dry gas and can be used for certain oxychlorination reactions or other industrial purposes for which elemental chlorine is now used. The potassium bisulfate is used, with more sulfuric acid, to make monopotassium phosphate and phosphoric acid. A simplified equation, representing phosphate rock as tricalcium phosphate, is:

$$\begin{aligned} &KHSO_4 + 2H_2SO_4 + Ca_3(PO_4)_2 + 6H_2O \\ &\quad \rightarrow 3CaSO_4 \cdot 2H_2O + KH_2PO_4 + H_3PO_4 \end{aligned}$$

After removal of the gypsum by filtration, the mixture of phosphoric acid and monopotassium phosphate can be ammoniated to produce a chloride-free NPK fertilizer such as 8–48–16, or monopotassium phosphate can be separated from the phosphoric acid by precipitation in methanol solution and marketed separately. If desired, the monopotassium phosphate can be converted to a pyroorpolyphosphate by pyrolysis.

The obvious advantage of the process is the recovery of HCl in a useful form. Another advantage is that a standard wet-process phosphoric acid plant can be adapted with a minor change. Numerous other processes for making potassium phosphate from KCl have been proposed, patented, or developed on a commercial scale, but none is known to be in commercial use for fertilizer purposes.

Compound Fertilizers

The preceding sections of this chapter have described the production of nitrogen fertilizers, phosphate fertilizers, and potash fertilizers from basic raw materials. These fertilizers, without further processing, are referred to as “straight materials.” Worldwide, in 2009/2010 (Fig. 24.15), straight materials accounted for about 213 million mt—about 55% of the total production. The remaining 45% was in the form of “mixed” fertilizers, that is, fertilizers that contain more than one of the major plant nutrients nitrogen, phosphorus, and potassium. Application of nitrogen fertilizer as straight material is much more prevalent than phosphate fertilizer. The reasons for these patterns are both agronomic and economic.

Agronomically, large applications of nitrogen alone—for example, in the rice production and “side dressing” of

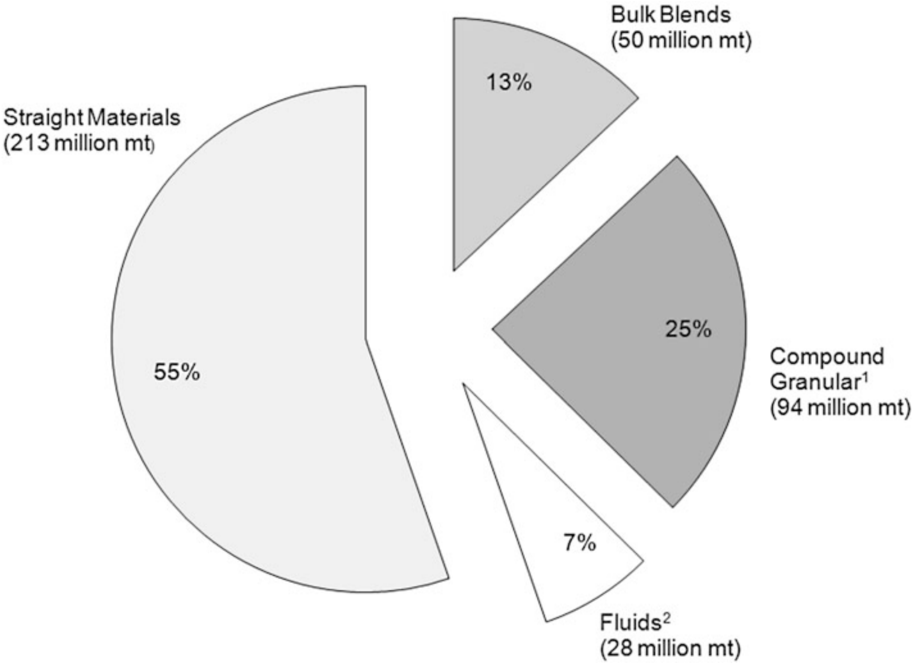


Fig. 24.15 Estimated world production of fertilizers by product type. (1) Steam/water, chemical (including ammonium phosphate), and compacted granular products. (2) Single- and multiple-nutrient fluids, including anhydrous ammonia. (3) As reported by IFA, nutrient production (2009/2010 reporting year) amounted to 163.7 million mt ($N + P_2O_5 + K_2O$). Product mt were calculated by multiplying 2.35 times the nutrient. *Source*: Derived from IFA and IFDC data files

corn—often are of considerable importance. For phosphate, there usually is no similar agronomic need for single-nutrient application; so application in mixtures, with the resultant elimination of multiple handling, is preferred. In the United States and Europe, with highly mechanized farming practices, the present trend is toward "precision" agriculture, which involves variable rates of application of nutrients at different locations in the same field based on the nutrient status. This method of application requires straight fertilizers rather than compounds [24].

Mixed fertilizers can be divided, by physical characteristics and production methods, into the following categories, all of which will be discussed in some detail:

- Nongranular mixtures
- Compound granulars
- Bulk blends
- Fluid mixtures

Nongranular Mixtures

Until about 1950, essentially all the mixed fertilizers available to farmers were of the nongranular type. Production was by simple batch weighing and blending of available nitrogen, phosphorus, and potassium ingredients, usually in locations that were convenient with respect to rail or water receipt of raw materials and out-shipment of bagged mixtures to dealers in farm areas. The source of phosphate almost invariably was nongranular SSP produced at the mixing location from in-shipped phosphate rock and locally produced sulfuric acid. In-shipment of elemental sulfur was required to produce the acid, usually by the now obsolescent "chamber" process. The major nitrogen ingredient used in the mixtures was fine-crystal by-product ammonium sulfate, and the potash source was nongranular potassium chloride. In some plants the superphosphate was "ammoniated" by reaction with ammonia solution, either in a separate operation before mixing with the other ingredients or during the mixing operation. Such ammoniation was beneficial in reducing the corrosiveness (acidity) of the superphosphate and in improving its physical condition. The ammonia also provided nitrogen at relatively low cost, but the amount that could be absorbed by the superphosphate was limited. Organic materials often were included in the mixtures as nutrient sources, bulking agents, or physical conditioners. Materials used included fish meal, packinghouse wastes, dried blood, tobacco stems, cottonseed hulls, and others. Mineral materials such as kaolin clay, diatomaceous earth, and vermiculite sometimes were included as conditioners. Batch-type drum mixers of 1- to 3-ton capacity were used with the ammonia solution piped directly into the mixer. Mixture grades were low by present-day standards, and storage and handling properties of these nongranular mixtures were generally poor. Caking problems, in particular, were considerable. Production of this type of fertilizer now persists in only a very few locations.

Compound Granulars

Continuous Ammoniator–Granulator Process. Following the close of World War II in 1945, farmer demands developed for much increased quantities of fertilizer and for products with better handling properties. Granulation, which provides an increase in particle size and the elimination of fines, was a promising avenue for the physical improvement that was pursued by researchers in both the United States and Europe. A milestone in this movement to granulation was the development, patenting, and demonstration by TVA, in 1954, of a continuous rotary-drum granulation process for mixed fertilizers [25, 26].

The rotary-drum granulator in this process is essentially the same as was pictured for the DAP granulation process. In early versions of the process, feed to the drum of SSP consisted of (1) crushed oversize product and fines, (2) nongranular SSP, (3) nongranular potassium chloride, (4) ammonium sulfate crystals, (5) sulfuric acid, and (6) anhydrous ammonia or an ammonia-containing nitrogen solution. The acid and the ammonia or ammoniating solution were sparged under the rolling bed of solids in the drum with a sparger arrangement, as shown in Fig. 24.16. With the proper balance of liquid and solid feeds, granulation occurs as a result of the rolling action in the action in the drum, and it is completed, with hardening, in a rotary-drum dryer.

This process was rapidly and widely adopted by mixed-fertilizer producers throughout the United States and later adopted by producers in Europe, Latin America, and Asia.

The process was particularly attractive because: (1) the raw materials were essentially the same as materials that were already available at most batch-type nongranular plants, (2) higher proportions of relatively cheap anhydrous ammonia or nitrogen solution could be used, (3) continuous-type processing allowed higher production rates, and (4) product physical properties were much improved as a result of granulation. Today, the same basic process, but with numerous modifications, is used in most mixed-fertilizer granulation plants. However, the number of plants has dwindled in favor of bulk blending and fluid fertilizers. One early modification to the process was the addition of phosphoric acid as a feed material, either along with or in place of sulfuric acid. Other feed materials used in some variations include TSP, MAP, ammonium nitrate, and urea. The use of urea has become more attractive in compound fertilizer production in Asia because higher grades can be achieved. The molten urea or concentrated urea solution is sprayed onto the bed of materials in the granulator and the wet product requires gentle drying. Most urea-based compound fertilizers are more hygroscopic, requiring dehumidified bulk storage facilities before the product is bagged [27].

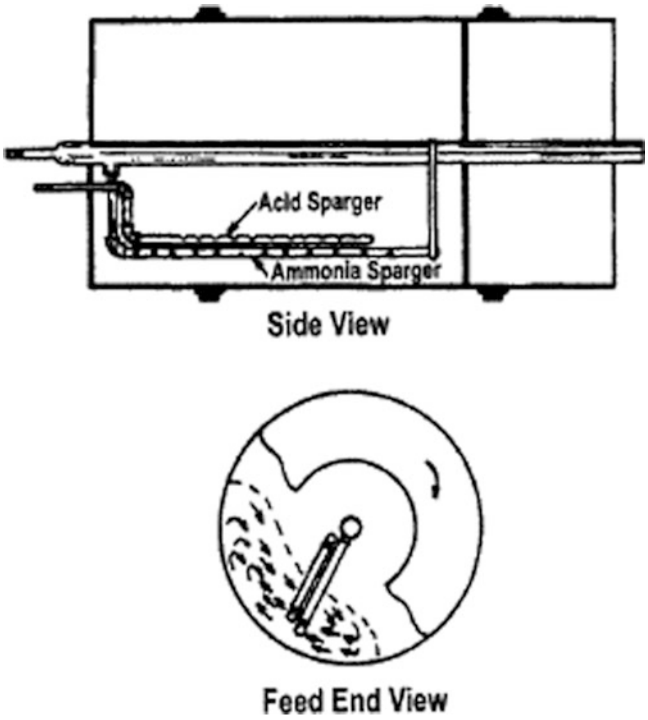


Fig. 24.16 Sparger location under rolling bed in TVA-type ammoniator-granulator *(Courtesy TVA)*

Use of Preneutralizer (Slurry Granulation). A major improvement to the basic TVA granulation process was the incorporation in most plants of a preneutralizer vessel for prereaction of phosphoric acid with ammonia. The procedure, as applied to the production of DAP, was described earlier and was pictured in Fig. 24.13. As in the DAP process, the preneutralizer used in the production of mixed fertilizers is operated at an NH_3:H_3PO_4 mole ratio of about 1.4 to give maximum ammonium phosphate solubility and thus maximum fluidity of the slurry pumped to the granulator. Subsequent further addition of ammonia in the granulator sharply decreases solubility and thus promotes solidification and hardening of granules. As in ammonium phosphate production, the iron and aluminum impurities normally present in wet-process phosphoric acid contribute significantly to the hardening of granules during ammoniation [17]. Sulfuric acid, when used, is fed directly to the granulation drum, not to the preneutralizer. The use of the preneutralizer allows the inclusion of higher proportions of acid in formulations. Also, better use is made of the heat of the ammonia–acid reaction for the evaporation of water, with resultant savings in dryer fuel.

Use of a Pipe Reactor (Melt Granulation). A more recent development that has been adopted in most U.S. granulation plants involves the substitution of a pipe reactor or a pipe-cross reactor for the preneutralizer vessel as the locale for the ammonia–acid reaction [28–32].

A simple pipe reactor differs from the pipe-cross reactor in that only one acid inlet port, usually for phosphoric acid, is provided.

The salient feature of the pipe reactor and pipe-cross reactor systems is that the heat of the ammonia–acid reaction

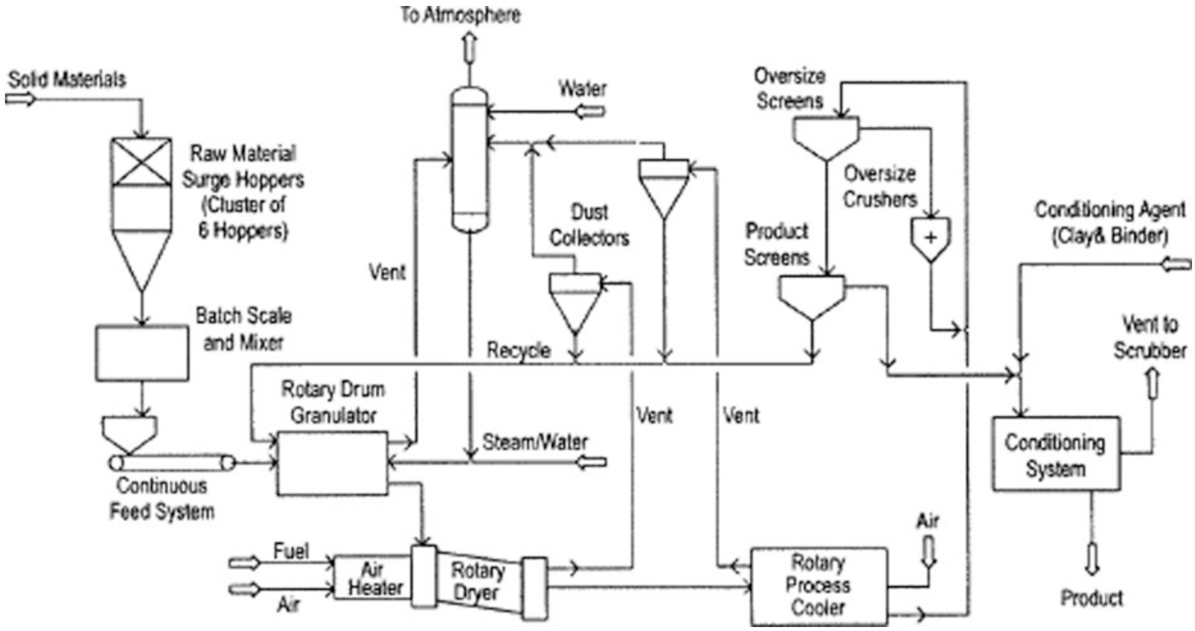


Fig. 24.17 Typical NPK fertilizer granulation plant using steam and/or water *(Source: Fertilizer Manual)*

is confined to the pipe and is efficiently utilized there to vaporize essentially all the moisture in the feed acid. The resultant steam exits the pipe discharge end in the granulator and is swept away in an air stream. The ammonium phosphate melt, which is of very low moisture content, also exits the pipe and is discharged onto the rolling bed of dry solids in the granulator. This melt serves as the binder for granule formation. The reaction temperature in the pipe reactor usually is held below 150°C to avoid a buildup of troublesome scale in the pipe. In view of this temperature limitation, there is little or no conversion of the phosphorus to polyphosphate forms. By allowing the pipe temperature to rise to about 200°C or higher, polyphosphate can be formed without scaling, but this mode of operation seldom is used in mixed-fertilizer granulation. TVA has used such high pipe temperatures to produce 11–57–0 and 28–28–0 grade ammonium polyphosphate granules in which 15–25% of the P_2O_5 is in polyphosphate form [28]. The physical properties of these products are exceptionally good, and the polyphosphate content makes them particularly suitable for dissolution to form fluid fertilizers. Because of the very low moisture content of the pipe reactor melts, even at the lower operating temperatures, the normally expensive drying of finished granular product usually is not needed: passage of the product through a cooler with an air sweep is sufficient for final drying. This results in a major fuel saving, which is a very significant economic advantage.

Steam Granulation. A simple method for granulating dry mixtures of fertilizer materials was developed in Europe in the early 1950s and still is employed there to some extent. In this procedure, which is pictured in Fig. 24.17, the feed materials are finely ground to ensure their incorporation in the granules. A dry mixture is made and then is moistened in a rotary drum or pan-type granulator with either water spray or steam, or both. This added moisture forms a salt solution by partial dissolution of the mixture ingredients. The granules form because of the plasticity and rolling action in the granulator. Initially, these granules are very soft; but, with further rolling and drying in a subsequent drying operation, salt bridging imparts the required granule hardness. In contrast to the TVA ammoniation–granulation process, the steam granulation process usually causes no significant reaction between mixture ingredients. In Europe, this feature is of considerable significance when superphosphate is used as a feed material. Ammoniation of superphosphate, which occurs in the TVA ammoniation–granulation process, converts part of the water-soluble P_2O_5 content of the superphosphate to water-insoluble but citrate-soluble forms that are beginning to be accepted as marketable phosphate fertilizer in some European countries. The citrate-soluble forms are accepted as effective, marketable phosphate forms in the United States.

Compaction Granulation. This process involves the use of mechanical force to form granules (dense particles) from finely divided solid materials [7]. In this process (Fig. 24.18) the fine materials are weighed and, depending on the desired grade, thoroughly mixed. The mixture is then continuously fed to a pressure-roll compaction machine

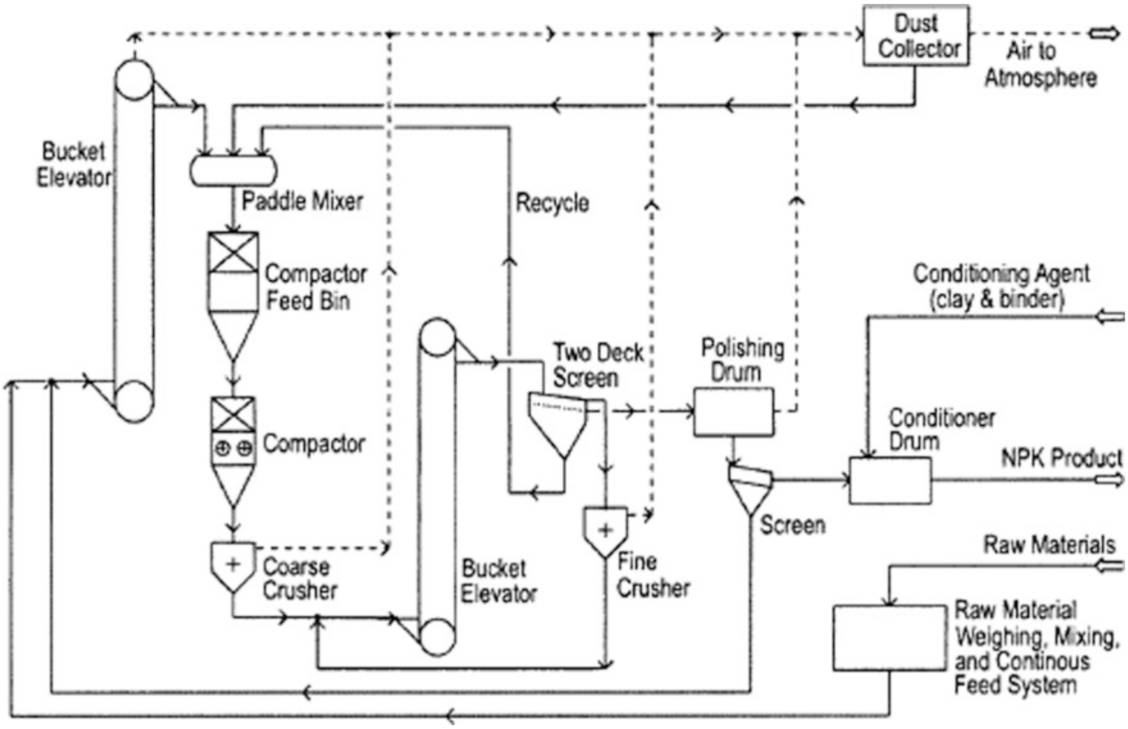


Fig. 24.18 Typical NPK fertilizer compaction granulation plant *(Source: Fertilizer Manual)*

Table 24.3 Typical pressing forces for fertilizer compaction[a]

Feed material	Pressing force[a] (kN/cm)[b]	Water content (%)	Feed particle size (mm)
Ammonium sulfate	100–120	0.5–1.0	<1.0
Potassium chloride			
Feed temperature >120°C	45–50	Nil	<1.0, with max. of 3% <0.06[c]
Feed temperature 20°C	70	Nil	
Potassium sulfate			
Feed temperature >70°C	70	1.0	<0.5
Potassium nitrate	100	0.5–1.0	<1.0
Calcium nitrate	60	Nil	<1.0
Urea	30–40	Nil	1–3
Compound fertilizer containing			
No raw phosphate rock	30–80	<1.0	<1.0
Raw phosphate rock	>80	<1.0	<1.0
Urea	30–40	<1.0	<1.0

Source: Fertilizer Manual, United Nations Industrial Development Organization and International Fertilizer Development Center, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998

[a] Indicated pressing force is for machine having 1.0-m diameter rollers

[b] 1 kN/cm = approximately 0.1 tonne/cm

[c] Size criteria apply to material feed temperature as indicated

that forms the powdered mixture into a dense, hard sheet of material usually about 2–3 cm thick and 60–100 cm wide. The compaction machine consists of two horizontally opposed rollers that turn inward toward each other. The shaft of one roller is fixed while the other is movable. The pressure from a hydraulic system is applied to the movable shaft. The pressure is dependent on the fertilizer materials. The compacted sheet of material is crushed in a controlled manner to form smaller particles. The crusher discharge is screened to obtain the desired product-size fraction. The oversize material is crushed and recycled to the screens and the undersize material is returned to the compaction machine after thorough mixing with fresh feed. This is a dry process and does not involve chemical reaction. Some typical pressing force for fertilizer materials is shown in Table 24.3.

Compacted fertilizers have a more irregular shape as compared with the more smooth and round granules

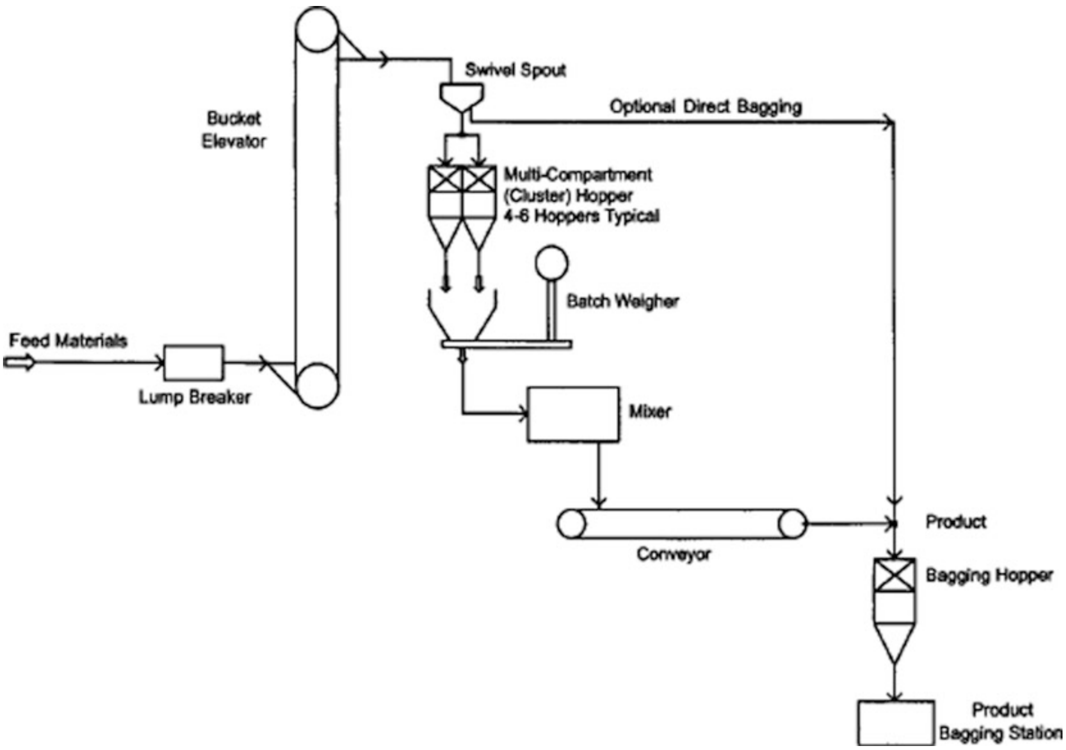


Fig. 24.19 Typical batch-type NPK fertilizer bulk-blending plant *(Source: Fertilizer Manual)*

produced by other granulation processes. This irregular shape may cause some resistance among users who are accustomed to the more traditional spherical fertilizer granules. However, compacted fertilizers are produced and used in Europe.

Bulk Blends

History and Growth. In the "compound" (mixed) fertilizer granulation processes just described, the starting materials are nitrogen, phosphorus, and potassium in the form of (1) nongranular solids, (2) liquids, or (3) gas (ammonia). When nitrogen, phosphorus, and potassium materials are all available from basic producers in granular form, as they now are, granular mixed fertilizers of almost any nutrient proportion (grade) can be made by simple proportioning and dry blending of granular nitrogen, phosphorus, and potassium materials; further granulation or other processing is not required. This is the basis of the presently popular "bulk-blending" system of mixed fertilizer preparation and distribution.

In 2008/2009 about 52 million mt of bulk-blended fertilizers was produced. This corresponded to about 13% of the global fertilizer production.

The simplicity of the bulk-blending procedure is illustrated in Fig. 24.19. The operations consist of (1) inshipment of granular fertilizer materials, (2) temporary storage, (3) proportioning, (4) blending, and (5) movement of the finished blend to the field for spreading.

Key factors in the rapid growth of bulk blending have been: (1) convenience in providing the farmer with a wide choice of plant-food ratios or grades on a "prescription" basis in relatively simple, small mixing plants located in the immediate farm area; (2) economy due to handling in bulk rather than bags; and (3) provision of spreading and other farmers services by the local blender.

Typical blending plants produce and market only 2,000–6,000 tons of blend per year. However, the present trend is for larger capacity plants covering a 30-mile radius of farming area designed to serve relatively large areas. The capital investment in blending plants is comparatively low. Blending plants operate on a quite seasonal basis, and often the operators have supplemental business enterprises.

Intermediates Used in Bulk Blending. The bulk-blending system came into being only after suitable nitrogen, phosphorus, and potassium intermediates became available in granular form. In particular, the introduction of granular DAP of 18–46–0 grade in about 1960, as described in a previous section, was a major catalyst to bulk blending. Characteristics of that product that make it an ideal material for shipping and subsequent blending are: (1) a high plant food content of 64%, (2) a high phosphorus content of 46%,

(3) excellent storage and handling properties, (4) compatibility with essentially all other fertilizer materials, and (5) economy of production near phosphate deposits. Most NPK bulk blends are prepared with DAP as the phosphate source. Granular TSP is used for no-nitrogen (PK) grades, but for NPK blends TSP has the disadvantage of being incompatible (reactive) with urea.

In the preparation of bulk blends, it is of great importance to ensure that all ingredients of the blend are reasonably matched in particle size [33]. The use of unmatched materials results in poor mixing, segregation during handling, and uneven distribution on the field. Nitrogen materials used with the greatest success in bulk blends are (1) granular urea, (2) granular ammonium nitrate, and (3) large-crystal or granulated ammonium sulfate. Prilled ammonium nitrate can also be made of large-enough particle size for good blending, but prilled urea normally is of smaller size than desirable. The chief potash material used in blends is granular-grade potassium chloride, which is sized specially by potash producers for use in blends. Although granular-size potash usually consists of irregular crusher fragments rather than well-rounded granules, studies have shown that this particle shape difference does not promote segregation or otherwise impair blend quality [34]. A somewhat smaller-size, coarse-grade potassium chloride sometimes is used in blends because it is cheaper, but segregation problems are invited. Potassium sulfate and potassium magnesium sulfate are made in granular form for the blending of specialized formulations.

The addition of micronutrients to bulk blends poses special problems because of the small quantities usually added. Although many micronutrients are available in granular form, intergranule spacing on the field can be excessive when the granular micronutrient constitutes only a very small proportion of the blend. Procedures have been devised and used for adding finely powdered micronutrients to blends together with a small amount of liquid binder [35]. The result is that each blend granule is coated with micronutrient, and thus the field distribution is better.

Operating experience and TVA studies have identified several combinations of materials that are incompatible in blends and should be avoided: (1) unammoniated superphosphates react with urea, and the result is wetting and caking and (2) urea in contact with ammonium nitrate forms an extremely soluble hygroscopic mixture that liquefies in the presence of only minute amounts of moisture. Therefore, these two materials can never be used together in the same blend. Essentially all other combinations of available fertilizer materials are compatible in blends.

Equipment and Operating Technique. There are different arrangements of equipment in the blending plants throughout the world. The operations in such plants generally consist of: (1) receipt of the intermediate granular components in bulk by truck or railroad car, (2) storage of the intermediates separately in bins, (3) removal of the granular blend components from storage and weighing in desired proportions, (4) mixing to obtain uniformity, and (5) discharge of the mixed batch directly into a spreader truck or into a holding bin. The most popular type of mixer is the batch-type rotary drum.

Such mixing plants can be highly automated so that the output with a mixer of only 1-ton capacity can be as much as 15–20 tons/h. Mixers of several types are available with capacities ranging between 5 and 8 tons per batch; in some cases batch mixer capacities can be 10 tons per batch [36].

Problems in Bulk Blending. Bulk blending is a simple practice that can provide uniform mixtures in almost any desired proportions. However, unless proper materials are used and certain handling precautions are taken, segregation can occur, and nonuniform blends will result. The factors involved and the precautions that are necessary have been described in several publications [3]. The main requirements are the use of granular materials with a well-matched range of particle size and the handling of the mixture after blending in ways that will minimize segregation. The mixing step is seldom a problem; it is quite easy to obtain a uniform blend by the use of proper materials and a few minutes of mixing. However, maintaining this uniformity during handling on the way to the farm and onto the soil requires precautions. Handling procedures that may cause segregation include coning (as occurs if the blend is allowed to drop from a discharge pipe onto a pile), vibration in hauling vehicles, and ballistic action imparted by some types of spreaders.

The problem of adding small quantities of micronutrient materials to bulk blends was mentioned earlier. This problem is being met by either using micronutrients in granular form or coating the blend particles with finely ground micronutrients and a liquid binder [35].

Fluid Mixtures

Mention was made earlier of the significant quantity of nitrogen applied to farmlands in the form of urea–ammonium nitrate solution and aqua ammonia. Since the late 1950s, there has also been a growing practice of applying mixed fertilizers (NP and NPK) in fluid form. In fact, the growth of the fluids market paralleled bulk-blended products. In 2002/2003 about 25 million mt of fluids was produced and corresponded to 7% of the world production. North America, and Europe are major markets for fluids. Fluids are also used in high-value crops in some countries of the Middle East and Latin America.

Numerous advantages have been cited to explain the growth of fluid mixed fertilizers [3]. Of these, the most significant probably is the ease and the precision with which fluids can be applied to the soil. When placement of the fertilizer is important, as in row fertilization, fluids often have an advantage. Also important is the adaptability of fluids to the homogeneous incorporation of micronutrients, herbicides, and insecticides. Another factor that encouraged the adoption of fluid mixtures was the prior existence of equipment and facilities for handling and applying nitrogen fluids; it was only logical for this practice to be expanded to include the preparation and application of mixtures. There are some claims of agronomic superiority for fluid fertilizers over similarly applied solids, but such claims are debatable and are not widely accepted.

Solution Fertilizers. Fluid fertilizers in use today can be divided into two major categories: (1) solutions, in which the plant food content is entirely in solution, and (2) suspensions, in which the plant food content exceeds its solubility, and the excess is held in suspension, usually by the addition of a suspending agent. The solution-type mixtures were the first to achieve commercial importance. Beginning about 1950, the practice of making solution mixtures was begun with pure electric-furnace orthophosphoric acid (54% P_2O_5) as the source of phosphorus. Ammoniation of this acid yielded an 8–24–0 grade ammonium orthophosphate solution that could be shipped to mixing plants in farm areas and used as a base for blending with other fertilizer materials. Blending usually was with urea–ammonium nitrate solutions (28–32% N) and a pure grade (solution grade) potassium chloride.

With the system that employed these raw materials, numerous advantages of fluid mixtures were established, but a need to improve the economics of the procedure was obvious. Pure electric-furnace acid is much more expensive than wet-process acid, so ways were sought to use the cheaper but impure wet-process acids. Direct ammoniation of wet-process acid causes precipitation of numerous impurities as an intolerable, gelatinous sludge. Also, plant-food solubilities are undesirably limited in the orthophosphate system, so the concentration (grade) of the product made with electric-furnace orthophosphoric acid was limited to 8–24–0. A series of developments by TVA led the way to solving both of these problems.

The problem of limited plant-food solubility was much alleviated by TVA's introduction of superphosphoric acid in 1957 [37]. Superphosphoric acids are acids of high P_2O_5 content (up to 80% P_2O_5) that contain significant proportions of polyphosphates (nonortho, poly merized phosphates). The superphosphoric acid first produced by TVA was a pure electric-furnace product of 76% P_2O_5 content, of which about 50% was in polyphosphate form. With ammoniation of this acid, it was possible to produce an 11–37–0 grade base solution instead of the 8–24–0 grade maximum made from ortho acid. Demonstration use of 11–37–0 base in a number of mixing plants showed that solution mixtures of relatively high concentration could be made. Also, it was found that the polyphosphate content enhanced the solubility of micronutrient additives by a sequestering action.

The unfavorable economics of the electric-furnace acid process, however, discouraged the commercial production of electric-furnace superphosphoric acid; so attention was directed toward concentration of the cheaper wet-process phosphoric acid to give wet-process superphosphoric acid. Satisfactory concentration methods were developed [38], but the effects of acid impurities on the handling properties of the concentrated acids limited the final P_2O_5 concentration and hence the polyphosphate content. The best handling properties were obtained with acids made from calcined rocks and with P_2O_5 concentration limited to 68–70% P_2O_5. In acid of this concentration, from 20 to 35% of the P_2O_5 is in the polyphosphate form, and the acid is referred to as low-conversion superphosphoric acid. Such acid today is the chief source of P_2O_5 for solution-type fertilizers.

The most popular scheme by which solution-type fertilizers are made from low-conversion, wet-process superphosphoric acid is outlined in Fig. 24.20. The first step is ammoniation of the acid to produce ammonium polyphosphate base solution. By carrying out this ammoniation in a pipe reactor, a temperature of over 200°C is developed, which promotes an increase in polyphosphate content to 65–70% of total P_2O_5 in the ammoniated acid [39, 40]. The grade of the resultant base is usually 10–34–0; an 11–37–0 grade can be made, but its storage properties are not as good, especially in cold weather. A sketch of a typical pipe-reactor system for the production of high-polyphosphate base solution from low-conversion, wet-process superphosphoric acid is shown in Fig. 24.21.

More than 100 U.S. plants use this pipe-reactor process to produce an estimated two million metric tons of 10–35–0 or 11–37–0 grade product per year. Other countries such as Belgium and France also use this process to produce ammonium polyphosphate solution. A considerable amount of the solution is used for direct application in the U.S. wheat belt and other areas in which potassium is not deficient. The rest is used in small mix plants to produce mixtures. The ammonium polyphosphate base solution is mixed with nitrogen solutions containing 28–32% N (urea–ammonium nitrate solutions) and sometimes also with potash to produce such clear liquid grades as 21–7–0, 7–21–7, and 8–8–8. As with bulk blending, the solution-mixing operation usually is carried out in small mixing plants located in farm areas (Fig. 24.22).

Suspension Fertilizers. Suspension fertilizers were developed chiefly as a means of breaking the grade barriers imposed on solution fertilizers by solubility limitations. The low solubility of potassium chloride was particularly troublesome, in that high-potash solution grades were not possible.

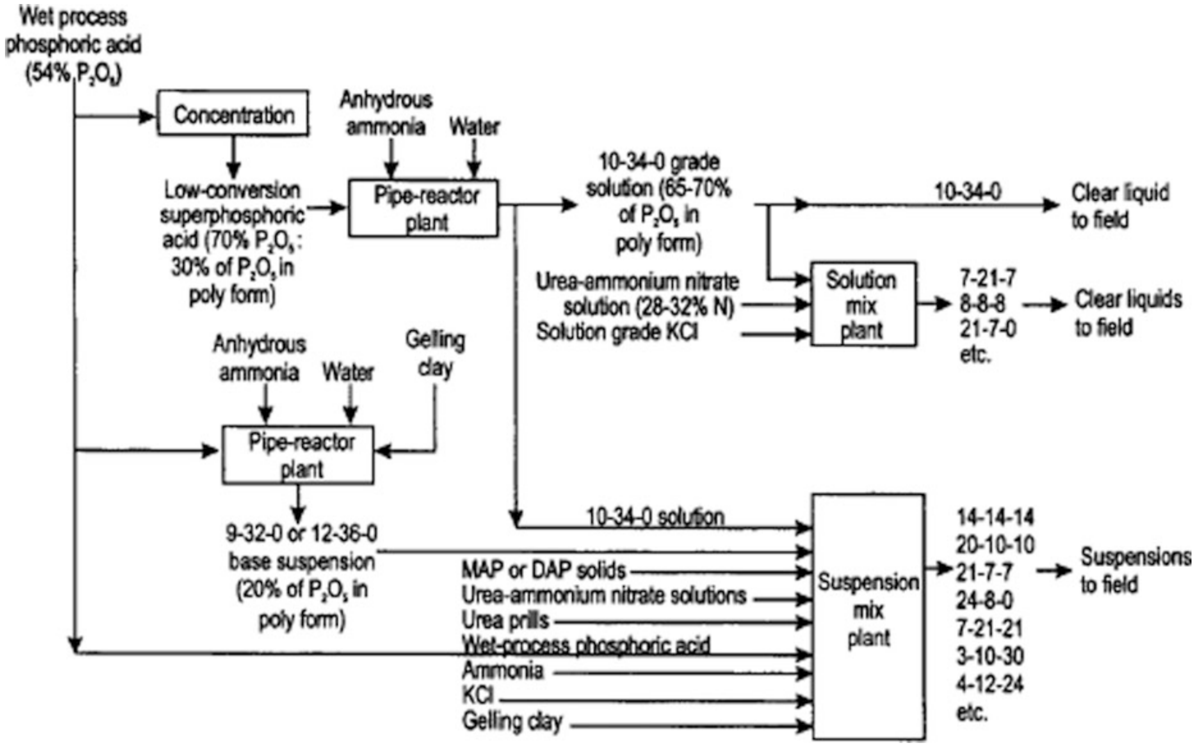


Fig. 24.20 Production methods for solution-type and suspension-type mixed fertilizers

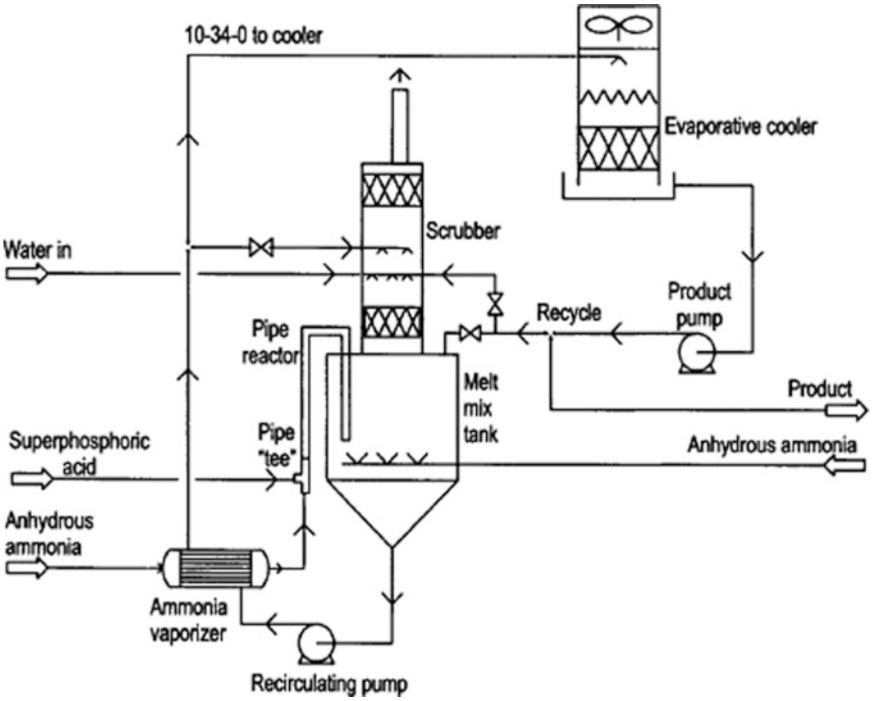


Fig. 24.21 Typical pipe-reactor plant for production of ammonium polyphosphate solution *(Source: Fertilizer Manual)*

The feasibility of preparing and handling suspension-type fertilizers was demonstrated by TVA and others in 1959 [41, 42]. Two essential requirements for the preparation of a successful suspension were shown to be: (1) small particle size of the suspended solids and (2) the addition of a gelling-type clay as a suspending agent along with sufficient

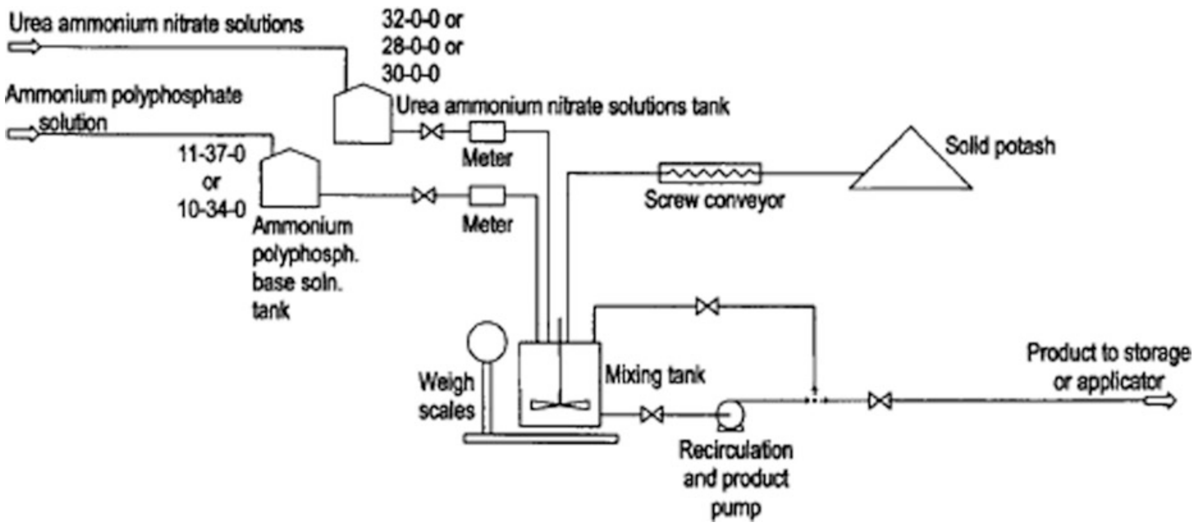


Fig. 24.22 Liquid fertilizer mix plant *(Source: Fertilizer Manual)*

high-shear agitation to properly disperse the clay. The most satisfactory gelling-type clay is attapulgite [43]. Attapulgite is a hydrated magnesium aluminum silicate, which is composed of needle-shaped crystals so small that they can be seen only with the use of high magnification. In the dry state, these crystals bond together in bundles as larger particles. These bundles must be disrupted and the individual crystals freed before they can assume the structural arrangement that is characteristic of a gel. When this disruption of the crystal bundles is carried out in the presence of an electrolyte, such as any of the common fertilizer salts present in fluid fertilizers, the individual crystals immediately assume the regular structural orientation of a weak gel, and thus provide the desired suspending action.

Wetting of the clay by the fluid fertilizer causes some of the required disruption of crystal bundles and freeing of individual crystals but not enough to provide good gelling without high-shear agitation. The addition of 2–3% of attapulgite is sufficient for most fertilizer suspensions. Instead of providing high-shear agitation for the entire batch of suspension, some operators carry out a pre-gelling operation in which high-shear agitation is applied only to the clay plus a limited amount of water or fertilizer solution. The pre-gelled mixture then is added to the major body of fertilizer with only mild mixing. Pre-gelled clay is also marked as liquid clay and is transported in tank cars. The provision of a high-shear mixer is essential not only for gelling of the suspending clay but also for disintegration of any granular products such as DAP or MAP that might be used in the suspension. A high-shear recirculation pump is also an aid to gelling of the clay.

Figure 24.23 shows the layout of a typical suspension mix plant in which all the major suspension ingredients are received as solids. This is presently a popular mode of operation. Raw material costs for such a plant often are less than for bulk blending because nongranular materials can be used. Piping for the optional addition of phosphoric acid and ammonia is shown. Such addition develops heat, which hastens the disintegration of some solids.

The storage and handling properties of suspensions can be enhanced by the inclusion of some polyphosphate in the suspension. Low-temperature storage properties, in particular, are enhanced. One popular method of providing polyphosphate is to supply all or part of the P_2O_5 as 10–34–0 ammonium polyphosphate base solution, such as is used in solution fertilizers. A second method is the use of a pipe reactor to prepare 9–32–0 or 12–36–0 grade ammonium polyphosphate base suspension from relatively inexpensive wet-process orthophosphoric acid (54% P_2O_5), ammonia, and gelling clay [44]. About 20% of the P_2O_5 in the 9–32–0 form or 12–36–0 product is in polyphosphate form. The clay content is 2%. When these base solutions are used for mixing with other suspension ingredients, their clay content often is sufficient to provide all the clay required in the final suspension. Commercial acceptance of 9–32–0 or 12–36–0 base suspension preparation has not yet occurred. At present, apparently it is most economical to furnish polyphosphate as 10–34–0 ammonium polyphosphate solution.

Controlled-Release Fertilizers

The term “controlled-release fertilizer,” as used in the following discussion, refers to fertilizers that release, either by design or naturally, their nutrient content over an extended

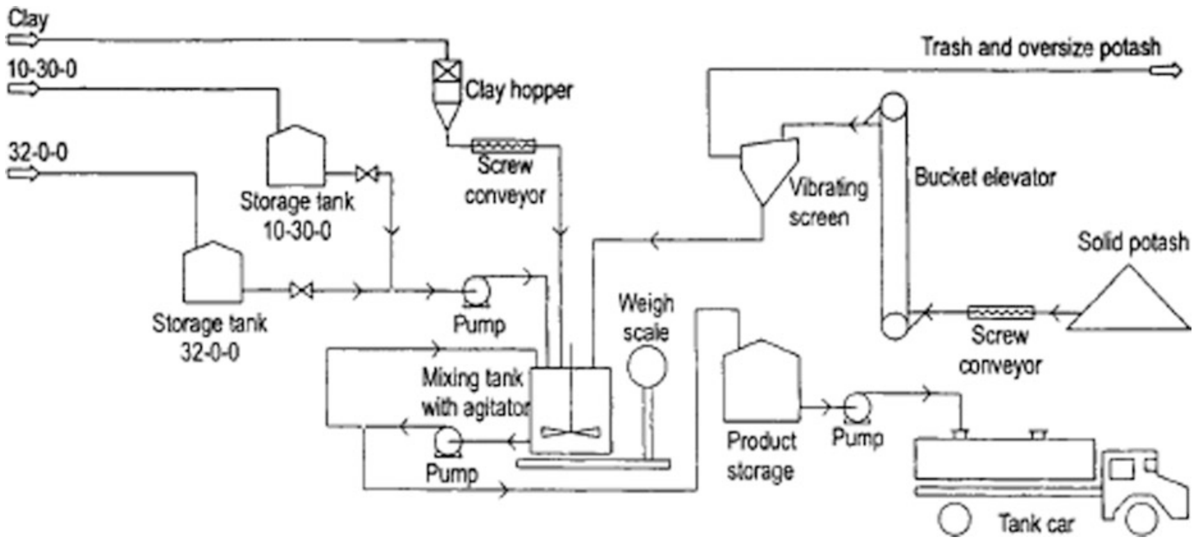


Fig. 24.23 Typical mix plant for production of suspension mixtures *(Source: Fertilizer Manual)*

period of time. Potential advantages claimed for controlled-release fertilizers are increased efficiency of uptake by plants; minimization of losses by leaching, fixation, or decomposition; a reduction in the release of nutrients and by-products to air and water; and avoidance of burning of vegetation or damage to seedlings. In spite of this impressive list of potential advantages, the controlled-release fertilizers have mainly been applied to high-value crops and/or ornamental plants and golf courses. However, there has been progressive increase of controlled-release fertilizers on rice in Japan and more recently corn in the United States and rice in China. The combined total controlled-release fertilizer production in North America (73%), Western Europe and Japan totaled 713,000 tons or 86% of total production in 2009. Agriculture in the "three" regions accounted for only 24% of the consumption. However, in Japan 90% of controlled-release fertilizer is used on agricultural crops. Large and growing markets are emerging in Asia, particularly in China where increase food demand combined with land and labor shortage is expected to raise controlled-release fertilizer consumption to more than one million tons/year.

Controlled-Release Phosphate Fertilizers. Several controlled-release phosphates have been produced commercially. Important among them are ground phosphate rock, basic slag, and fused CMP. These products have been discussed earlier.

Most of the insoluble or slightly soluble phosphate materials are used because they are less expensive than soluble phosphates or because they supply other elements. These products are rarely agronomically superior to soluble phosphates. However, some of them are valued for specialty uses because they may be safely placed in contact with seeds or roots in some situations. Slightly soluble phosphates are found to be useful in tropical agriculture where conditions are quite different from those in temperate zones.

Controlled-Release Nitrogen Fertilizers. Because of economics and environmental considerations, the slow-release nitrogen fertilizers are much more important than phosphate (or potash), particularly under certain soil and climatic conditions. In most cases unutilized phosphate and potash remain available for subsequent crops. In contrast, there is seldom much carryover of fertilizer nitrogen from one crop to the next. Nitrogen applied as fertilizer is lost from the soil through: (1) leaching or surface drainage, (2) volatilization as ammonia, and (3) decomposition. Because of these losses, nitrogen use efficiency by crops seldom exceeds 50 percent and may be as low as 20 percent under flooded rice conditions when grown under conditions of intermittent flooding or poor water control. Thus, improved utilization is particularly important because it would favorably impact the cost of production and simultaneously reduce the release of nitrogen compounds to air and water [45].

Slightly Soluble Materials. One group of controlled-release nitrogen fertilizers comprises chemical compounds that are inherently only slightly soluble in water or soil solution. Urea–aldehyde compounds are the principal representatives of this group that are produced commercially. Isobutylidene diurea (IBDU) is produced in Japan through the reaction of urea with isobutyraldehyde in a 2:1 mole ratio. When pure, it contains 32.18 percent N [46]. Crotonylidene diurea (CDU), also called cyclodiurea, is produced in Japan and Germany through the reaction of urea with crotonaldehyde or acetaldehyde. The pure compound contains about 32 percent nitrogen.

Urea–formaldehyde reaction products, usually called "ureaform," are produced by about six manufacturers in the United States and several other countries. Unlike IBDU and CDU, ureaform is not a definite chemical compound. It contains methylene ureas of different chain lengths; the solubility increases with decrease in chain length. It usually contains about 38 percent nitrogen.

Perhaps the most useful component of ureaform as a slow-release nitrogen material is trimethylene tetraurea; shorter chain lengths are too rapidly decomposed, and longer chains are highly resistant to decomposition. Unfortunately, it is very difficult to produce commercially a pure trimethylene tetraurea; ureaform contains both shorter and longer chains.

A typical ureaform may contain 30 percent of its nitrogen in forms that are soluble in cold water (25 C). The quality of the remaining 70 percent is judged by the percentage that is soluble in hot (boiling) water as determined by prescribed analytical procedures. At least 40 percent of the nitrogen insoluble in cold water should be soluble in hot water for acceptable agronomic response; typical values are 50–70 percent.

Ureaform is produced by reaction of urea with formaldehyde in aqueous solution in the presence of an acid that catalyzes the reaction. The degree of polymerization depends on the mole ratio of urea to formaldehyde and on the pH, time, and temperature of reaction. The reaction is stopped by neutralizing the acid. Production processes have been described by Clark et al. [47] and by Kolterman and Rennie [48].

In Japan, ureaform is also produced as a component of compound fertilizers by the addition of formaldehyde solution to the urea solution fed to the granulator [49]. In one process a methylene urea reactor is used to react urea and formaldehyde solution to produce a slurry that feeds directly to a granulation plant with other materials to make compound fertilizers [50]. Some slow-release nitrogen is formed, and the physical properties of the products are improved.

In the United States, ammoniating solutions containing urea, formaldehyde, and ammonia are used to prepare compound fertilizer containing water-insoluble nitrogen. The composition of one such solution is: urea—35.3 percent, formaldehyde—8.0 percent, CO_2—8.4 percent, NH_3—25.0 percent, and H_2O—23.3 percent. Other solutions with high formaldehyde content (up to 60%) are available for use with additional urea [51].

The addition of sulfuric or phosphoric acid (or superphosphate) neutralizes the ammonia and catalyzes the formation of methylene urea. Products containing about 20–30 percent of their nitrogen in a water-insoluble form are produced. Good control of conditions in the granulator and dryer is needed to produce a satisfactory quality of water-insoluble nitrogen. In some cases, a prereactor is used prior to the granulator [50].

The consumption of ureaform in the United States is approximately 75,000 tons/year, not including that formed in compound fertilizers from solutions. Most of the production goes into compound fertilizer for specialty uses, such as for lawns, flower gardens, and golf courses. The wholesale price of ureaform per unit of nitrogen is about three times that of urea, ammonium nitrate, or ammonium sulfate.

In tests made by TVA and IFDC of many other organic nitrogen-containing compounds as fertilizers, some proved readily available, some slowly available, some inert, and some toxic. Among those that showed promise as slow-release fertilizers were oxamide, glycouril, cyanuric acid, ammeline, and ammelide. For lack of an economical process, none of these materials has been manufactured commercially for fertilizer use.

Magnesium ammonium phosphate ($MgNH_4PO_4$) is a slowly soluble source of nitrogen although it contains more phosphate than nitrogen. The commercial product, called "Mag Amp," contains a variable amount of water of crystallization and some impurities; a typical grade is 8–40–0. A similar product containing potash is marketed under the trade name "En Mag" in the United Kingdom. The grade is 5–24–10. The rate of solution of these slightly soluble materials is controlled by the size and hardness of the granules.

Coated Soluble Materials. A wide range of materials and techniques has been explored with the object of making controlled-release fertilizers by coating soluble fertilizer materials with plastic films, resins, waxes, asphaltic materials, or other barriers. The first commercial production of coated controlled-release fertilizer utilizes a process developed by Archer Daniels Midland Company (ADM). The main component of the coating is a copolymer of dicyclopentadiene with a glycerol ester [52]. Applied in several layers that vary in composition, the coating releases fertilizer solution by osmotic exchange with moisture from the soil. The coated granular fertilizer was manufactured by ADM under the trade name Osmocote starting in 1964. It now is produced and marketed by Sierra Chemical Company under exclusive license from ADM. Three grades of Osmocote are available: 14–14–14, 18–9–9, and 16–4–8. The weight of the coating ranges from 10 to 15 percent of the gross weight. Osmocote products are recommended for turf, floriculture, nursery stock, and high-value row crops.

Sulfur-coated urea (SCU) is a controlled-release material developed by TVA during the 1960s and 1970s [53]. Sulfur was selected as the coating material on the basis of economy and efficiency after many coating materials were tried. Urea was selected as the material to be coated because of its high nitrogen content, low cost, and commercial availability.

Initial studies indicated that sulfur alone was not adequately resistant to moisture penetration. However, it was discovered that the addition of an oil wax sealant over the sulfur coating provided a successful coating. The process has been commercialized and the product is mainly used for turf and golf courses.

In this process granular urea is preheated in the first rotating drum to 77 C with electric radiant heaters to prevent the sulfur from freezing too rapidly on the granules. The molten sulfur is atomized and sprayed onto the rolling bed of granules in the second drum. Wax is applied in the third drum. The coated granules are then cooled in a fluidized bed cooler, and an absorbent powder is applied in a fourth drum to eliminate the tacky condition of the wax and to prevent the granules from floating when applied in water (as in a rice paddy).

The rate of dissolution of coated urea can be varied by varying the thickness of the sulfur coating and, to some extent, by varying the amount of sealant. In addition to the controlled-release characteristics, coated urea has excellent storage and handling characteristics and might even be handled in bulk in humid climates. Also, it may be blended with TSPs or SSPs, which are incompatible with uncoated urea. Also the sulfur coating is a slowly available source of sulfur, which would be an advantage in sulfur-deficient areas.

The sulfur-coating technique may be applied to fertilizers other than urea. Experimental work has included a coating of DAP, potassium chloride, potassium sulfate, and various compound fertilizers, including a pan-granulated urea-potassium chloride product of 32–0–16 grade.

Nitrification Inhibitors. Nitrogen fertilizers in the ammonium form are immobilized by sorption on clay particles and hence are resistant to leaching. However, the ammonium form is converted to the nitrate form under aerobic conditions by microbiological processes known as nitrification. The rate of conversion depends on the temperature; below 10 C it is very slow, but at 20 C and higher it is rapid. The nitrate form is more readily available to most plants although some plants (e.g., rice) can readily use the ammonium form. Thus, some of the purposes of controlled release (resistance to leaching, prevention of denitrification losses, and delayed availability) may be attained by delaying nitrification of ammonium nitrogen.

Numerous organic chemicals have been identified as nitrification inhibitors. The inhibition is due to the toxicity of these chemicals to organisms that convert ammonium nitrogen to nitrite, which is the first step of the nitrification process.

One of the nitrification inhibitors that has received attention in the United States is 2-chloro-6-(trichloromethyl) pyridine which is promoted under the name N-Serve. The minimum concentration of N-Serve in the soil for delaying nitrification at least 6 weeks is said to range from 0.5 to 10 ppm.

The N-serve literature points out that a nitrification inhibitor is helpful only when conditions favor high nitrogen loss from the soil. Such conditions are heavy rainfall or heavy irrigation, coarse-textured soil, and soil in the pH range where nitrification readily occurs. Nitrate thus formed will leach down into anaerobic soil layers.

Potassium azide (2–6% KN3) dissolved in anhydrous ammonia was effective in lowering the rate of nitrification in tests conducted by the U.S. Department of Agriculture in cooperation with agricultural experiment stations in the states of Louisiana and Washington. PPG Industries, Inc., is exploring this and other agricultural uses (e.g. as a herbicide in rice culture) for the azide.

Another nitrification inhibitor that has received much attention in Japan is 2-amino-4-chloro-6-methyl pyrimidine. It is manufactured by Mitsui Toatsu Chemicals, Inc. (formerly Toyo Koatsu Industries, Inc.) under the trade name AM. Other nitrification inhibitors that have been used in Japan in compound fertilizers are sulfathiazole, dicyandiamide, thiourea, N-2, 5-dichorphenyl succinamide, 4-amino-1, 2, 4-triazole hydrochloride, and guanylthiourea [50].

In flooded or waterlogged soils, as for example, in a continuously flooded rice paddy, anaerobic (reducing) conditions exist in the soil below the surface, and hence nitrification of ammoniacal nitrogen does not occur in this layer. Therefore, when ammonia, ammonium salts, or urea (which hydrolyzes in the soil to ammonia and carbon dioxide) is placed in the reducing zone of the soil, it remains in the ammonium form and is more or less immobilized by sorption on the clay. Some of the advantages of controlled release may be attained by placement of ammoniacal nitrogen in the reducing zone of the rice paddy soil. However, this placement is difficult since applicators quickly become clogged with mud and wet fertilizer. To facilitate subsurface placement, large granules or briquettes of urea have been produced experimentally that can be pressed into the soil by hand (or foot) (Fig. 24.24). Experiments with this placement have shown substantially increased efficiency of nitrogen utilization by rice as compared with the usual broadcast placement.

IFDC has designed a village-level briquetting machine that can produce urea briquettes ranging in weight from 1 to 2.7 g (Fig. 24.25). The capacity of these machines varies from 20 kg/h to 1 ton/h and can be operated intermittently. As of 2004 more than 1,000 of these machines are operating in Bangladesh, providing briquettes for nearly 500,000 ha of rice paddy. Using this technology the farmers are able to increase rice yields by about 20% while using 25% less urea. This technology is now being tested in Vietnam, Cambodia, and several countries in Africa. A hand-held applicator for deep placement of briquettes is also being tested in Bangladesh.

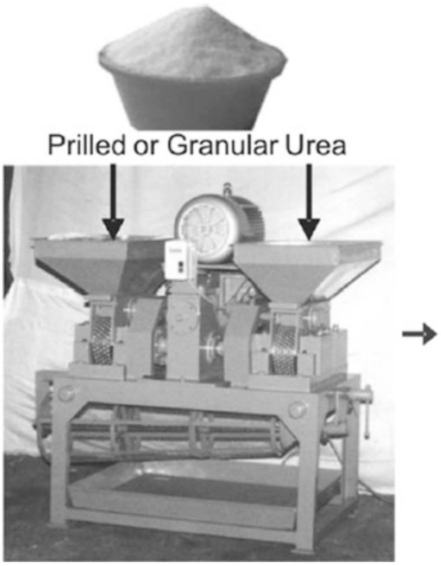


Fig. 24.24 Compacted urea product and briquetted urea particles of three particle sizes (from *left*)

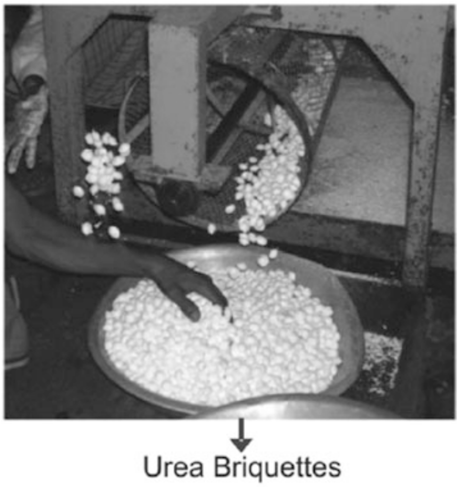


Fig. 24.25 A village-level briquetting machine

Physical Quality of Fertilizers

Physical quality is an important determinant used by a farmer-consumer in selecting a particular fertilizer. The acceptability of a fertilizer in the marketplace depends not only on its nutrient content but also on its physical quality. Although the available nutrient content and its effect on crop response is the ultimate reason for the use of fertilizers, physical quality is often the reason for selecting one fertilizer over another. It is important that fertilizers remain free-flowing (noncaking), relatively nondusty, and be able to withstand a reasonable amount of exposure to normal atmospheric humidity to withstand transportation, distribution, and application. Many production techniques, including various physical conditioning treatments, have been developed to provide favorable physical properties [54].

The quality of a fertilizer product is basically judged by two criteria: chemical properties and physical properties. The chemical properties (nutrient content, form, and availability to the crop) and their effectiveness for crop growth are the main reasons for using a fertilizer. In contrast, physical properties of a fertilizer product are important with respect to processing, handling, storage, and field application.

The consumer, through simple observation, cannot ascertain whether a fertilizer actually meets the guaranteed nutrient content; however, federal or state laws normally regulate the nutrient content thereby giving some protection to the consumer. In contrast, physical condition usually is a matter only of agreement between purchaser and supplier. Normally fertilizer regulatory laws do not regulate fertilizer physical properties. However, the consumer can certainly see whether a fertilizer is caked or free-flowing, dusty or dust-free, wet and sticky or dry. The most frequently encountered problems resulting from deficiencies in physical properties are caking (agglomeration or lump formation), dustiness, poor flowability, excessive hygroscopicity (moisture-absorptioncharacteristics), and segregation (non uniformity of composition throughout a fertilizer lot).

The physical quality of a fertilizer can be evaluated by many different methods that can vary from simple observation to sophisticated techniques. Producers have adopted and/or developed methods to evaluate their particular products. International, national, and regional organizations have adopted, developed, and/or published procedures intended for general application to a variety of fertilizer products and raw materials [3, 55, 56].

Next Generation of Fertilizers

As discussed earlier, fertilizers have an important role in production and productivity increases. Most of the fertilizers in use today were developed between the 1940s and 1980s by scientists of TVA, an agency of the U.S. Government, along with those from the fertilizer industry and state land-grant university research and extension systems. The bulk of R&D was conducted on the TVA Reservation in Muscle Shoals, Alabama, USA. Unfortunately, since the early 1990s, public funds for new fertilizer technology R&D programs have essentially disappeared. Further, fertilizer companies are currently investing almost nothing in R&D. Not surprisingly, over the past 25 years, no new, significantly more efficient fertilizer products have been developed that are cost-effective for use on food crops by farmers in less developed countries. Yet efficiency of fertilizer technology is closely linked with key factors determining the future of humankind—food security, availability of natural resources, and the environment.

The world's resources of rock phosphorus (P) and potassium (K) from which fertilizers are made are finite. Nitrogen production is extremely energy intensive: it takes the energy equivalent contained in four barrels of oil to make one ton of urea, i.e., the equivalent of 480 million barrels of oil is needed to produce today's supply of 120 million metric tons (mmt). Moreover, current fertilizer technologies result in great losses either at the production or utilization stage. Up to two-thirds of the urea applied is wasted due to crop use inefficiencies, and the current technologies used to produce phosphate concentrate from mined ore generate significant losses of phosphate.

The potential impact of improving the efficiency of fertilizer technology is vast. For example, a gradual 50% improvement in average nitrogen use efficiency in cereal crops by 2050 would:

- Reduce total accumulated nitrogen losses from 1,281 to 817 mmt.
- Save $403 billion (at $400/mt for urea).
- Reduce by the equivalent of four billion barrels of oil the amount of energy used to produce urea fertilizer.
- Reduce N losses significantly:
- Volatile N_2O-NO_X by about 974 mmt of CO_2 equivalents;
- Leaching to freshwater and marine ecosystems with the latter particularly affected by N.

Most significantly, more efficient fertilizer technologies can help better meet future food security at farm, national, and global levels. Farmers will have to apply less fertilizer to get more yield, while improving their profits. National governments will be able to reduce their importation of fertilizer and/or, in case of subsidies, the drain on their budgets, while increasing agricultural production. At the global level, food security would increase while negative impacts on environment and resource availability would decrease.

Driven by these issues, the International Fertilizer Development Center (IFDC) launched the Virtual Fertilizer Research Center (VFRC) to create the next generation of fertilizers and production technologies to help feed the world's growing population and provide sustainable food security. It is *virtual* in the sense that it is intended to be comprised of the work of multiple research institutions cooperating via communications technology to advance a unified research agenda.

References

1. Smil V (1997) Cycles of life. Scientific American Library, New York, NY
2. Food and Agriculture Organization of the United Nations (FAO) (2002) FAOSTAT database. http://faostat.fao.org. Rome, Italy
3. United Nations Industrial Development Organization (UNIDO) and International Fertilizer Development Center (IFDC) (1998) Fertilizer manual. Kluwer Academic Publishers, Dordrecht, The Netherlands
4. Herring JR, Fantel JR (1993) Phosphate rock demand into the next century: impact on world food supply. Nonrenewable Resources 2(3):226–241
5. British Petroleum–Amoco (1998) Statistical Review of World Energy
6. Van Kauwenberg, British Petroleum (2010) BP Statistical Review of World Energy, June, 2010. http://www.bp.com/statisticsalreview
7. International Fertilizer Development Center (IFDC) (1983) Compaction—alternative approach for granular fertilizers. IFDC Technical Bulletin T-25
8. McVicar MH (1970) Using commercial fertilizers, 3rd edn. The Interstate Printers and Publishers, Inc
9. Hignett TP (1971) History of world fertilisers and manufacturing processes. Ind Chem Manufacturer 9(4):13–17
10. Young RD, Achorn FP (1978) Trends in U.S. fertilizer technology. TVA Bull. Y-133. Tennessee Valley Authority, Muscle Shoals, AL, pp 5–7
11. (1980) The role of phosphorus in agriculture. Am. Soc. Agron., Crop Science Soc. Am., and Soil Science Soc. Am. Madison, WI, pp 198–199
12. Hignett TP (1965) Nitrophosphate process advantages and disadvantages. In Proc. 15th Annual Meeting Fert. Ind. Round Table. pp 92–95
13. Getsinger JG, Houston EC, Achorn FP (1957) Production of diammonium phosphate by continuous vacuum crystallization. J Agr Food Chem 5:433–436
14. Young RD, Hicks GC, Davis CH (1962) TVA process for production of granular diammonium phosphate. J Agr Food Chem 10:442–447
15. U. S. Patent 3,153,574
16. (1970) New Developments in Fertilizer Technology, 8th Demonstration. TVA Bull. Y-12. Tennessee Valley Authority
17. Young RD, Hicks GC (1967) Production of monoammonium phosphate in a TVA-type ammonium phosphate granulation system. Comm Fert 114(2):26–27
18. Lehr JR, McClellan GH (1973) Phosphate rocks; important factors in their economic and technical evaluation. CENTO Symposium on the Mining and Beneficiation of Fertilizer Minerals. November 19–24, pp 192–242
19. Pelovski Y, Garrett MK. Partially acidulated phosphate rock fertilisers—agronomic, environmental and production aspects. Proceedings No. 364, The Fertiliser Society
20. Yates LD, Nielsson FT, Fox EJ, Magness RM (1953) Enriched and concentrated superphosphate. Ind Eng Chem 45(3):681–690
21. Walthall JH, Bridger GL (1943) Fertilizer by fusion of rock phosphate with olivine. Ind Eng Chem 35(7):744–777
22. Thompson WH (1971) New route cuts costs for potassium orthophosphates. Chem Eng 78(8):83–85
23. Drechsel EK (1973) Potassium phosphates: the new generation of SUPER phosphates. Paper presented to the American Chemical Society. Division of Fertilizer and Soil Chemistry, Chicago, Illinois, August 28

24. Giese A (1998) Environmental responsibility on the farm—the impact on the fertilizer and agriculture industry. In: Environmental Challenges of Fertilizer Production—An Examination of Progress and Pitfalls. Proceedings of an International Workshop, JJ Schultz and EN Roth (Eds). IFDC, Muscle Shoals, AL, September 17–19
25. Yates LD, Nielson FT, Hicks GC (1954) TVA continuous ammoniator for superphosphates and fertilizer mixtures. Part I and Part II of Farm Chemicals (August 1954)
26. U. S. Patent 2,741,545
27. (1998) Environmental challenges of fertilizer production—an examination of progress and pitfalls. Proceedings of an International Workshop, JJ Schultz and EN Roth (Eds). Muscle Shoals, AL, September 17–19
28. (1972) New Developments in Fertilizer Technology, 9th Demonstration. TVA Bull. Y-50. Tennessee Valley Authority
29. (1976) New Developments in Fertilizer Technology, 11th Demonstration. TVA Bull. Y-107. Tennessee Valley Authority
30. Medbery J (1971) Experience with pipe-cross reactor. Proc. Fert. Ind. Round Table. pp 52–55
31. Baggett KJ, Brunner DJ (1977) MFA's experience with TVA cross-pipe reactor. Proc. Fert. Ind. Round Table. pp 64–70
32. Fittell RS, Hollingworth LA, Forney JG (1977) Manufacture of granular ammonium phosphates using a pipe reactor process. Proc. Fert. Ind. Round Table. pp 70–81
33. Hoffmeister G, Watkins SC, Silverberg J (1964) Bulk blending of fertilizer material: effect of size, shape, and density on segregation. J Agr Food Chem 12:64–69
34. Hoffmeister G (1962) Compatibility of raw materials in blended fertilizers—segregation of raw materials. Proc. 12th Annual Meeting Fert. Ind. Round Table. pp 83–88
35. Young RD (1969) Providing micronutrients in bulk-blended, granular, liquid and suspension fertilizer. Comm Fert 118(1):21–24
36. Young RD (1971) Production of compound fertilizers from intermediates in local plants. TVA Bull. Z-30. Tennessee Valley Authority, pp 12–20
37. McKnight D, Striplin MM (1958) Phosphoric acid of high concentration. J Agr Chem 13:33–34
38. Phillips AB (1963) Concentration wet-processes acid to superphosphoric acid. Farm Chem 126:36, 38, 62
39. Meline RS, Lee RG, Scott WC (1972) Use of pipe reactor in production of liquid fertilizers with very high polyphosphate content. Fert Soln 16(2):32–45
40. U. S. Patent 3,775,534
41. Walters HK Jr (1959) Salt suspension fertilizers. Comm Fert 25–26
42. Newsom WS Jr (1960) Suspension fertilizers. Solutions 30–35
43. Sawyer EW, Polon JA, Smith HA (1960) The stabilization of suspension fertilizers with colloidal attapulgite. Solutions 36–43
44. Mann C II, McGill KE, Jones TM (1982) I&EC Product Research and Development. pp 488–495
45. Trenkel ME (1997) Improving fertilizer use efficiency: controlled-release and stabilized fertilizers in agriculture. International Fertilizer Industry Association, Paris, France
46. Hamamoto M (1966) Isobutylidine diurea as a slow-acting nitrogen fertilizer and studies in this field in Japan. Proceedings No. 90. The Fertiliser Society, London
47. Clark KG, Yee JY, Love KS (1948) New synthetic nitrogen fertilizers, preparation and properties of urea-form. Ind Eng Chem 40(7):1178–1183
48. Kolterman DW, Rennie WW (1960) Ureaform fertilizers. In: Sauchelli V (ed) The chemistry and technology of fertilizers. Rheinhold Publishing Co, New York, NY, pp 48–54
49. Ando J (1970) Developments in granulation of mixed fertilizers in Japan. Proceedings of the 20th Annual Meeting Fertilizer Industry Round Table. Memphis, TN, pp 85–93
50. Kuwabara M, Hayamizu S, Hatakeyama A (1977) Trends in urea-based compound fertilizer technology. In: Granular fertilizers and their production. British Sulphur Corporation, London, England, pp 125–147
51. Nielsson FT (1973) IMC experience in specialty fertilizer manufacture. Proceedings of the 23rd Annual Meeting of the Fertilizer Industry Round Table. Washington, DC, pp 33–42
52. Powell R (1968) Controlled release fertilizers. Noyes Development Corporation, Park Ridge, NJ
53. Young RD (1974) TVA'S development of sulfur-coated urea. TVA Bull. Y-79. Tennessee Valley Authority, Muscle Shoals, AL
54. Rutland DW (1991) Fertilizer caking: mechanisms, influential factors, and methods of prevention. Fertilizer Research 30:99–114
55. Hofstee JS (1993) Physical properties of fertilizer in relation to handling and spreading. Wageningen Agricultural University, Wageningen, The Netherlands
56. International Organization for Standardization, Central Secretariat, 1, rue de Varembe, Case Postale 5b, ch-1211, Geneva 20, Switzerland
57. Bremner JM, Yeomans JC (1986) Effects of nitrification inhibitors on denitrification of nitrate in soil. Biol Fertil Soils 2(4):173–179

Sulfur and Sulfuric Acid

25

Gerard E. d'Aquin and Robert C. Fell

Sulfur

Sulfur is one of the few elements that is found in its elemental form in nature. Typical sulfur deposits occur in sedimentary limestone/gypsum formations, in limestone/anhydrite formations associated with salt domes or in volcanic rock [1]. A yellow solid at normal temperature, sulfur becomes progressively lighter in color at lower temperatures and is almost white at the temperature of liquid air. It melts at 114–119°C (depending on crystalline form) to a yellow liquid which turns orange as the temperature is increased. The low viscosity of the liquid begins to rise sharply above 160°C, peaking at 93 Pas at 188°C, and then falling as the temperature continues to rise to its boiling point of 445°C. This and other anomalous properties of the liquid state are due to equilibria between the various molecular species of sulfur which includes small chains and rings.

Sulfur is also found as sulfide minerals in combination with iron or base metals (e.g., pyrites) and as sulfates in combination with alkali metals (copper) and alkaline earths (e.g., gypsum). Hydrogen sulfide, with its "rotten egg" odor, and carbon dioxide are the principal "sour" components of sour gas. Crude oil and coal include a variety of complex organic species containing sulfur. Such compounds are removed from the liquid fuels by treatment with hydrogen to convert the sulfur component to hydrogen sulfide which is removed from the fuel stream. Removal of sulfur from fuels for technical and environmental reasons represents the largest current source of elemental sulfur.

World elemental sulfur production in 2009 was almost 48 million metric tons, up 3 million tons since 2003. Other sources (pyrites roasting in China and cleaning smelter gas-effluent globally) have added 25 million tons to the sulfur balance (Table 25.1).

Currently, almost all sulfur is derived from hydrocarbon processing. Minor quantities are still mined. Poland has a capacity of 700,000 t with a long history in Frasch mining. Mexico restarted one of its idled mines in 2010 (Jaltipan) with an output of about 100–200 t per day. Iraq's Mishraq mine is said to be trying to restart. Most sulfur mines are associated with oil deposits found in conjunction with geologic formations known as salt domes. In situ cross contamination with hydrocarbons (as at the Jaltipan and Mishraq mines) leads to the production of "dark" sulfur with technically and commercially unacceptable levels of carbon contamination. Volcanic sulfur deposits yield bright sulfur, but it is often contaminated with toxic metal oxides. Small quantities of several specialty sulfurs are produced for a variety of other applications, including fungicides, bleaching, fumigation, pharmaceuticals, rubber manufacture, and cutting oils. Applications of elemental sulfur as a fertilizer are growing globally. The Sulphur Institute (www.sulphurinstitute.org) recently launched a program entitled: *Sulfur—An Advantaged Element*. Additionally, an excellent publication by the USGS entitled "Materials Flow of Sulfur" is available at http://minerals.USGS.gov/minerals/pubs/commodity/sulfur/index.html.

Transportation and Storage

Although all sulfur is produced and most is consumed in a molten state, the preponderance of international commerce outside of North America and Europe uses solid sulfur. Vancouver was once the world's leading sulfur exporting port. Volumes peaked at 6.3 million tons in 2004, declining to 3.6 million tons in 2010 [2]. Qatar should surpass Vancouver no later than 2013 given expansion plans and continued decline in Canada's "commercially available" sulfur output.

G.E. d'Aquin
Con-Sul, Inc., Professional Consulting Services, Bigfork,
MT 59911, USA
e-mail: consulinc@aol.com

R.C. Fell
Monsanto Enviro-Chem Systems, St. Louis, MO, USA

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_25,

Table 25.1 Elemental sulfur production in all forms (million metric tonnes)

World sulfur production by type	1999	2009
Mined/Frasch	3.2	0.3
Recovered sulfur, of which:		
From oil	15.6	22.7
From gas	21.9	23.3
From other	1.1	2.4
Sub total	41.8	48.7
Other forms		
Pyrites		
Other		
Sub total	15.3	19.2
Total	57.1	67.9

Source: *Sulphur Magazine*, Jan/Feb 2011; US Geological Survey 2000, 2010

Prior to bulk transport, solid sulfur is normally "formed" into a variety of shapes associated with individual solidification processes. These will be discussed in the following section. "Formed" solid sulfur is easily transported by truck, railcar, barge, or ocean vessel. It is classified as a nonhazardous material "granular formed sulfur". Caution must be exercised in order to avoid fugitive dust and spillage. Sulfur dust is highly visible, will become corrosive, and under certain circumstances, explosive.

In the case of marine transport, which may last several weeks, the risk of corrosion during transit must be addressed.

Vessel holds must be coated prior to loading, to prevent direct sulfur-steel contact. A new coat of paint covering all steel is good. However, nicks and scratches quickly expose steel. Various products and means of coating holds are employed by knowledgeable companies prior to loading. This preparation avoids a range of potentially severe consequences and cannot be bypassed without significant risk to cargo and vessel. Spraying water with surfactants, or aqueous foam, on sulfur as it is being loaded or unloaded is another key safety measure which mitigates dust dispersion. Water used in dust suppression as sulfur is loaded will move to the bottom of holds during transit. Daily pumping of this residue to sea is necessary to avoid the accumulation of up to several feet of standing water in the bottom of the hold. Prior to loading and following discharge, formed sulfur is often stored in the open air.

In contrast, domestically produced sulfur destined for internal use in the USA, Canada, and Europe is almost exclusively transported in liquid form. This has led to the establishment of an extensive sulfur infrastructure consisting of molten sulfur terminals, tanker ships and barges, rail tank-cars, and tank-trucks. In the cases of marine transport, vessels range from 23,000 to 9,500 t and barges from 1,000 to 2,500 t. Overseas, molten shipments can be received in Morocco, Tunisia and one port in India. Japan prohibits solid sulfur, even in the case of export activities. This has led to a fleet of coastal vessels, some as small as 1,000 t, for exporting Japanese molten sulfur to users in China, Taiwan, and formerly, Korea.

Molten sulfur in ships is maintained in a liquid state using heating coils. When transported by rail tank-car, at least a portion of the sulfur cools and solidifies during transit. It is re-liquefied prior to discharge. The re-melting process can take up to 3 days (i.e., in the United States with sulfur coming from Canada during winter). Re-melt is accomplished by passing low pressure steam through coils located between the car's inner tank and outer skin. Care must be taken to use a low heating temperature to avoid turning sulfur into a viscous material. Trucks normally operate within distances which preclude cooling and solidification. Nevertheless, limited coiling allows for emergency steaming. In Tampa, FL, many sulfur truck trailers are designed for round trip carriage. An inner tank carries sulfur from the port and, depending on the design, phosphoric acid, solid fertilizers, or phosphate rock in an outer container on the return.

Certain countries have regulations that require molten sulfur be treated (de-gassed) at the production point to reduce the hydrogen sulfide level below 10 ppm weight. The City of Vancouver requires formed sulfur transshipped at the port to test below 30 ppm weight. Neither US nor Canadian governments have a national maximum hydrogen sulfide content standard for sulfur.

Storage of molten sulfur requires insulated tanks equipped with heating coils through which steam or glycol is constantly circulated. The tank's internal surfaces, particularly the roof, must be protected against corrosion and carefully monitored. Corrosion occurs most often when moisture enters through tank vents while it empties, condenses on the wall, and mixes with vapors of sulfur and its compounds. In recent years, CSI Controls has developed an enhanced system to heat sulfur tanks externally to prevent sulfur from freezing on the walls. It also evaporates all moisture from the walls and roof. Tanks, pumps, and lines must also be properly insulated and heated. A means of injecting steam into the tank void must be provided for fire control. Finally, depending on environmental regulations, tank vents may require scrubbing systems. In the case of marine terminals, these must be sized to accommodate large volumes of air as vessels discharge. "Formed" solid sulfur, on the other hand, can be stored in a pile exposed to the open air. Runoff water containment and neutralization is required around such sites. Some localities may have maximum H_2S content standards, protective walls, enclosed storage, windbreakers, and sprinkling systems to prevent blowing dust. Another means of storage, employed to minimize cost and maximize volume in any given area, is to "vat" or "block" the sulfur. In this case, a properly constructed impermeable "base pad" surface is prepared. This is then enclosed (surrounded by) with aluminum plates (forms) which are 50–75 cm high. A thin layer of molten sulfur is poured onto the pad and allowed to solidify. The process is repeated until the level of sulfur reaches the top of the aluminum forms wall. Another set of forms is then attached to the top edge of the solidified sulfur and the process is repeated. In this manner, sulfur "blocks" over 10 m high and several hundred meters long are created. Sulfur from blocks is reclaimed on-site, by melting, or by mechanical means followed by melting or shipment as the "crushed bulk" material. Despite its apparent simplicity, innumerable technical issues affect every aspect of the block construction, pouring, maintenance, and reclamation. Among these are insuring no liquid, dust, or gas effluent escapes from the site; as well as corrosion; overcoming the effects of sun, rain, and freezing; insuring proper solidification of the sulfur being poured; wall degradation; etc. Companies experienced in creating sulfur blocks and providing equipment are Devco USA (http://www.devcousa.com/), Enersul Inc. (http://www.enersul.com), and Hazco Environmental Services (http://www.hazco.com). Considerable information on these matters can be obtained from *Sulfur*, a BC Insight.com publication.

A key final consideration when storing solid sulfur is the inevitable presence of sulfuric acid. Sulfur can become naturally contaminated with sulfuric acid through the action of thiobacilli thiooxidans [3] or continuous exposure to direct sunlight [4] and the presence of moisture. Recent research has demonstrated the short-term effectiveness of certain bactericides in delaying bacterial colonization. Nevertheless, discrete pockets of weak (highly corrosive) sulfuric acid should always be presumed to exist within a sulfur storage pile. Hydrochloric acid, may also be present when solid sulfur has been transported by ocean vessel [5]. Ocean deliveries of solid sulfur should be neutralized to avoid potentially disastrous corrosion of downstream equipment.

Solidification and Melting

For many years, the standard industry practice was to ship solid sulfur as a "crushed bulk". The words adequately describe its appearance and dustiness. *Crushed bulk* is the least acceptable type of solid sulfur. It is created when using earthmoving equipment to recover sulfur which has been allowed to solidify in blocks, sheets, or vats. Obviously, the product is extremely dusty. Very few producers (Russia, Kazakhstan, and Iran) still ship crushed bulk. Most consumers refuse to accept shiploads even at significant price discounts due to dustiness, related contamination, and product loss. "Lump and Crushed" sulfur is classified as a hazardous material by the International Maritime Organization (IMO).

Slate, another type of formed sulfur, was the sulfur industry's first attempt at developing a type of solid sulfur which was less dusty than "crushed bulk". Slate is produced by pouring a thin (usually up to 3 cm) layer of molten sulfur onto a moving belt while simultaneously cooling the belt's underside. Solidified sulfur breaks off into irregular shapes when falling off the end of the belt, leading it to resemble pieces of "peanut brittle." Owing to its greater dustiness, slate sulfur is a less desirable form of formed sulfur and faces commercial resistance. Slate is classified as nonhazardous by the IMO.

Best Available Current Technologies (BACT) for solidifying sulfur creates a pelletized, prilled, or granular product. These types of "formed" sulfur were developed in the late 1970s through today. They minimize breakage and sharp edges which create dust during handling. As the product names imply: *pelletized* sulfur is a pellet generated when a discrete amount of sulfur is deposited and cooled on a conveyor belt. It has the shape of an egg yolk on a plate. The most common processes are offered by "Sandvik" [6] and "Berndorf" [7] but other licensors exist. *Prilled* sulfur is created when sulfur droplets are cooled by a surrounding fluid, which may be air or water. *Air-prills* are created by spraying sulfur from the top of a tower into a strong updraft of cool air. It is an extremely smooth product which resembles prilled urea or ammonium nitrate. No air-prilling facilities have been built in almost three decades due the process' inevitable emission of sulfur dust to the environment. High volume air scrubbing systems to eliminate dust pollution have not been successful. In addition, the last two installations at Jubail, Saudi Arabia, were destroyed by a fire believed to have been ignited by static electricity. *Wet-prills* are created when sulfur droplets enter a cool water bath. They have a slightly irregular texture caused by water's more rapid cooling action. Devco [8] USA, Martin Sulphur [9] and Enersul [10] have fairly similar technologies and products. Current processes create a homogeneous product which resembles small nuggets. At this writing, all sulfur-forming operations in the United States produce wet-prill. *Granular* sulfur, as the name implies, is created through the gradual buildup of successive layers of sulfur around a central core. As a result, the ultimate product size can be controlled. Enersul [10] possesses the most widespread granulation system, with installations in Canada, the Former Soviet Union, Europe, and the Middle East. Kaltenbach-Thuring [11] provides another type of granulation technology. Most recently (4/11), Martin Sulphur [9] has announced a proprietary (patent pending) granulation process.

Despite developments in forming technology and dust suppressants, moisture remains the primary tool for dust mitigation. To that end, a mixture of water and surfactant should be applied during all bulk transfer operations. A US Patent was recently issued to Con-Sul, Inc.'s Gerard d'Aquin for a revolutionary solid sulfur ship unloading system which

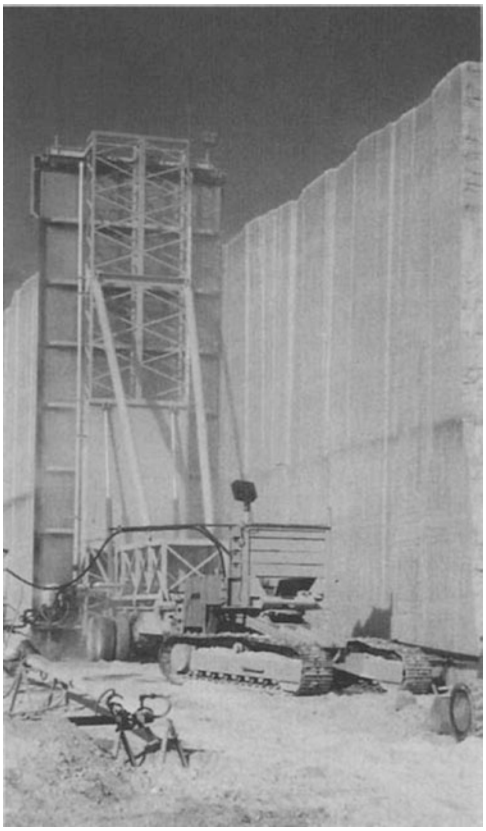


Fig. 25.1 An ellethorpe melter. (*Courtesy* Con-Sul, Inc.)

eliminates dust, controls acidity, and de-waters most types of formed sulfur [12].

Melting block inventory is most often accomplished using a track-mounted Ellethorpe melter (Fig. 25.1). This resembles a rectangular ironing-board mounted on tracks. The "hot" portion is placed vertically against the blocks side. Sulfur melted by the rectangular surface is collected at the base and piped to filtration (if appropriate) and storage. Mechanical reclamation, normally from inside the block to minimize dust emissions, crushes sulfur and transfers the lumps to a static pit or above ground tank melter.

Problems can surface. Melting or re-melting solid sulfur that has been transported by vessel requires proper knowledge and equipment. Contaminants, including sand, dirt, rocks, and rusted metal, which can mix with solid sulfur during transit and storage must be removed. Sulfuric acid is another issue: weak acid is almost always present in solid sulfur that has been stored for some time, let alone stored, transported by vessel, and then stored again. While customers can require a specification "below 100 ppm free acid" at load-port, they may receive several times that level when measured at the destination melter. The currently accepted practice is to mix lime or a liquid oxidizer with the material being melted to neutralize all acids. Liming, in turn, requires a filtration stage—large installations use steam jacketed pressure leaf filters, containing a diatomaceous earth medium over a stainless steel mesh, to remove the lime/ash residue. An alternative is to allow ash to settle prior to transfer from an intermediate pit.

Carbon compound molecules contained within sulfur will be converted to an extremely hard (Car-Sul) material when solid sulfur is re-liquefied. Car-Sul, which can be filtered, abrades sulfur burner spray nozzles [13]. Users of sulfur in the sulfuric acid industry therefore prefer receiving sulfur which has never been solidified rather than incurring the foregoing operating costs and losses of sulfur in filtrate residue. Owing to the volume of production, legislation, and a developed liquid infrastructure, most sulfur consumed in North America, Europe, Japan, and Korea is never solidified.

Development of the Sulfur Industry

Early mankind doubtless found elemental sulfur in volcanic craters, encrusting the edges of hot sulfur springs, and embedded in limestone formations. They discovered that it would burn and used it for medicinal purposes, as a bleach, fumigant, colorant, and incense. These uses are mentioned in ancient writings. The Romans produced incendiary weapons from sulfur. In the thirteenth century, the Chinese invented gunpowder using sulfur, nitrate, and charcoal.

The earliest commercial sulfur came from limestone deposits. Those in Sicily and the Italian mainland developed world markets in the eighteenth and nineteenth centuries. Traditional mining methods were used to produce sulfur ore, which was burned slowly in a pile (Calcarone) to yield crude sulfur [14]. Steam smelting in autoclaves came into use about 1859. In 1890, Robert Gill built a multi-chamber furnace to improve the process's production rate and efficiency. Italian monopoly of the sulfur markets continued until the early 1900s when the Frasch process brought previously unrecoverable sulfur deposits on the North American Gulf Coast into production. Oil exploration efforts in Texas and Louisiana in the late 1800s uncovered sulfur deposits in limestone at depths of 200–300 m, but mining was complicated by the intervening layers of quicksand and the presence of hydrogen sulfide gas. Numerous conventional mining attempts at Sulphur, LA, failed.

Finally, in December 1894, Hermann Frasch demonstrated the hot water process for mining underground sulfur deposits. With its favorable economics, the Frasch process completely displaced the Italian sulfur industry. The ready availability of low-cost sulfur opened the way for commercial sulfuric acid production by burning sulfur. This eventually supplanted the long-standing iron pyrite combustion process for sulfuric acid production by eliminating its extensive gas-cleaning operations.

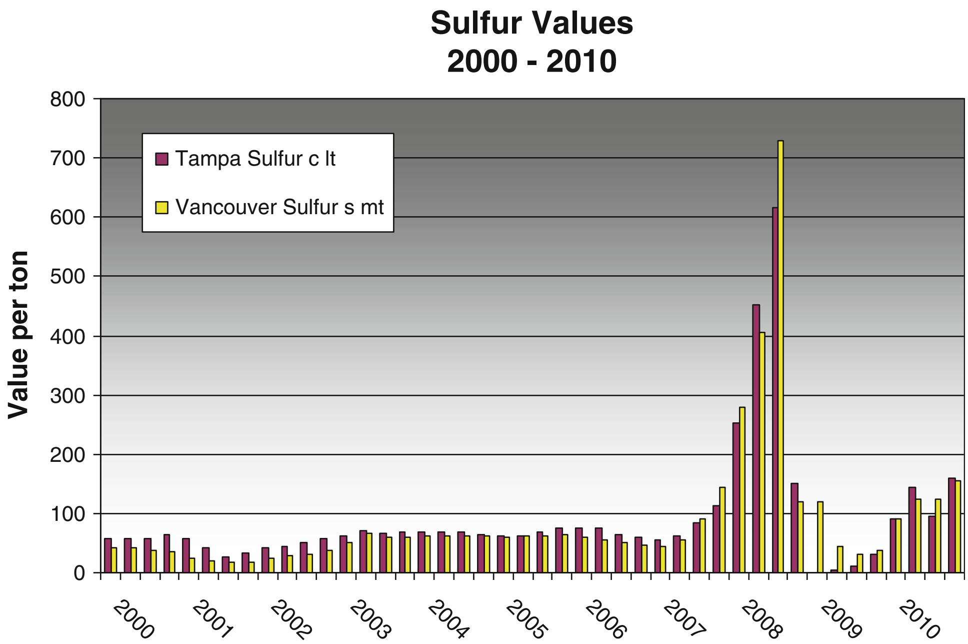


Fig. 25.2 Sulfur values 2000–2010. (*Courtesy* Con-Sul, Inc.) *Note*: 1Q 2009 Tampa price \$0 c = contract; s = spot

In 1883, the Claus process for producing sulfur from hydrogen sulfide through partial combustion over an iron oxide catalyst was patented. It had limited success over the following 50 years as a method for producing sulfur and its primary use was to remove the sour component in gas processing.

The number of Claus installations grew during the second half of the twentieth century. An increase in sour gas and liquid petroleum feedstocks, processing and, more recently, environmental regulations led to greater quantities of sulfur being removed from hydrocarbon fuels. Sulfur extracted from oil and gas has increased dramatically since the 1970s when environmental legislation took hold. In 1985, Frasch and recovered sulfur in the United States each accounted for about five million tons of production. By 1995, US Frasch output had declined to 3.1 million tons and recovered represented 7.3 million tons. In 2001, faced with oil companies disposal of recovered sulfur at low prices and rising natural gas costs to heat mine-water, US Frasch mining ceased [15]. Frasch sulfur technology, with its colorful and inventive history, lasted just over 100 years in the United States. For the world, removing sulfur from hydrocarbons provided virtually all of the 48 million tons of elemental sulfur generated during 2010 [16].

The cessation of US Frasch output at the end of the 1990s brought the demise of associated infrastructure and dissipated the industries commercial focus. Since 2001, sulfur producers have been unable to effectively match supply with consumer demand through inventory management.

The consequence has been a series of value fluctuations (most notably declines in 2001, 2004, and 2008–2009) due to declines in phosphate fertilizer output which led to a collapse of sulfur prices [17] (see Fig. 25.2).

Contrary to normal economic reasoning, fluctuations in price are not due to large imbalances between demand and commercially available supply. Rather, they are due to small imbalances, compared to the 48 million tons of annual output. Even small fluctuations cannot be accommodated in the industry's limited commercial storage facilities. Refineries and sour-gas processing plants must dispose of their sulfur output. Volumes in excess of the limited storage capacity available normally forces refineries to cease operations due to environmental regulations. Therefore, any reduction in demand and a smooth outflow to consumers, will lead to aggressive price discounting.

While Frasch sulfur companies maintained a working storage capacity of approximately seven million tons of solid sulfur at several locations in the early 1980s, total US solid block storage is currently below 600,000 t. Such storage also helped reduce upward movements in price when weather or accidents restricted deliveries. As sulfur went through periods of low value, producers in more distant locations (Canada, Russia, Saudi Arabia, Kazakhstan) have resorted to blocking the material (see section "Solidification and Melting"). This occurred during the late 1960s along the Alberta Foothills area (East of the Rocky Mountains from the US border to well North of Edmonton, Canada).

In the late 1970s the area contained upward of 22 million tons of elemental sulfur. With US stockpiles, it served as the world's reserve. Over the past three decades Alberta's stockpiles have been as low as 4 million tons and as high as 16 million tons. Today, the figure stands at about 12

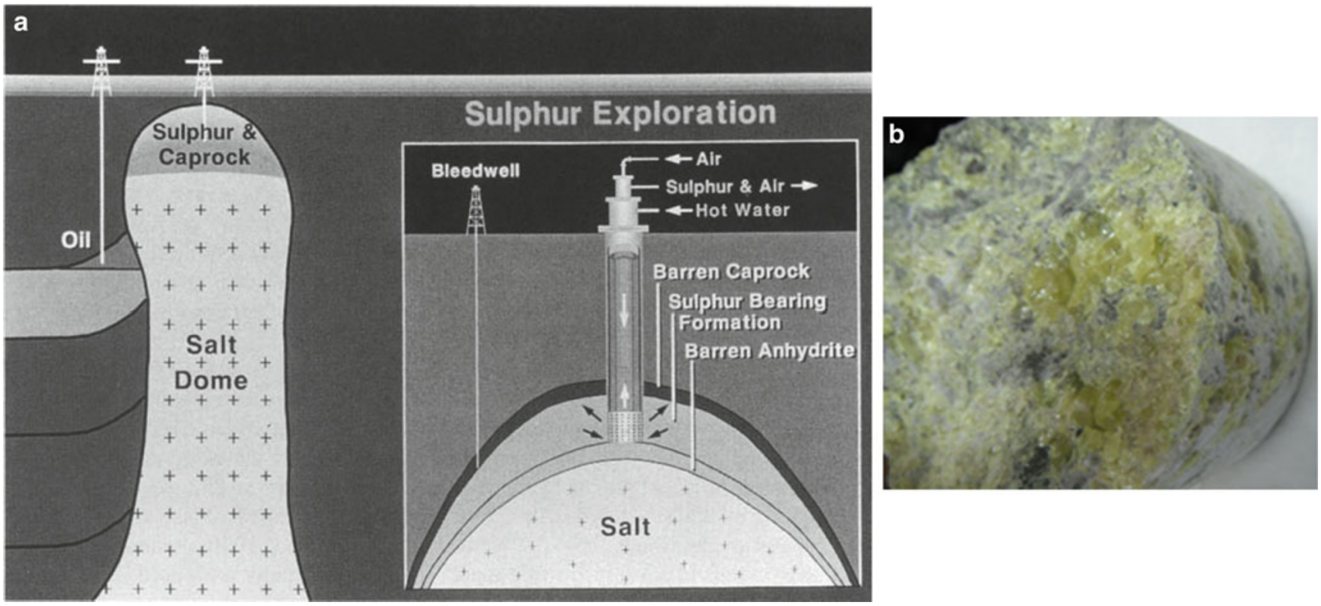


Fig. 25.3 (**a**) Frasch process for mining sulfur from salt-dome formations. (*Courtesy* Freeport Sulphur Co.). (**b**) Sulfur-bearing limestone

million tons; 3 million in the Foothills and 9+ million in the Oil Sands producing area. Sulfur produced and stored in the Oil Sands, as differentiated with the limited production and storage in the Foothills, has only a limited "commercial availability" due to constraints in transportation infrastructure and costs. Changing these conditions requires capital investment and governmental approval. Neither one is likely given the volatile nature of sulfur demand and prices. Therefore, Syncrude's nine million ton block of sulfur and parts of the 900,000 t produced per year are commercially unavailable. In stark contrast, had production and inventory been located in the Foothills, and had reclamation and transportation existed, it would have been shipped to markets in 2009, 2010, and 2011.

Other areas where a significant volume of sulfur inventory accumulation has taken place when prices do not justify the cost of commercial disposal are Saudi Arabia (1970s to early 1980s and 2000–2009), Russia (depleted in 2011) and Kazakhstan (still with 6 million tons of inventory).

Sulfur industry practice when calculating an outlook of annual supply/demand balances uses total production less demand with no differentiation for "commercially available" or "commercially unavailable" supply sulfur. Similarly, no distinction is made for "commercially available sulfur" inventory at Alberta Foothills locations and non-commercially available at Syncrude's Oil Sands location.

One means of disposing of unwanted sulfur streams is re-injection of H_2S and CO_2. The process has long been used in Canada and the USA on a small scale. The largest project is currently at ExxonMobil's La Barge, Wyoming gas plant where up to 400,000 t per year of elemental sulfur has been removed from the market since 2005. Huge projects to re-inject sulfur at the Kashagan field in Kazakhstan and at various Arab Gulf locations are underway or contemplated.

Sulfur is also produced from sulfide ores (pyrites) by thermal decomposition in the absence of air, by roasting/smelting under reducing conditions, or by reaction of the ore with SO_2. Hydrometallurgical processes have produced sulfur from metal pyrites as a by-product.

Except for China, pyrites roasting no longer accounts for significant quantities of sulfur values. And even China has embarked on a rapid conversion to using elemental sulfur in former pyrites roasters in order to lower production costs and pollution. Canadian sulfur exporters, who initiated this trend with the development and introduction of innovative technology, have gained the most. The Canadian exports of sulfur to China rose from 31,000 t in 1995 to a peak of 3.9 million tons in 2005 [18]. Since then the volume declined to 1.5 million tons in 2010.

Sulfur Production Processes

Despite the economically driven closure of the Main Pass 299 sulfur mine in 2000 (just 8 years after initial output), the Frasch process remains the most effective method for extracting sulfur from native deposits. Certain constraints on the geological formations required by the Frasch process limit its use to deposits along the Gulf of Mexico, Poland, the Former Soviet Union, Mexico, and Iraq. Other sulfur deposits may yield to the Frasch process. Figure 25.3 shows the structure needed for Frasch mining from salt domes.

The sulfur-bearing limestone must have sufficient porosity to allow sulfur to migrate upon melting. Both the caprock and the underlying anhydrite formation should be impervious to prevent the loss of the hot water pumped into the mine. These salt-dome deposits are typically lens shaped and are 175 m thick with diameters of a few hundred meters up to several kilometers. Some horizontal deposits exist and have been exploited.

A sulfur well consists of a casing and three concentric pipes reaching into the sulfur-bearing strata. The outer 8–10 in. pipe carries 165°C water pumped into the formation to melt the sulfur. An inner cement-lined 3–6 in. pipe is used to transport the melted sulfur to the surface. Compressed air is passed through the l inch tube in the center to air lift the sulfur. Without the air lift, the molten sulfur would rise only part way in the middle pipe. The compressed air produces a low-density sulfur froth that rises to the surface. The superheated water melts the sulfur in the vicinity of the well, forming a molten sulfur pool at the bottom of the well. As production continues, the formation fills with water. To continue production, bleed wells are drilled at the periphery of the formation to allow for discharge of the cooled mine water. In some mine fields, sufficient mine water is lost to the geological formation to provide for continued production. To limit mine water loss and reduce water heating costs, mud, or synthetic foam is sometimes pumped into the formation to seal major crevices.

Although most US Frasch mines were located inland, Freeport Sulfur Company pioneered offshore sulfur mining in the 1960s [19].· The company utilized offshore oil drilling techniques to access several shallow water deposits. Developing a means to utilize saltwater to provide heated mine water presented the company with unique challenges in corrosion and scaling control. The now defunct Main Pass 299 facility located in 200 ft of water 10 miles from the Mississippi Delta, commenced production in 1992 and closed in 2000.

Sulfur produced from salt-dome structures can be quite pure, but it often contains up to 1% of bituminous residues, which render it dark and can make it unacceptable for sulfuric acid production. Some purification is obtained by filtering the dark sulfur through diatomaceous earth. Nevertheless, the most effective means of meeting maximum commercial specifications for carbon is through blending dark sulfur with bright recovered sulfur devoid of carbon. That practice was widely used by Louisiana Frasch producers. To improve filtration, carbonaceous material is converted into larger, filterable, particles (Car-Sul) by treating the sulfur with heat or sulfuric acid. Freeport's submerged combustion distillation process was used from 1966 until 1979 to purify Frasch sulfur with up to 2% carbonaceous material. It was also going to be used to purify sulfur extracted at Mishraq, Iraq in 1990.

Recovered Sulfur

Hydrogen sulfide is recovered from natural gas or refinery gases by absorption in a solvent or by regenerative chemical absorption [20].· In either case a concentrated hydrogen sulfide stream is produced that is treated further by the Claus process. A typical Claus plant has a feed stream of at least 45% H_2S, but with modifications can handle streams containing as little as 5% H_2S. For gas streams with low concentrations of hydrogen sulfide, direct conversion of the hydrogen sulfide to sulfur is accomplished in a solvent system, as in the Stretford or CrystaSulf processes.

The Claus process is based on the reaction of H_2S with sulfur dioxide according to the highly exothermic reaction:

$$2H_2S + SO_2 \rightarrow 3S + 2H_2O$$

In practice, sulfur dioxide is produced in situ by partial oxidation of the hydrogen sulfide with air or oxygen in a furnace. In the split-flow arrangement, one-third of the H_2S stream is burned and then recombined with the remaining two-thirds before entering the Claus reactor. In the straight-through version, the entire H_2S stream is sent through the burner and the extent of H_2S combustion is controlled by the air feed rate.

A flow diagram for a typical Claus process is shown in Fig. 25.4. The hydrogen sulfide is burned in a fuel-fired furnace (950–1,250°C) with air to produce sulfur and a gas stream containing H_2S and SO_2. Process controls maintain the H_2S:SO_2 ratio near 2, in accordance with the stoichiometry. Heat is removed from the gas stream in a waste heat boiler to control the process gas temperature. The process gas is passed through one or more catalyst beds to convert the H_2S and SO_2 to sulfur, which is then removed in condensers between each bed. A high temperature shifts the equilibria toward the reactants, whereas a low temperature causes sulfur condensation on the catalyst bed, leading to decreased catalyst activity, requiring bed switching/regeneration to achieve modem recovery standards. In practice the temperature is closely controlled for each Claus reactor, with higher temperatures at the first reactor where compounds such as COS and CS_2 are converted to S. Much research and development into H_2S conversation has occurred in the past decades. Three-state Claus units are now capable of achieving recoveries of 98+ percent, a significant improvement over the 85–90% range of the late 1980s [21].·

Figure 25.4 shows a typical sulfur recovery plant based on the Claus process, and Fig. 25.5 provides a photographic overview of an installation. The tail gas from the Claus reactors may be further processed to remove any remaining sulfur compounds. Combined H_2S removal efficiencies of 99.5–99.99% are achievable [20]. This may be done by low-temperature Claus-type solid-bed processes (e.g., the Sulfreen process), wet-Claus absorption/oxidation processes

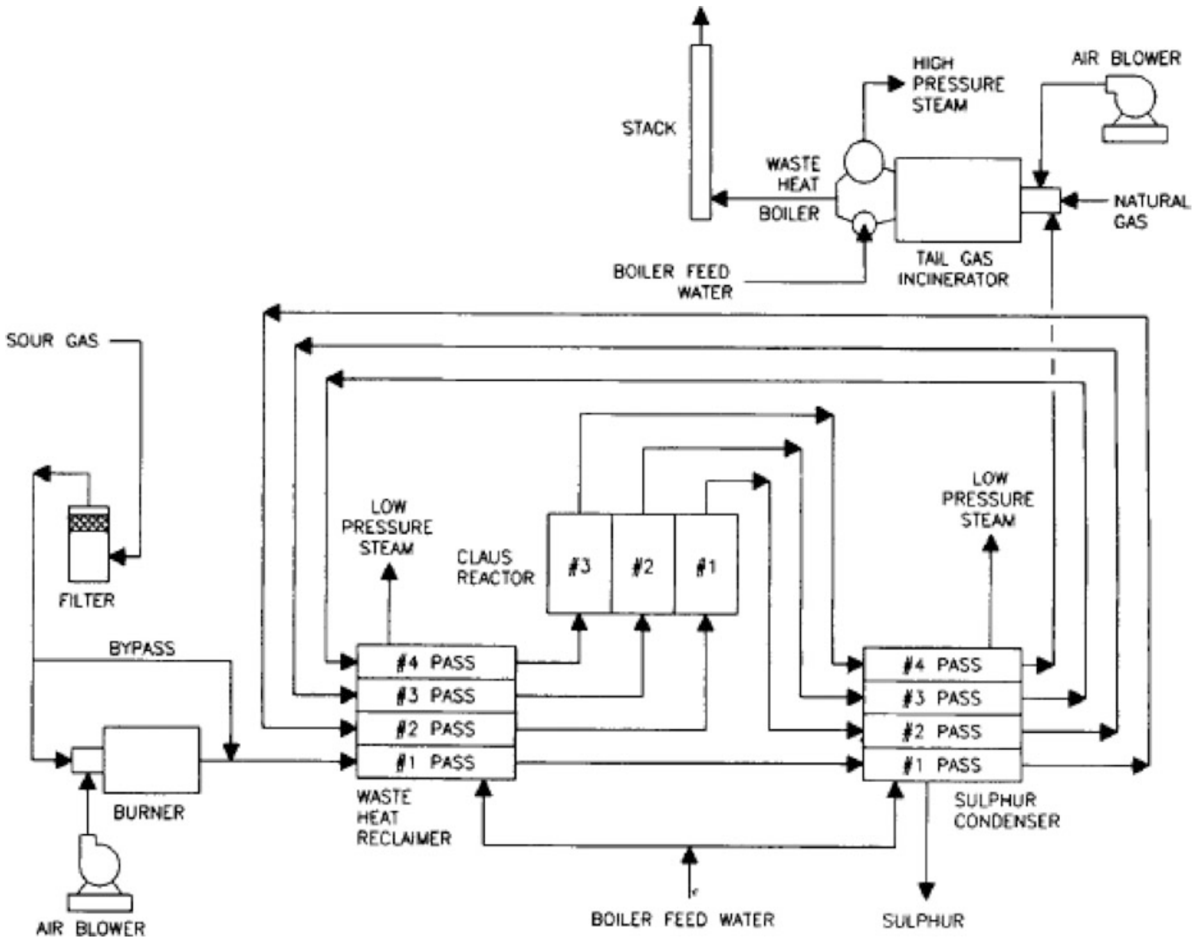


Fig. 25.4 Flow diagram for a typical Claus process

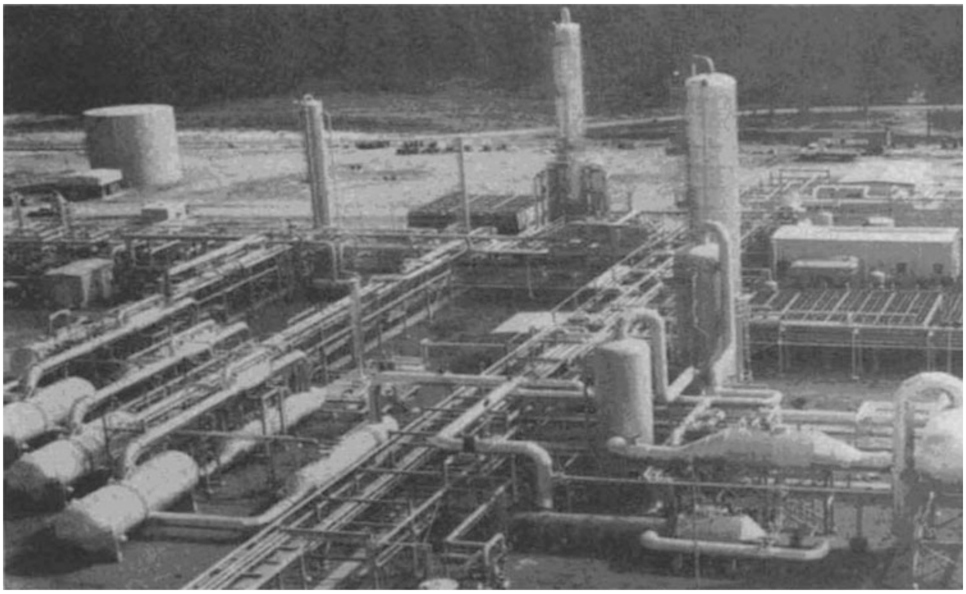


Fig. 25.5 Republic Refining Co. 440 metric ton per day Claus sulfur recovery unit (*left*) and SCOT tailgas cleanup unit (*right*). (*Courtesy* Ortloff Engineers, Ltd.)

Table 25.2 Estimated US sulfur supply/demand (million metric tonnes)

Supply	‘00	‘05	‘06	‘07	‘08	‘09	‘10
Production							
Frasch	0.8						
Recovered: oil	6.4	6.9	7.0	7.0	6.9	7.0	7.1
Gas	2.3	1.8	1.4	1.3	1.1	1.2	1.1
Subtotal	8.7						
Total	9.5	8.7	8.4	8.3	8.0	8.2	8.2
Imports to US							
Canada	1.51	2.01	2.03	2.10	2.18	1.50	2.30
Venezuela	0.16	0.37	0.41	0.37	0.36	0.20	0.02
Mexico	0.51	0.47	0.49	0.44	0.37	0.53	0.42
Other	0.03	0.00	0.00	0.02	0.10	0.00	0.40
Total	2.21	2.85	2.93	2.93	3.01	2.23	3.14
Exports	−0.8	−0.9	−1.10	−1.20	−1.14	−2.10	−1.40
Total supply	10.91	10.65	10.23	10.03	9.87	8.33	9.94
Consumption	11.19	10.75	10.12	10.30	10.00	8.20	10.22
Inventory change	−0.28	−0.10	0.11	−0.27	−0.13	0.13	−0.28

Source: North American Quarterly Review, April 1, 2011, Con-Sul, Inc.

(e.g., the Clauspol 1500 process), or hydrogenation of the off-gas to form H_2S for recycle (e.g., the SCOT process). Residual sulfur compounds in the tail gas are then incinerated to SO_2. The residual SO_2 in the oxidized tail gas may be scrubbed by any of several processes (e.g., the Wellman-Lord process) before being vented to the environment. It is feasible to bring the H_2S content of the treated tail gas to as little as 150–10 ppm weight, depending on the solvent used [22].

Production and Consumption of Sulfur

World production of sulfur in all forms reached 68 million tons in 2009 [23]—up from 57.1 million tons in 1999 [24]. The lower growth rate than anticipated at the end of the 1990s and early 2000s arose principally from delays in implementing Caspian and Arabian Gulf sour oil and gas projects. Schedules slipped due to technical difficulties (Caspian—Kashagan field, Government issues at Tenghiz) and cost escalation/over-ambition and economic setbacks (Arab Gulf). Non-elemental sulfur output has seen a significant change in composition. Volumes arising from environmentally mandated stack-gas effluent recovery have increased. Within the mix, pyrites-derived production continues to decline, though only slightly now that China is the lone remaining producer of substance.

Table 25.1 summarizes the dramatic shifts in the sources of sulfur during this period. It shows the demise of Frasch and native sources of supply and a 9 million ton increase in recovered volume in 10 years. Added sulfuric acid and SO_2 recoveries from effluent gas added four million tons.

US output from 2000 to 2010, see Table 25.2, changed dramatically. Elemental sulfur production declined 1.3 million tonnes. Graph 25.1 provides a historical perspective of changes from 1980 to 2000.

Canada’s sulfur production is concentrated in Alberta. Table 25.3 provides output data for the three western most provinces. Over the past decade, sulfur generated in the region collapsed from 8.5 million tons in 1999 to 5.8 million tons in 2010—a drop (2.7) million, or (31%). The primary causes were:

Exhaustion of sour gas deposits, some under exploitation for more than 30 years.

Failure of new sour gas deposits to yield the needed replacement output—the Caroline gas plant area being just one.

Appreciation of the Canadian dollar having the impact of lowering Canadian Gas producers’ revenues on gas sent to the United States.

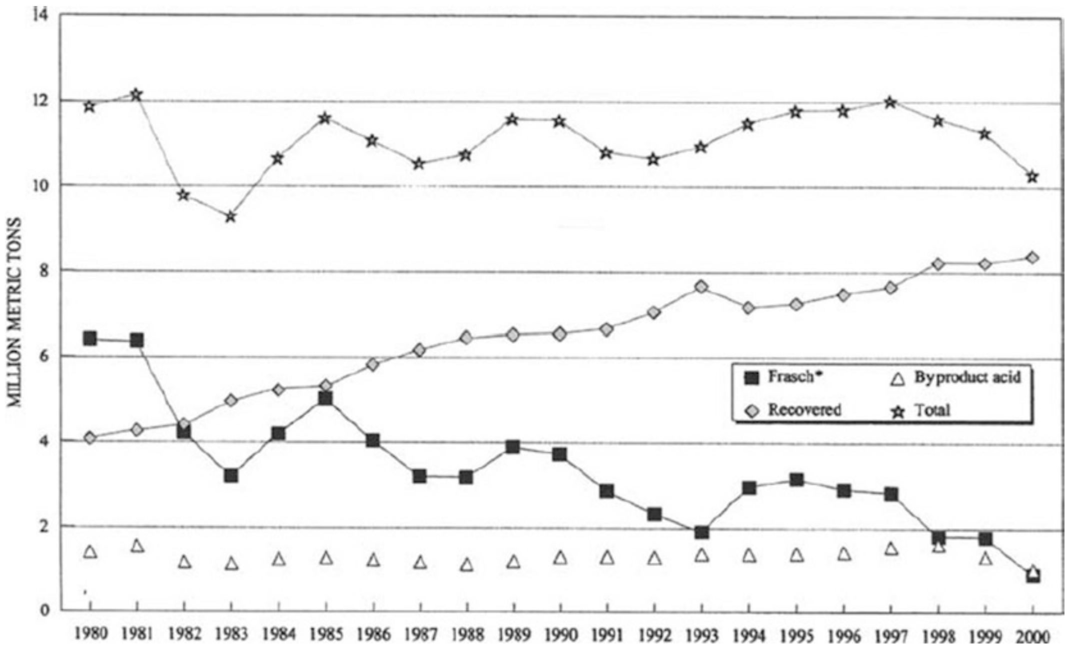
A climate of falling US gas price leading to sour gas exploration and exploitation ceasing.

Delays in the development of oil sands upgrading facilities which were expected to provide replacement volumes of sulfur. Wild cost inflation in the mid-2000s, inadequate physical, human, and regulatory resources, the financial bubble of 2007 and its subsequent collapse, all contributed.

From 1999 to 2003, and in 2009, production exceeded export shipments to the USA and Overseas. As a result, inventories in the oil sands region of Alberta grew significantly. From 2004 forward production was below export demand, leading to withdrawal of sulfur from inventory located in Alberta’s foothills.

Regional shipments and local use fell during the time frame from 7.3 million tons to 6.1 million tons—a 1.2 million tons annual decline.

Canadian exports by rail, in liquid form, to the United States and overseas, as a solidified material, through the port of Vancouver also declined.



Graph 25.1 Trends in the production of all forms of sulfur in the United States (Adapted from *Ober J,* Op Cit.)

Table 25.3 Western Canada sulphur summary[a] (million metric tons—MMT)

	'10	'09	'08	'07	'06	'05	'04	'03	'02	'01	'00	'99
Production	5.8	6.3	6.6	7.3	7.6	7.6	8.0	7.8	7.9	7.9	8.4	8.5
Shipments	6.1	6.0	7.1	7.5	8.0	8.3	8.5	7.5	7.3	6.5	7.0	7.3

[a] *Note*: Alberta, British Columbia, and Saskatchewan

Loss of the US Frasch companies' commercial structure affected North American markets dramatically. Recovered producers, devoid of storage capabilities and marketing experience, and needing to dispose of sulfur to avoid affecting refinery operating rates just pushed it out the door. Sulfur prices embarked on a decline which culminated with spot sulfur prices falling to the $10s per ton in Tampa, FL, and Vancouver, Canada, by mid-2001. In contrast, prices during 1990 were $140/t and $108/, respectively. A second collapse occurred in 2008–2009, leading to a Tampa price of zero during the first quarter of 2009 (see Fig. 25.2).

The United States remains the largest producer and consumer of sulfur in all forms. US production and consumption data are provided in Table 25.2. Fertilizers and agrochemicals represent the largest use for sulfur, 69%. Phosphatic fertilizers consume 7.1 million metric tons, 57%, in the form of sulfuric acid. Agro-chemicals account for 1.3 million metric tons, 10%, all in the form of elemental sulfur. Sulfur, surprisingly, represents the largest fungicide and pesticide product applied in the United States. Refining use, for alkylation, is the second largest segment of consumption. It is also one of the most difficult segments to track accurately. This arises from refineries not reporting a portion of their production and using that material directly for internal use. Accordingly, a portion of actual sulfur output does not get reported within the scope of "production of sulfur," causing total uses to exceed production by 1.1 million metric tons.

US production of elemental sulfur has fluctuated dramatically in recent years due to the cessation of Frasch mining operations. Output was 9.5 million tons in 2000, 8.7 million tons in 2005, only 8 million tons in 2008 and is estimated to have rebounded to 8.2 million tons in 2010.

Sulfuric Acid

Sulfuric acid is the largest-volume chemical manufactured in the world and its consumption is often cited as an indicator of the general state of the nation's economy [25]. According to the US Geological Survey's Mineral Industry Survey, about 27 million tons of sulfuric acid were produced in the United States in 2008, of which approximately 74% was used in fertilizer production. Its use extends to nearly every major chemical sector. This versatile acid is truly the "workhorse" of the chemical industry (Table 25.5).

Pure sulfuric acid is an oily, water-white, slightly viscous liquid with a melting point of 104°C and a boiling point of 279°C. It is infinitely miscible with water, forming sulfuric acid solutions characterized by their weight percent H_2SO_4. Oleum may be formed by dissolving SO_3 in sulfuric acid to attain fuming sulfuric acid, with concentrations nominally greater than 100% H_2SO_4. Historically, sulfuric acid concentrations were determined by measuring the solution density using hydrometers calibrated in degrees Baume (Be). This practice is waning, although some specifications and tables of properties still include this measurement (Table 25.4).

Uses of Sulfuric Acid

The primary industrial uses of sulfuric acid are in phosphate fertilizer manufacture, petroleum refining, copper ore leaching, synthetic rubber and plastics, and pulp and paper mills [26]. It is used as a solvent, a dehydrating agent, a reagent in chemical reactions or processes, an acid, a catalyst, and an absorbent, and in many other applications. In spite of its wide usage, sulfuric acid rarely is contained in the final product. Sulfuric acid ends up as gypsum in phosphate fertilizer manufacture, for example. In many other processes the sulfuric acid is converted to a waste product that requires disposal or reuse. Because disposal of waste sulfuric acid and of sulfur compound air emissions are becoming increasingly unacceptable environmentally, the recycle and capture of sulfur values from waste sulfuric acid has become widespread.

Table 25.4 Estimated annual export summary—total Canadian exports (million metric tons)

	‘11	‘10	‘09	‘08	‘07	‘06	‘05	‘04	‘03	‘02	‘01	‘00
Vancouver	3.4	3.6	4.1	4.7	5.2	5.8	6.1	6.3	5.3	5.1	5.2	5.3
US	2.6	2.3	1.5	2.2	2.1	2.0	2.0	2.0	2.0	1.9	1.1	1.5
Total	6.0	5.9	5.6	6.9	7.3	7.8	8.1	8.3	7.3	7.0	6.3	6.8

Nearly all sulfuric acid is manufactured by the contact process in which sulfur trioxide is absorbed into 93–98% H_2SO_4. The acid may be sold at various strengths, usually depending on the requirements of the consumer. It is generally marketed on a 100% basis, but normally is shipped as 93% H_2SO_4 (66°Be), as 98% acid, or as 20–22% fuming oleum. Table 25.6 shows common acid strengths and end uses. Concentrated acid may be stored in mild steel tanks, but dilute acid must be contained in lead-lined or plastic tanks. Bulk shipments of concentrated acid are made in steel tanks on ships, tank barges, or railcars. Reagent grade acid is commonly sold in 5-L glass bottles.

Development of the Sulfuric Acid Industry

Sulfuric acid is formed in nature by the oxidation and chemical decomposition of naturally occurring sulfur and sulfur-containing compounds. It is made by the action of bacteria (*Thiobacillus ferrooxidans* and *thiooxidans*) on coal wastes or iron disulfide discarded on refuse dumps at coal and copper mines; it is produced in the atmosphere by the oxidation of sulfur dioxide emitted from the combustion of coal, oil, and other substances; and it also is formed by chemical decomposition resulting from geological changes.

Although there were vague references to “spirits” expelled from alum by Arabian alchemists in the tenth century and by Roman alchemists in the thirteenth century, the first distinct mention of sulfuric acid has been credited to Basil Valentine in late 1400 [27]. He burned sulfur with

Table 25.5 US sulfur and sulfuric acid sold or used (thousand metric tons of sulfur content)

End use	Elemental S 2008	Sulfuric acid 2008	Total 2008	Form
Copper ores	49	456	456	H_2SO_4
Other ores		124	124	H_2SO_4
Pulp mills and paper products		187	187	H_2SO_4
Inorganic and other chemical products		331	380	S/H_2SO_4
Synthetic rubber and other plastic materials		69	69	S/H_2SO_4
Cellulosic fibers including rayon		31	31	S/H_2SO_4
Industrial organic chemicals, etc.		180	180	S/H_2SO_4
Nitrogenous fertilizers		58	58	
Phosphatic fertilizers		5,690	5,690	
Pesticides		15	15	S
Agricultural chemicals	2,050	28	2,080	S
Other chemical products		92	92	S
Petroleum refining and other petroleum and coal products	2,990	244	3,230	S
Metals leaching		12	12	H_2SO_4
Storage batteries (acid)		24	24	H_2SO_4
Exported sulfuric acid		7	7	H_2SO_4
Total identified	5,090	7,540	12,700	
Unidentified	684	132	816	
Grand total	5,770	7,680	13,500	S/H_2SO_4
As sulfuric acid		27,489	46,002	H_2SO_4

Source: USGS minerals yearbook, sulfur, 2008

Table 25.6 Acid strengths and end uses

Percent H_2SO_4	°Be	Percent Oleum (% Free SO_3)	Specific Gravity	Uses
35.67	30.8		1.2700	Storage batteries, electric utilities
62.18–9.65	50–55			Normal superphosphate and other fertilizers
77.67	60.0		1.7059	Normal superphosphate and other fertilizers; isopropyl and *sec-butyl* alcohols
80.00	61.3		1.7323	Copper leaching
93.19	66.0		1.8354	Phosphoric acid, titanium dioxide, steel pickling, regenerating ion exchange resins
98–99				Chlorine drying, alkylation, boric acid
104.50		20	1.9056	Surfactants, nitrations
106.75		30	1.9412	Hydrofluoric acid
109.00		40	1.9737	Explosives
111.24		50	1.9900	Reagent manufacture, organic
113.50		60	1.9919	Sulfonations, blending with
114.63		65	1.9842	Weaker acids

Source: Chemical economics handbook, SRI International, Dec 1990

saltpeter in glass retorts or bell jars with a little water, and he also calcined copper as (ferrous sulfate heptahydrate) with silica, with both processes yielding sulfuric acid, although he took them to be different substances. These processes for making sulfuric acid continued until 1746 when Dr Roebuck constructed a lead chamber in England for sulfuric acid manufacture. This marked the beginning of the "chamber process" for sulfuric acid, which was to continue in use for the next two centuries [28].

The first lead chamber was 1.8 m^2, and 8:1 mixtures of sulfur and saltpeter in iron carts were rolled into it and burned with intermittent admission of air [29]. As in glass retorts, the sulfur trioxide that was formed combined with water to produce sulfuric acid, which condensed on the walls and collected in pans. Steam was introduced into the chambers in 1774, and continuous addition of air was begun in 1793. It then was recognized that the sulfurous acid from the burning of sulfur was oxidized by air and needed saltpeter only as a catalyst. In 1827, Gay-Lussac invented a tower for recovering the nitrogen oxides escaping from the chamber. The nitrogen oxides were condensed in sulfuric acid but could not be economically recovered from the acid until the invention of the Glover de-nitrating tower in 1859. The introduction of these two towers completed the chamber process except for various refinements to reduce costs. The process could produce acid with up to 77% H_2SO_4 but generally yielded strengths in a 65–68% range. Higher-strength acid was produced by boiling chamber acid to remove water. A platinum still for producing concentrated sulfuric acid from chamber acid was first built for the Harrison Works in Philadelphia in 1814.

The developing markets for sulfuric acid in the late eighteenth century increased the demand for Sicilian sulfur. By 1832, sulfur prices had risen to $80/t, and stocks rose; then in 1833 the market broke, with the price at $15/t. Sicilian government attempts to stabilize sulfur at $70/t failed, in part because of an 1833 discovery by a French chemist that sulfur dioxide could be obtained by roasting pyrites in a furnace. Process improvements in pyrite roasting by 1870 made pyrites competitive with sulfur as a raw material. By 1880, with the single exception of the United States, the sulfuric acid industries had gone to a pyrites basis. By 1909, virtually all US sulfuric acid came from pyrites and as by-product acid from zinc and copper smelters. In the late 1890s, development of the Frasch process for sulfur mining lowered the price of sulfur. This change, coupled with new developments in sulfur burners, led sulfuric acid manufacturers back to sulfur as a raw material source. During World War I gypsum was burned with coal to produce SO_2 for sulfuric acid production. This process continued in Europe until the 1950s, with the last plant (in Great Britain) shut down in 1975. Today sulfuric acid production from pyrite roasting remains important in Spain, the Former Soviet Union, China, Japan, South Africa, Turkey, and some eastern European nations [30]. The Bosveld No. 2 plant in South Africa operated on gypsum until the early 1980s when it was shut down.

The contact process for sulfuric acid dates from 1831, when a Briton patented a method for converting sulfur dioxide to sulfur trioxide by passing the gas through a heated tube filled with finely divided platinum. The sulfur trioxide was adsorbed in chamber acid to produce concentrated and fuming sulfuric acids. Commercialization of the new process was delayed by lack of markets for the concentrated acid and poor understanding of the process parameters. About 1870, demand for fuming sulfuric acid spurred German development of the contact process. In 1901, BASF reported that the governing principles for the successful manufacture of contact acid were well understood. These included cleaning of the sulfur dioxide gas stream, use of excess oxygen, and temperature control of the catalyst bed. However, the reliance of the contact process on feed acid produced by the chamber process limited its development until about 1930. It became known that the acid in the final SO_3 absorption tower had to be kept in the range 98.5–99.0% H_2SO_4 to be effective. Accurate dilution techniques to allow recycle of the absorber acid were developed in the late 1920s. The contact process then could continuously produce diluted acid, and no longer required chamber acid. The development of poison-resistant vanadium catalysts for SO_2 conversion allowed the use of smelter gases in the contact process. With its lower capital and operating costs, the contact process then supplanted the chamber process for sulfuric acid production. By 1940, the United States had equal numbers of chamber and contact processes. Today no commercial chamber plants are operating in the United States.

Since the 1940s most developments in the contact process have focused on energy recovery and pollution abatement. The production of SO_2, its conversion to SO_3, and the dilution of H_2SO_4 are exothermic processes that are exploited to reduce energy costs at sulfuric acid plants. Sulfur emissions have been sharply reduced by using two SO_3 absorbers, although one absorber is the economic choice. Tail gas scrubbing processes have been developed to further reduce sulfur emissions from sulfuric acid plants. No single absorber plants have been built in the United States since the 1970s, although there were still some in operation as recently as 2000. As of the early 1970s, US EPA regulations limited new sulfur-burning sulfuric acid plant emissions to 2 kg SO_2 and 0.075 kg acid mist per metric ton of H_2SO_4 produced, and metallurgical plants to 650 ppm volume SO_2. During the 1990s some local requirements exceeded the EPA regulations and several plants were constructed with SO_2 limited to 100 ppm SO_2 (about 0.5 kg SO_2 per metric ton H_2SO_4 produced). One double absorption plant followed by an ammonia based DynaWave scrubber, built under Monsanto Enviro-Chem license in Ulsan, Korea, started operation in 1999 with SO_2 at 30 ppm.

Manufacture of Sulfuric Acid by the Contact Process

The basic steps in the contact process are: (1) production of sulfur dioxide; (2) cooling and, for smelters, cleaning of the process gas; (3) conversion of the sulfur dioxide to sulfur trioxide; (4) cooling of the sulfur trioxide gas; and (5) absorption of the sulfur trioxide in sulfuric acid [31]. Figure 25.6a is a photograph of a contact process plant. A simplified diagram of a double absorption contact sulfuric acid process is shown in Fig. 25.6b. Because sulfur dioxide is produced by several processes, it is convenient to separate the discussion of sulfur dioxide production from its conversion to sulfuric acid.

Sulfur Dioxide Production

Sulfur is converted to sulfur dioxide by burning molten sulfur with dried air in a sulfur burner to yield a 1,000–1,200°C gas stream containing 10–12% SO_2. The burner is mounted at one end of a sulfur furnace, and the gas passed through a waste heat boiler at the other end. The gas temperature is reduced to 420–440°C on residual oxygen level limits the formation of nitrogen oxides, which otherwise would be very high at the combustion temperature of 1,750°C. The combustion gas then is cooled to 620–650°C, and additional dried air is injected to burn the residual sulfur in an afterburner. A second waste heat boiler cools the process gas to 420–440°C before sending it to the converter.

Pyrites and other iron sulfides are roasted to produce an iron oxide cinder and an off-gas containing 7–14% SO_2 which is contaminated with varying amounts of arsenic, lead, zinc, and other metal oxides. The off-gas must be cleaned before it is sent to the converter. Various types of pyrite roasting equipment have been used in the past, including shaft furnaces, multiple hearth roasters, rotary kilns, and dust roasters. Fluid-bed roasters now are widely used for their superior process technology, throughput rates, and economics. The roasting process must be controlled between 850 and 940°C. At lower temperatures the reaction is incomplete, whereas at higher temperatures the iron oxides and sulfides form a eutectic melt that inhibits the reaction rate. Fluid-bed roasters surpass other types in temperature control and temperature uniformity throughout the bed. Fluid-bed technology for SO_2 production was introduced in the early 1950s by Dorr Oliver in the United States and by BASF in the Federal Republic of Germany.

Copper, lead, zinc, and other sulfide ores may be processed by roasting or smelting. Roasting or sintering of sulfide ore is essentially identical with pyrite roasting. Sulfur melting generally occurs at higher temperatures. Older reverberatory furnace smelters produce off-gas with only 1–2% SO_2 too low for its economical recovery as acid. By using an oxygen-enriched air feed, the off-gas can be raised to 6–8% SO_2 Bath smelters (Mitsubishi, Noranda), where the ore concentrates are heated and reacted in the slag/matte melt, produce an off-gas with 10–20% SO_2 [32]. Flash smelters (lnco, Outokumpu), which involve suspension and reaction of the concentrates in an oxidizing gas stream, operate at 1,200–1,300°C and produce a waste gas with 10–15% SO_2 Oxygen enrichment of the feed air can raise the SO_2 level in the off-gas to 30–80%. The Kivcet process smelts with pure oxygen and produces off-gas with 80–85% SO_2 [33]. Normally strong gases are diluted with air to 14% SO_2 in order to limit outlet temperatures in the converter first pass to less than 640°C and to provide sufficient oxygen to convert the SO_2 to $SO_{3.}$ In 1996 the original Olympic Dam sulfuric acid plant in South Australia was modified to

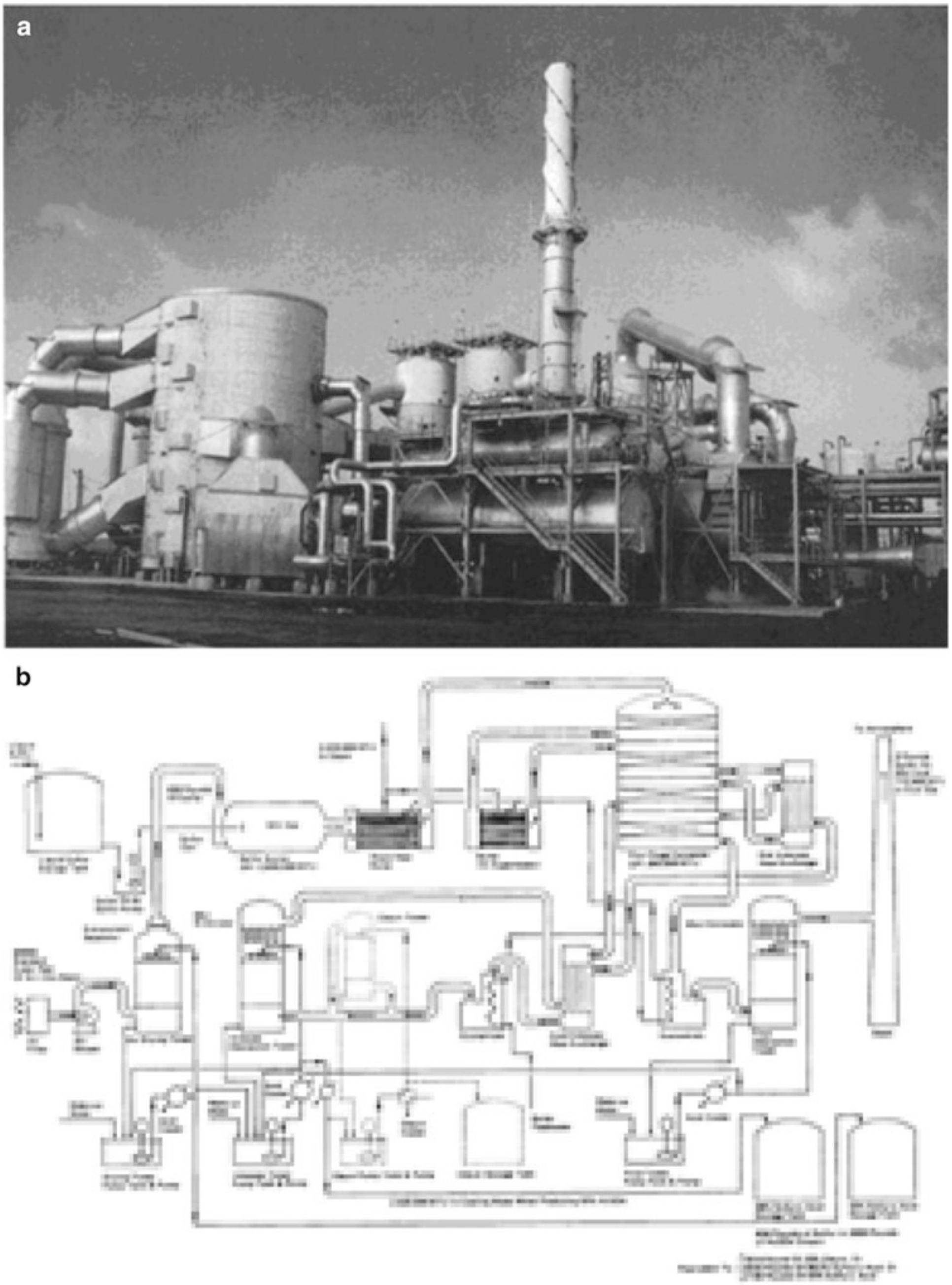


Fig. 25.6 (**a**) Modern double absorption sulfuric acid plant with view of sulfur furnace in foreground. (*Courtesy* Monsanto Enviro-Chem.). (**b**) Schematic flow diagram of a double absorption contact process for sulfuric acid manufacture

operate with 18% equivalent SO_2 gas strength. This plant used a Monsanto pre-converter and a cesium promoted catalyst [34]. Figure 25.7 is a photo of a metallurgical gas sulfuric acid plant.

Off-gas from roasting and smelting operations may contain dust, SO_3, halogens, NO_x, arsenic and other toxic metal fumes, and mercury [35, 36]. These components must be removed from the gas stream before it is sent to the converter. Although SO_3 is produced in the converter, its presence in the cooled gases upstream of the converter will cause excessive corrosion by forming sulfuric acid mists. The cleaning plant steps are: (1) hot-gas dedusting; (2) wet scrubbing; (3) gas cooling; (4) mist removal; and, if necessary, (5) mercury removal. The hot gases generally are passed first through a waste heat boiler to reduce the temperature to 250–400°C. Cyclones followed by hot-gas electrostatic precipitators (50–90 kV) remove nearly all (99+%) of the dust. The gas then is contacted with weak (5–30%) H_2SO_4 in an open spray tower which removes metal vapors and additional solids, cools the gas to 50–80°C, and converts SO_3 to acid mist for later removal. If higher levels of particulate removal are required, a venturi, Swemco, or DynaWave scrubbers may be used. The saturated gas is next cooled in a packed tower or shell and

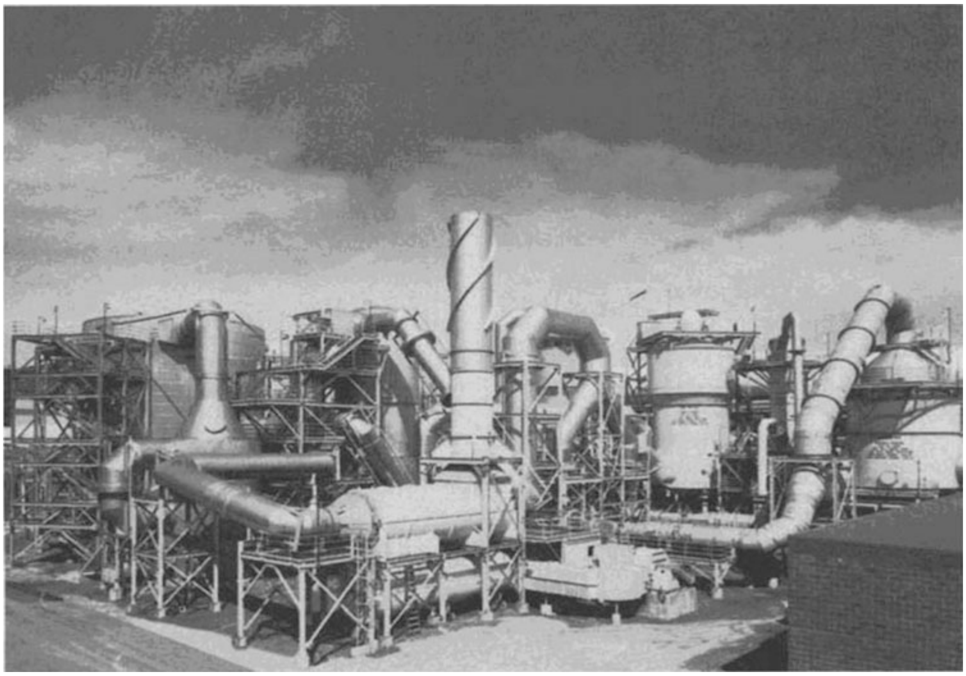


Fig. 25.7 Modern metallurgical sulfuric acid plant with view of preheating furnace in foreground. (*Courtesy* Kennecott, Monsanto Envior-Chem, and Manly Prim Photography)

tube heat exchanger to condense excess water. Shell and tube heat exchangers of graphite or alloy construction are generally used only in smaller size sulfuric plants or where the gas temperature has a tight approach to available cooling water. Silica packing is used in the towers, or sodium silicate is added to the weak acid circuits to remove fluorides as flourosilicic acid.

In cases where the fluoride levels are very high, additional liquid–gas contacting stages are provided to reduce the gas phase concentration of fluorides. Acid mist is removed in wet electrostatic precipitators fitted with lead tubes and star wires or, more recently, with FRP or PVC tubes specially treated to maintain a conductive liquid film surface (Lurgi) [37] and composite wires fabricated with barbs to promote corona discharge [38]. If present, mercury is removed in an additional tower by scrubbing with mercury chloride solution (Boliden) [39], hydrochloric acid solution, or 70–85% sulfuric acid.

Waste sulfuric acid sludges from petroleum refineries are disposed of by conversion to SO_2 for production of fresh sulfuric acid. The heavy organic components of the sludges can be decomposed thermally at 800–1,300°C (Lurgi, Monsanto Enviro-Chem) or reductivity at 200–600°C with coal in a rotary kiln (Chemico). Thermal decomposition is accomplished in a fuel-fired vertical or horizontal furnace. The acid sludge also can be injected into fluid-bed pyrite roasters as a means of disposing of the acid and reducing fossil fuel consumption. Dilute acid sludges must be concentrated to 60–75% H_2SO_4 for economical conversion. This is generally done by using waste heat from the decomposition process. The sulfur dioxide gas stream from these processes requires cleaning, as described above for roasting the smelting plants.

Calcium sulfate may be decomposed to cement clinker and sulfur dioxide gas in a coke-fired rotary kiln at 900–1,400°C (Miiller-Kiihne) [40]. However, the unfavorable economics of this process relegate it to countries that do not have other sources of sulfur. Phosphogypsum (gypsum produced by the acidulation of phosphate rock) may be decomposed in this way as a means of recycling the sulfur values in the large waste phosphogypsum piles at fertilizer plants (OWS-Krupp and FIPR/Davy McKee).

This process is hampered by the more extensive gas cleaning requirements for decomposing phosphogypsum as compared with natural gypsum. During the early 1980s there was some interest in recycling phosphogypsum. By 2000, because of the high cost of these processes, there was little commercial interest. Environmental forces also are behind the recycling of ferrous sulfate from metals industry pickling liquors. This “green salt” is decomposed to sulfur dioxide and iron dioxide in pyrite roasters. Elemental sulfur, coal, or fuel oil may be used as supplementary fuels.

Refinery waste gases may be burned to eliminate hydrogen sulfide and other sulfur containing contaminants. Streams containing small amounts of H_2S or constituents unsuitable for Claus plants may require combustion to SO_2 as the means of disposing of the toxic gas. The resulting effluent gas usually is low in SO_2 and contains water vapor and carbon dioxide. Flue gases from fossil fuel power plants also fall into this category. Recovery of the sulfur values from these dilute gases usually is driven by environmental considerations rather than economics. In the United States, power plant flue gas often is scrubbed with lime to convert the SO_2 to gypsum sludge for landfill disposal. Alternatively, the BergbauForschung process recovers sulfur dioxide by dry adsorption on activated coke at 130°C. The SO_2 is released by heating the coke to 600–650°C. Sulfur dioxide can be absorbed in a sodium sulfite solution (Wellman-Lord) to produce sodium bisulfite. Pure moist sulfur dioxide can be recovered by heating the sodium bisulfite.

Single Versus Double Contact Process

The single absorption contact process for sulfuric acid is characterized by four main process steps: gas drying, catalytic conversion of SO_2 to SO_3, absorption of SO_3, and acid cooling. The maximum SO_2 conversion for a single absorption plant is about 97.5–98%. By adding a second SO_3 absorber with one or two catalyst beds between absorbers, the SO_2 conversion can be increased to 99.5–99.8% or even as high as 99.9% with a cesium promoted catalyst, resulting in lower SO_2 emissions. The so-called double absorption process is now the industry standard.

If water vapor is present in the gas stream or the gas temperature or metal surface temperatures drop below the dew point, liquid acid is formed by condensation of H_2SO_4 vapor rather than by absorption of SO_3. Therefore, the SO_2 laden process gas sent to the converter must be dry to protect the downstream process equipment against corrosion. The drying generally is done in a packed tower with recirculating concentrated (93–98%) sulfuric acid kept at 50–60°C by indirect cooling. The tower acid stream is heated by condensation of the water and by dilution of the acid. The tower temperature is used to control the moisture level of the gas sent to the converter. Acid from the drying tower is cross flowed to the absorber or is sent to storage tanks for shipment. When sulfur is burned, the combustion air to the sulfur burner is dried because the combustion of sulfur does not produce water. Off-gases from pyrite roasters and metallurgical smelters are dried as part of the gas cleaning process. For roaster gases with low SO_2 concentrations, a predryer may be added upstream of the main dryer.

Oxidation of SO_2

Oxidation of SO_2 to SO_3 is accomplished in multi-stage, fixed-bed catalytic converters equipped with interstage boilers or heat exchangers to remove the heat of reaction. Typically, four stages are compartmented within a single vertical converter, which may be brick-lined, steel, or cast iron. Newer converters are stainless steel, and some have five stages for higher conversion. Isothermal tubular converters are no longer suited to modem high-capacity plants. The extruded cylindrical catalyst pellets are usually 4–9% V_2O_5 with alkali metal sulfate promoters on a silica carrier (diatomaceous earth, silica gel, or zeolites). The reaction temperature for vanadium catalyst is generally 410–440°C. In the late 1980s a cesium promoted catalyst became commercially available from Topsoe and Monsanto. These low-strike catalysts operate at 360–400°C. Higher temperatures (~600°C) reduce the SO_2 conversion and lead to structural damage of the catalyst. High-pressure drops across catalyst beds from catalyst dust formed during processing require periodic catalyst removal and screening to remove dust. Ring-shaped catalysts developed by Topsoe and others have lower dust pressure drops and are now in wide use. Other catalyst shapes used are ribbed rings and cylinders. The usual catalyst loading per 1 t/day sulfuric acid capacity is 150–200 L in a double absorption plant and 200–260 L in a single absorption plant. Bayer developed and operates fluid-bed converters that utilize special 0.3–1 mm abrasion-resistant catalysts.

Absorption of SO_3

Sulfur trioxide from the converter is absorbed in 98% H_2SO_4 recirculated counter currently through a packed tower maintained at 60–80°C by indirect cooling. The optimum concentration of the absorber acid is near the H_2O–H_2SO_4 azeotrope, 98.3% H_2SO_4, where the SO_3, H_2SO_4, and H_2O vapor pressures are at their lowest values. Absorption efficiencies in excess of 99.9% generally are obtained. On leaving the converter, the process gas is cooled first with feed gas in a gas-gas heat exchanger and then with boiler water in an economizer to 180–220°C before it enters the absorber. An impingement separator, or Teflon or glass fiber mist eliminator, is placed in the top of the absorber to remove acid mists. If oleum is produced at the plant, it is made in a separate oleum tower upstream of the absorber. A portion of the SO_3 stream to the absorber is diverted to the oleum tower where it is absorbed in a recirculating stream of oleum.

In double absorption plants an intermediate absorber is placed between the second and third (or between the third and fourth) converter beds. By removing SO_3 from the gas stream at this intermediate point, higher SO_2 conversions are attained in the downstream converter beds, and the overall SO_2 conversion is increased. The cooled gas from the intermediate absorber is reheated by hot converter gas in gas–gas heat exchangers before returning to the converter. An oleum tower may be placed before the intermediate absorber.

Acid Cooling

Absorption of SO_3 in concentrated sulfuric acid and the formation of H_2SO_4 from SO_3 and H_2O produce heat in the absorber, as does acid dilution from the addition of makeup

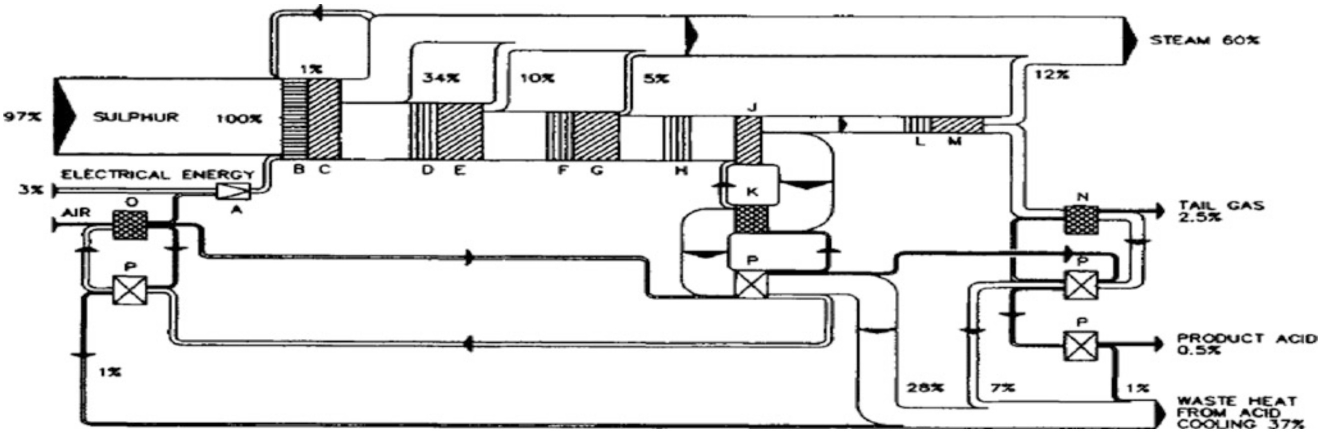


Fig. 25.8 Sankey energy flow diagram for a 1,000 t/day sulfur-burning double absorption sulfuric acid plant (feed gas 10 % SO_2). *A* blower; *B* sulfur furnace; *C* waste heat boiler; *D* catalyst bed 1; *E* steam superheater; *F* catalyst bed 2; *G* boiler; *H* catalyst bed 3; *J* intermediate heat exchangers; *K* intermediate absorber; *L* converter bed 4; *M* economizer; *N* final absorber; *O* air dryer; *P* acid coolers. (*Courtesy* Lurgi GmbH, Frankfurt, Germany)

water. Process control requires that the acid be cooled before it is recirculated to the dryer or absorber towers or sent to storage. Earlier acid coolers of parallel banks of stacked, irrigated, cast iron sections have been largely replaced by stainless steel shell and tube or plate exchangers, with or without anodic protection. Hastelloy, Sandvik SX, ZeCor, and Saramet alloys and Teflon linings are also used in acid piping and coolers [41–44].

Tail gas emissions are controlled by improving the SO_2 conversion efficiency and by scrubbing the tail gas. In a double absorption process plant, a five-bed converter has 0.3% unconverted SO_2, as compared with 0.5% for a four-bed converter. A Lurgi Peracidox scrubber may be used to remove up to 90% of the residual SO_2 in the tail gas from a double absorption plant. Hydrogen peroxide or electrolytically produced peroxymonosulfuric acid is used to convert the SO_2 to H_2SO_4 in the Lurgi scrubber.

Other Modifications to the Sulfuric Process [45]

Tail gas from single absorption plants may be absorbed on activated carbon (Sulfacid) or scrubbed with ammonia (Monsanto AMMSOX) or sodium sulfite (Wellman Lord). Metallurgical acid plants differ from sulfur-burning plants in that the cleaned SO_2 process gas must be heated before it is sent to the converter. Many of these plants have weak SO_2 streams that require large gas–gas heat exchangers for temperature control. Four plants in the Former Soviet Union processing 2–4% SO_2 use an unsteady-state oxidation process in which the cold (40–70°C) SO_2 gas is reacted on hot catalyst beds without intervening heat exchangers [46, 47]. As the temperature front moves through the bed to the exit side, the flow is reversed. Cycle times are 30–120 min, and single-bed conversions of 80–90% are reported as compared with 55–60% for conventional processes at higher exit temperatures.

The thermal capacity of a 1,000 t/day sulfuric acid plant is about 63 MW This heat liberation must be controlled in a manner that maintains optimum gas temperatures in the converter system and optimum acid temperatures in the dryer and absorber circuits. Tail gas emissions also are affected by the energy balance. Figure 25.8 shows an energy flow diagram for a contemporary sulfur-burning sulfuric acid plant. About 97% of the total energy input derives from burning sulfur, and 3% comes from the electricity consumed to drive the gases through the plant. Most plants can recover 55–60% of the energy as high-pressure steam (40–60 bar, 400–480°C), but about 40% is lost as waste heat dissipated to the environment in the form of hot water from acid coolers.

During the late 1970s acid plants were optimized to generate more steam. Steam can be produced at pressures up to 80 bar from high-temperature sulfur burners such as the Lurgi two-stage combustion system. The high-pressure steam is let down to low pressure steam through a turbogenerator that co-generates electricity. In double absorption plants, economizers were installed upstream or downstream of the heat exchanger servicing the intermediate absorber. Boiler feed water is preheated to 90–95°C in this economizer to increase steam production. The energy production from acid plants was increased to 70% by installing low-gas-temperature economizers, low pressure-drop catalysts, and suction drying towers, by increasing the SO_2 feed gas concentration, and by preheating the boiler feed water with hot acid. Further energy recovery requires higher operating temperatures for the absorbers and acid coolers. Venturi concurrent absorbers operating at acid temperatures of 130–140°C are installed in several plants. To recover acid

Table 25.7 World production of new sulfuric acid for 1999 and 2009 (million metric tons 100 % H_2SO_4)

	1999		2008		Difference	
	From elemental sulfur	Total	From elemental sulfur	Total	From elemental sulfur	Total
Western Europe	8.5	17.5	7.1	14.2	−1.4	−3.3
France	1.4	2.2	0.5	0.7	−0.9	−1.5
Germany	1.2	3.0	1.4	3.3	0.2	0.3
Spain	0.3	2.8	0.9	2.6	0.6	−0.2
Central Europe and FSU	10.3	15.6	12.5	17.9	2.2	2.3
Poland	1.2	2	1.3	2.3	0.1	0.3
FSU	8.3	12	11.2	15.6	2.9	3.6
North America	37.6	46.2	25.2	30.8	−12.4	−15.4
Canada	1.6	4.8	1.1	4.4	−0.5	−0.4
US	33.2	37.1	24.1	26.4	−9.1	−10.7
Mexico	2.8	4.3	2.8	4.4	0	0.1
Latin America	5.6	10.2	9.8	18.5	4.2	8.3
Brazil	4.0	4.8	5.2	6.1	1.2	1.3
Africa	17.1	18.1	17.5	18.5	0.4	0.4
Morocco	8.5	8.5	8.9	8.9	0.4	0.4
Tunisia	4.8	4.8	4.3	4.3	−0.5	−0.5
South Africa	2.5	3.1	2.5	2.9	0	−0.2
Asia	16.5	42.0	27.2	68.9	10.7	26.9
China	4.3	21.6	20.8	53.8	16.5	32.2
Japan	2.2	6.9	1.7	7.2	−0.5	0.3
India	5.2	6.1	4.7	7.9	−0.5	1.8
Oceania	1.1	2.5	2.5	5.1	1.4	2.6
World total	102.8	158.7	113.1	191.3	10.3	32.6

Source: USGS, Con-SuI, Inc. and Freeport Sulfur Marketing Department estimates

heat directly as steam, it was necessary to increase the acid temperatures from 110 to 120°C to about 200°C.

As of 2005 Monsanto Enviro-Chem had built 21 Heat Recovery System (HRS) units since demonstrating the first HRS at Namhae Chemical in South Korea in 1987 [48, 49]. The process is based on 310 stainless steel, which resists corrosion in 98.5% H_2SO_4 at temperatures up to 220°C. The intermediate absorber at Namhae takes 194°C gas from the converter third stage economizer and absorbs the SO_3 in 199°C, 99% acid. Recirculated acid from the absorber is cooled from 220°C in a 10-bar HRS boiler. The added energy recovery for this process is reported to increase the total recovery to 90%. Monsanto's proposed Monarch process combines HRS technology with the wet catalytic converter process (Lurgi) to increase heat recovery and shift it to high-pressure steam production for electric power generation [50].

Other Sources of Sulfuric Acid

Spent sulfuric acid usually is diluted in the process in which it is used: titanium dioxide pigment processing, plastics manufacture, and so on. The dilute acid may be used in processes requiring dilute acid or may be concentrated for reuse by a number of vacuum evaporation processes (Simonson-Mantius, Chemetics) [51] or by thin-film evaporation (DuPont, Bofors). In the submerged combustion distillation process, water is evaporated from the dilute acid by forcing hot flue gases from a fuel-fired burner below the acid surface (Chemico). The concentration of 75% acid to 95–98% H_2SO_4 by the Pauling-Plinke process is done by feeding the 75% acid to a stripping column fitted with a stirred cast iron pot mounted in a furnace. The acid concentration in the pot must be kept above 80% to minimize corrosion.

Chemetics has developed a process for treating spent alkylation sulfuric acid with nitric acid to produce a sulfuric acid that can be used to acidulate phosphate rock, the major use for sulfuric acid. The organic contaminants are converted to carbon particles that are removed with the gypsum on filtration of the phosphoric acid. Special alloys are used in the fabrication of the acid reactor. Topsoe developed and, by the year 2005 had built, more than 45 wet sulfuric acid (WSA) process units. This process is especially suited for low-strength, less than 4%, SO_2 gas streams which would not be auto thermal or meet water balance conditions in the conventional dry sulfuric acid contact process. In the WSA process wet SO_2 gases pass through converter beds where the SO_2 is oxidized to SO_3. The SO_3 reacts with water vapor to form H_2SO_4 in the gas phase. The acid is condensed in proprietary WSA condensers. Sulfuric acid is produced at concentrations around 98% [52, 53].

Table 25.8 Sulfuric acid production from pyrites and other forms (million tons 100 % H_2SO_4)

	1999		2008		Difference	
	Pyrites	SOF	Pyrites	SOF	Pyrites	SOF
World	16.0	40.0	18.4	59.8	2.4	19.8
Western Europe	2.1	6.9	0.7	6.4	−1.4	−0.5
Eastern Europe FSU	1.0	4.3	0.5	4.0	−0.5	−0.3
North America		8.6		5.6		−3.0
Asia	12.3	13.4	17.0	29.2	4.7	15.8

Source: USGS, Con-Sul, Inc., Freeport Sulfur Marketing Department

Table 25.9 Sulfuric-acid-consuming industries in the United States (millions of metric tons, 100 % H_2SO_4)

Consuming industries	2008	1999	1989	1979
Agriculture	24.0	24.5	26.9	24.1
Phosphatic fertilizers	23.4	23.7	26.4	23.2
Nitrogenous fertilizers	0.6	0.6	0.3	0.5
Pesticides		0.1	0.1	0.2
Other agricultural		0.1	0.1	0.2
Other industries	0.6	7.2	11.5	14.1
Petroleum refining	9.6	1.6	2.1	2.4
Synthetic rubber and plastics	0.3	1.2	1.0	
Rayon and cellulose			0.2	0.3
Other chemicals	1.5	1.4	2.7	4.2
Copper ore processing		2.2	1.8	2.1
Uranium and other ore processing		0.3	0.4	
Iron and steel			0.2	0.9
Other metals	0.3	0.2	0.2	0.1
Pulp and paper	0.6	0.4	1.0	0.8
Pigments and paint	0.9	0.5	0.4	0.6
Other	0.9	0.9	1.6	1.6
Total	38.7	31.7	38.4	38.1

Source: U.S. Geological Survey Minerals Yearbook and U.S. Bureau of Mines: Sulfur Annual Report

Production and Consumption of Sulfuric Acid

The world production of sulfuric acid (2008 and 1999) is given in Table 25.7. Production increased 32 million tons. Output from elemental sulfur feedstock rose about 10% while that from "other" sources increased by 40%. Increased recovery from smelters were a major contributor in the latter's increase. While China reduced output from pyrites slightly, environmental recovery projects at smelters added a considerable volume.

China represents the most dramatic change: 21.6 million tons of output in 1999 versus 53.8 million tons in 2008. Of that, only 4 million tons were produced using elemental sulfur in 1999 vs. 21 million tons in 2008. The increase from effluent recovery was also impressive—from 5 million to 14 million tons in 2008 as China recognized its global environmental responsibilities and the positive economics of such investments.

Table 25.9 lists the sulfuric-acid-consuming industries in the United States and shows the trends in their acid consumption rates from the 1980s to date (Table 25.8).

Nearly all the sulfuric acid consumed in agriculture was reacted with phosphate rock (principally $Ca_9(PO_4)_6CaF_2$) to produce phosphoric acid. Granular phosphate fertilizers are produced by ammoniating phosphoric acid to yield mono- and di-ammonium phosphates; ammonium phosphate is also produced as a fertilizer. Petroleum refining is the largest non-fertilizer use for US sulfuric acid. The acid competes with hydrogen fluoride as a catalyst in petroleum alkylation reactions for gasoline production. Sulfuric acid acts as a catalyst in synthetic rubber and plastics manufacture. Copper ore leaching is carried out by distributing the acid over leach piles of the ore and collecting the copper-rich leachate for processing. Sulfuric acid from nearby smelters is normally used in copper ore leaching.

Acknowledgement The authors wish to acknowledge that major portions of this chapter are taken from the ninth edition version (1992) written by Dr Robin W. Strickland.

References

1. British Sulphur Corp: Ltd, *Sulphur*, No. 193,33–36 (Nov/Dec. 1987)
2. d'Aquin GE *North American Quarterly Sulfur Review*, Con-Sul, Inc., Bigfork; Jan 15
3. (a) Sulfur Magazine, No. 259, p 41, Nov–Dec 1998, Gerrie du Plessis. (b) Hyne JB, "The Forming, Handling and 1230 Transportation of Elemental Sulphur, Presented at Sulphur 1990 Conference, British Sulphur Corporation, London
4. (a) Ibid, 45(b). (b) Kemp E, Hyne JB, Rennie W (1971) Reaction of Elemental Sulfur with Water Under U.V Radiation. Internal. J. Sulfur Chem Part Al(1) 69–70
5. (a) Personal communication, Pennzoil Sulphur Co. Operations at Antwerp terminal. (b) Dowling, N.I., Clark PD, Hyne JB (1996) Understanding and Mitigating Corrosion During Handling and Transportation of Elemental Sulfur. Sulphur 1996 Conference, British Sulphur Corporation, London
6. Sandvik Process Systems, Inc., 21 Campus Rd., Totowa, NJ, USA
7. Berndorf Belt Systems USA, 920 Estes Avenue, Schaumburg, IL, USA
8. Devco International, Inc., 6846 S. Canton Ave., Suite 400, Tulsa, OK, USA
9. Martin Midstream Partners, P.O. Box 191, Kilgore, TX, 75663
10. Enersul, Inc., 7210 Blackfoot Tr. SE, Calgary, Alberta, Canada
11. Kaltenbach-Thuring SA, 9 Rue de l'Industrie, 6000 Beauvais, France.
12. d'Aquin GE (2002) Transporting sulfur pellets, U.S. Patent 6368029, Apr. 9, 2002
13. Hyne JB. Some Impurities in Elemental Sulphur-Origins and Elimination. Sulphur 1991 Conference, The British Sulphur Corp, London
14. Thieler E (1936) *Sulphur*. Theodor Steinkopff, Leipzig
15. Iraq and Poland were the only nations with operating Frasch mines in 2001. In 2012, Poland remains in production and Iraq is considering re-starting operations.

16. BC Insight (2010) *Sulphur Magazine*, Jan/Feb, 2011; US Geological Survey 2000
17. d'Aquin GE (Apr 2002, Apr 2011) *North American Quarterly Sulphur Review*, Con-Sul, Inc., Bigfork, MT
18. Ibid. 17.
19. British Sulphur Corp. London (Nov/Dec. 1987) *Sulphur*, No. 193, 26–30
20. British Sulphur Corp. London (Sept./Oct. 1987) *Sulphur*, No. 192, 30–39
21. Rameshi M. State of the art in gas Treating, Sulphur 2000 Conference, The British Sulphur Corporation, London
22. Ibid. 21.
23. Ibid. 16.
24. Ober J. Sulfur-2000 Annual report, Mineral industry surveys, U.S. Geological Survey, U.S. Dept of the Interior, Reston, VA
25. Sander UHE, Fischer H, Rothe U, Kola R. (1984) *Sulphur. Sulphur dioxide, and sulphuric acid*. Verlag Chemie, Weinheim, Germany
26. Heydorn B, Aguiar D, Ferguson A (1987) Sulfuric acid. In: Chemical economics handbook. SRI International, Menlo Park, CA
27. Fairlie AM (1936) *Sulfuric acid manufacture*. Reinhold, New York
28. Duecker WW, West JR (eds) (1959) *The manufacture of sulfuric acid*. Van Nostrand Reinhold, New York
29. Wyld W (1923) Sulphuric acid and sulphur dioxide raw materials. D. Van Nostrand, New York
30. The British Sulphur Corp. Ltd (1989) World sulphur and sulphuric acid atlas, 5th edn. Purley Press Ltd, London
31. Ibid. 25.
32. Sohn HY, George DB, Zunkel AD (1983) Advances in sulfide smelting. The Metallurgical Society of AIME, Warrendale, PA
33. British Sulphur Corp. London (1987) Sulphur, No. 190, 33–40
34. Bhambri N, Fell RC, Fries RM, Ritschel PM. Metallurgical Sulfuric Acid Plants for the New Millennium. In: Sulphur 1998 conference, British Sulphur Corporation, London
35. Ibid. 25.
36. Friedman L (1988) The Sulfuric Acid Plant Recent Advances-Reviews and Analysis. AIChE Meeting, Clearwater, FL.
37. British Sulphur Corp. London. (1990) Sulphur, No. 203, 24–32
38. Ibid. 36.
39. British Sulphur Corp. Ltd (1990) Sulphur 207:37–41
40. Ibid. 25.
41. Davies M, Hodgson DS, Rodda I (1988) Applications of SARAMET in H_2S0_4 plants. Sulphur 1988 Conference, British Sulphur Corporation, London
42. Douren L (1987) Sandvik SX-The Future Material for Concentrated Sulfuric Acid. Sulphur 1987 Conference, British Sulphur Corporation, London.
43. Holdschick H, Jurascheck MA (1988) Advances in the Development of Cast Materials for High Concentrated Sulphuric Acid at Temperatures up to 280°C. Paper read at Sulphur 1988 Conference, British Sulphur Corporation, London
44. Puricelli S, Fell RC, Randolph DR (2000) The State of Sulfuric Acid Technology. AlChE meeting, Clearwater, FL.
45. McAlister DR, Grendel RW, Schneider D, Shafer IR, Tucker IS (1990) A Sulfuric Acid Plant for the 1990s, Sulphur 1990 Conference, British Sulphur Corporation, London.
46. Blicharz M (1988) Different Methods of Sulphur Removal from Gas from Non-ferrous Metal Works. Sulphur 1988 Conference, British Sulphur Corporation, London.
47. Matros YS, Bunimovich GA (1990) Reverse Process of $S0_2$ Oxidation in Sulfuric Acid Production. Sulphur 1990 Conference, British Sulphur Corporation, London.
48. Shafer, IR (1988) Namhae Heat Recovery System Update. Fertilizer Industry Round Table, Baltimore, MD.
49. Smith RM, Sheputis I, Kim UB, Chin YB (1988) Sulfuric Acid Heat Recovery System (HRS) Operations at Namhae Chemical Corporation, Korea. Sulphur 1988, British Sulphur Corporation, London.
50. Ibid. 45.
51. Lebel G (1990) Pilot Plant Development of the Chemetics TiO_2 Waste Acid Recovery Process. Sulphur 1990 Conference, British Sulphur Corporation, London
52. British Sulphur Corp. Ltd (1997) Sulphur 253:67–69
53. Polk P (2005) e-mail to Fell, R.C., Topsoe Houston, Tx.

Salt, Chlor-Alkali, and Related Heavy Chemicals

26

Thomas F. O'Brien

Sodium Chloride

There are very many inorganic salts, but only one of them, sodium chloride, is commonly referred to by the simple name "salt." Sodium chloride is ubiquitous, and more than 14,000 uses have been tabulated [1]. It is a raw material in the production of many chemicals, including chlorine, caustic soda (sodium hydroxide), synthetic soda ash (sodium carbonate), sodium chlorate, sodium sulfate, and metallic sodium. Indirectly, it is also used to produce many other sodium salts and the very useful hydrochloric acid.

Pure sodium chloride crystals (halite) are colorless to white and contain 39.34% Na and 60.66% Cl. NaCl is rather soluble in water, and temperature has little effect on its solubility. A saturated solution holds 26.3% (w/w) NaCl at 0°C and 28.1% at 100°C.

Salt is a nutrient essential to both humans and animals. It has been a vital commodity throughout human history. Wars have been fought over its supply, words derived from "salt" appear in all languages and a vast number of place names, and the very spread of civilization depended on its availability.

Besides serving as a chemical feedstock, sodium chloride is used in food processing to add flavor and palatability and as a texturizing agent in bread and other baked goods. Salt also is used as a fermentation controller in dairy products, a bactericide in preserved foods, and a color developer in processed meats. A few of its other applications are hide-processing, leather tanning, fabric dyeing, and water softening. China and the United States are the leading producers of salt (Table 26.1). Between them, they account for more than 40% of the world's supply. Germany, India, Canada, and Australia form the next tier, producing more that 20% of the total. One of the largest uses in the United States is as a highway de-icing and anti-icing chemical [2]. Table 26.2 summarizes US consumption of salt by its principal uses. Annual consumption can fluctuate widely, as shown by the first column of Table 26.3. This is almost entirely due to variations in highway salt usage (column 2).

There are underground deposits of salt ("rock salt") throughout the world. NaCl also makes up about 3% of seawater [3]. The latter is an inexhaustible supply, used in many parts of the world to produce solid salt ("solar salt") by evaporation. In the United States, rock salt predominates. There are mines in Kansas, Louisiana, Michigan, New Mexico, New York, North Dakota, Ohio, Oklahoma, and Texas. Some Canadian provinces also mine salt deposits.

There are two distinct types of rock salt deposit, stratified beds and domes. In the first of these, the salt is in relatively shallow layers that often cover great areas. They may be roughly horizontal or disturbed by tectonic movements. Depths below the surface may range from 100 m to more than a kilometer. Shafts from the surface to the deposit typically are 6 m in diameter. Usually, the salt is in a number of strata, interspersed for the most part with shale or anhydrite. Total thickness of the deposit may be several hundred meters. Each stratum may be mined separately by drilling and blasting or by using a continuous machine. Mining is by the room-and-pillar method, with frequent crosscuts between rooms. A typical room may be 15 m wide and 8 m high.

The configuration of salt domes is very different. A dome forms when deep-lying salt is forced into plastic flow through faults in its overburden. The dimensions of a dome may be measured in kilometers, and it may contain many billions of tons of salt. In the United States, domes are found on or near the Gulf Coast. Mining is similar in principle to that practiced in bedded deposits, but individual blasts may be on a much larger scale.

Salt can also be extracted from underground deposits by solution mining. In this process, a small-bored shaft is drilled into the cavity or deposit. It is then fitted with concentric pipes, one to pump water into the deposit and the other to

T.F. O'Brien (✉)
800 Avondale Rd., Apt. 3J, Wallingford, PA, USA
e-mail: mudville9inc@aol.com

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_26,

Table 26.1 Estimated salt production, 2010

Country	Production[a]
China	60
United States	45
Germany	16.5
India	15.8
Canada	14
Australia	11.5
Mexico	8.8
Brazil	7
World total	270

[a] Millions of metric tons

Source: United States Geological Survey, Mineral Commodity Summary, *Salt*, 2011

Table 26.2 Salt consumption in United States, 2010

Market	% of total
Chemical industry	40
Highway de-icing	38
Distributors	8
Agricultural	4
Food	4
General industry	2
Water treatment	2
Others (incl. exports)	2

Source: United States Geological Survey, Mineral Commodity Summary, *Salt*, 2011

Table 26.3 US dry salt sales[a]

Year	Total	Highway	Non-highway
2005	32.0	20.5	11.6
2006	23.1	12.1	11.0
2007	31.7	20.3	11.5
2008	34.2	22.2	12.0
2009	27.3	16.5	10.8

[a] Millions of metric tons

Source: www.saltinstitute.org/Uses-benefits/Salt-in-industry, 2011-08-12

bring saturated brine back to the surface. The purity of the brine is enhanced because insoluble and very slowly soluble impurities, notably anhydrite (calcium sulfate), remain in the cavity as settled residue. The geometry of a bedded salt deposit (large area, small thickness) does not favor injection and withdrawal at essentially the same point. Instead, it is normal practice to operate galleries that contain numbers of pipes placed at horizontal intervals along the deposit. The dissolving fluid goes down one pipe, and the saturated brine rises through another some distance away. The cavity naturally formed as salt dissolves then takes the shape of the deposit much more efficiently.

Many processors mine their own deposits this way and feed the salt solution (brine) directly into their processes. Other companies specialize in large-scale solution mining and serve the processing industry as brine merchants. When it is not necessary to evaporate the brine down to solid salt, this is normally the cheapest form of supply.

The saturated raw brine can be handled in several ways. Evaporation in vacuum pans without chemical treatment produces evaporated granular salt, which usually assays more than 99% NaCl. The evaporation always is in multiple-effect systems, in which the vapor produced in one effect becomes the heating medium in another, or in vapor-recompression systems. The steam economy achieved in either case can be quite high. Evaporator condensate can be used as process water or as a source of steam.

Chemical treatment of raw brine before evaporation enhances the purity of the final product salt. For the most part, this involves the precipitation of calcium and magnesium by adding, respectively, soda ash and caustic soda:

$$\mathrm{Ca^{2+} + Na_2CO_3 \rightarrow 2Na^+ + CaCO_3}$$

$$\mathrm{Mg^{2+} + 2NaOH \rightarrow 2Na^+ + Mg(OH)_2}$$

After clarification of the brine to remove the precipitated mud, evaporation produces a salt that is more than 99.95% NaCl, often 99.99%. The higher purity is advantageous in all processes and allows the use of salt in new applications. Examples are food (human and animal), raw material for certain chemical reactions, and polishing pharmaceutical tablets. Purified salt can also be used in the pharmaceutical industry to make hemodialysis and intravenous fluids [4].

Sea salt, or solar salt, is harvested in many countries. Flat coastlines, warm breezes, plentiful sunlight, and little rainfall are all desirable, and operation may be seasonal. Solar salt works naturally are most common in low-latitude countries. Production also is favored in inland locations where brines more concentrated than seawater are available. The process is passive; salt water from lake or ocean is trapped into shallow ponds and allowed to concentrate under the influence of solar energy. In the case of seawater, about 90% of the volume must evaporate before salt begins to precipitate.

The salinity of seawater varies with its location, but a global average is about 3.5 wt.% total dissolved solids. Reported compositions frequently are normalized to this basis. The composition of seawater is complex, and total evaporation would produce solids that are only about 78% NaCl (anhydrous basis). Production of salt therefore involves a series of fractional crystallizations, and a solar salt plant will include a large number of separate ponds with different functions. These include (a) concentrating ponds, which allow suspended solids to settle and bring the brine close to saturation, (b) lime ponds, which allow species less soluble than NaCl (principally calcium compounds) to

Table 26.4 Purity of various types of salt (wt%, dry basis)

Component	Rock salt	Washed solar salt[a]	Purified vacuum salt
NaCl	93–99	99	99.95
Sulfate	0.2–1	0.2	0.04
Calcium	0.05–0.4	0.04	0.0012
Magnesium	0.01–0.1	0.01	0.0001
Iron	0.05–0.5	0.03	0.0001

Source: Akzo Zout Chemie brochure, *Electrolysis Salt* (*Sodium Chloride*)

[a] See refs. [3, 5] for recent improvements

Table 26.5 Salt usage by type, USA in 2009

Type of salt	% of total, 2009	Average price, 2010, \$/ton[b]
Rock salt	45	35
Salt in brine[a]	38	8
Vacuum pan salt	10	170
Solar salt	8	70

[a] Predominantly for chlor-alkali manufacture

[b] Average for all forms and quantities; bulk costs of solid salt will be lower

Source: United States Geological Survey, Mineral Commodity Summary, *Salt*, 2011

precipitate, (c) crystallizers, which produce salt crystals, and (d) bitterns ponds, which hold concentrated solutions rich in magnesium salts. Proper management of the process requires transfer of brine from pond to pond at the proper time, and it is not uncommon for seawater to require 2 years or more in process. Division of the stages of precipitation is not clear cut; thus, magnesium precipitation begins before sodium precipitation is complete. The cutoff point is a compromise between yield and quality. Typically, NaCl is collected until the liquor reaches a certain specific gravity, at which point the brine is transferred to the bitterns ponds. Recovery of the NaCl originally present in seawater generally is 70–80%.

The quality of solar salt is determined primarily by the care taken in separating the various products of crystallization and in harvesting salt from the bottom of the crystallizing ponds. Attempts to increase the size of the NaCl cut necessarily increase the level of contamination. Washing of the recovered salt in any case can improve its quality [5]. Magnesium deposited on the crystals and held in occluded bitterns is most easily removed. More vigorous processing can also remove inclusions of calcium, and as a result washed salts may contain more 99% NaCl. Table 26.4 compares this to typical rock and purified vacuum salts. Table 26.5 then breaks down US consumption by type of salt. It reflects the widespread use of on- or near-site solution mining by the chemical industry, followed by direct use of brine.

Particularly when salt is recovered from inland brines, the bitterns may contain valuable minerals and justify further processing. To take a single example, Dead Sea water is rich in bromine and is one of the world's major sources of the element (*infra*).

Sodium chloride is the ultimate source of many industrial chemicals, and this chapter treats a number of them. Sodium phosphates and phosphoric acid are important examples. They do not appear here but rather are covered in Chap. 23.

It is also worth noting that potassium chemicals have similar chemistry and that many of them are produced simply by replacing Na in the raw material with the scarcer and more expensive K.

Sodium Carbonate (Soda Ash)

Soda ash, or sodium carbonate, is derived commercially from three sources: trona/nahcolite ores, naturally occurring alkaline brines, and synthetic processes. Most common in the last category is the Solvay ammonia-soda process. Variations of this process operate in many parts of the world. The United States has 95% of all known deposits of trona. Since natural production has significant cost advantages, the United States no longer produces synthetic soda ash. Because natural reserves are limited elsewhere, synthetic processes still provide about three-quarters of the world's supply of Na_2CO_3. China is the leading producer, with about 40% of the total. The United States produces about 22% of all soda ash. Data on natural soda ash production and reserves are in Table 26.6. Market distribution in the United States is shown in Table 26.7.

The area near Green River, Wyoming, is the primary source of natural ash. The deposit contains over 10^{11} metric tons of trona, enough to satisfy global demand for over 2,000 years [6]. Significant amounts are also recovered from the alkaline brines of Searles Lake in California.

"Trona" is the dihydrate of sodium sesquicarbonate, $Na_2CO_3{\cdot}NaHCO_3{\cdot}2H_2O$. Processing is required to purify it and to calcine the bicarbonate portion. There are two approaches to this task. Purification may either precede or follow calcination. The older (sesquicarbonate) process first hot-leaches the crushed ore to extract the mineral. Purification involves clarification, filtration, and crystallization by cooling. The refined trona then is calcined to form soda ash. The newer (monohydrate) process first calcines the crushed ore to form a crude carbonate. This is dissolved and the solution evaporated to crystallize the monohydrate, $Na_2CO_3{\cdot}H_2O$. Mild calcination then yields the anhydrous soda ash. The monohydrate process is normally more energy efficient. Because trona has low solubility, the dissolving and separation steps in the sesquicarbonate process are at high temperature. The higher temperature and CO_2 concentration make the process fluids more corrosive. The disadvantage of the monohydrate process is contamination of the carbonate solution with organics and minerals. A purge of the crystallizer mother liquor is necessary, its size dependent on the extent of contamination.

Table 26.6 World production and reserves of natural soda ash in 2010[a]

Country	Production[b]	Reserves[c]
United States	10,000	23,000,000
Turkey	1,000	200,000
Botswana	250	400,000
Kenya	450	7,000
Mexico	–	200,000
Others	–	280,000
World total (rounded)	11,700	24,000,000

[a] Quantities are estimates in thousands of metric tons
[b] World production of synthetic soda ash was 34,300,000 metric tons
[c] Sodium carbonate basis; about 1.8 t of trona yield 1 t of Na_2CO_3
Source: United States Geological Survey, Mineral Commodity Summary, *Soda Ash*, 2011

Table 26.7 Soda ash market distribution—USA in 2010

Segment	% of total
Glass	46
Chemicals	29
Soap and Detergents	10
Distributors	6
Flue gas desulfurization	3
Pulp and paper	2
Water treatment	1
Miscellaneous	3

Source: United States Geological Survey, Mineral Commodity Summary, *Soda Ash*, 2011.

Mechanical mining of trona resembles that of coal or salt. The ore is disposed in nearly horizontal beds 2.5–5 m thick and found at depths of 250 to more than 1,000 m [7]. Solution mining is an alternative. The solubility of the bicarbonate can be enhanced by using a dilute solution of NaOH. This converts the bicarbonate portion of trona into the more soluble carbonate. Production of a 30% solution of Na_2CO_3 theoretically requires about a 6% solution of NaOH [9].

Figure 26.1 is a block flow sheet of the ammonia-soda process. The description that follows is keyed to the numbered blocks. NaCl enters the process as a solution, or brine. It must first have been purified to remove calcium and magnesium ions, so that they will not precipitate within the process, contaminating the product and fouling the equipment. Solutions of soda ash and caustic soda are added to the brine, precipitating $CaCO_3$ and $Mg(OH)_2$ according to the equations noted in the preceding section. The solids are flocculated, removed in a clarifier, and discarded. The purified brine goes to the ammonia absorber (1) in the main process.

The other major raw material is limestone. Normally, the best available limestone is calcined in vertical-shaft kilns (2) at high temperature to produce lime and carbon dioxide. Kilns typically are fired with foundry coke. Their design is highly efficient and produces a gas containing very little oxygen and more than 40% CO_2. Lime from the kiln is cooled by the entering combustion air, which is thereby preheated. Similarly, gas leaving the kiln preheats incoming limestone and fuel, again promoting energy economy. Economizing steps such as these are frequent in the Solvay process because of its normally large scale and high utility demands. The dry lime is slaked (3) before use in the process.

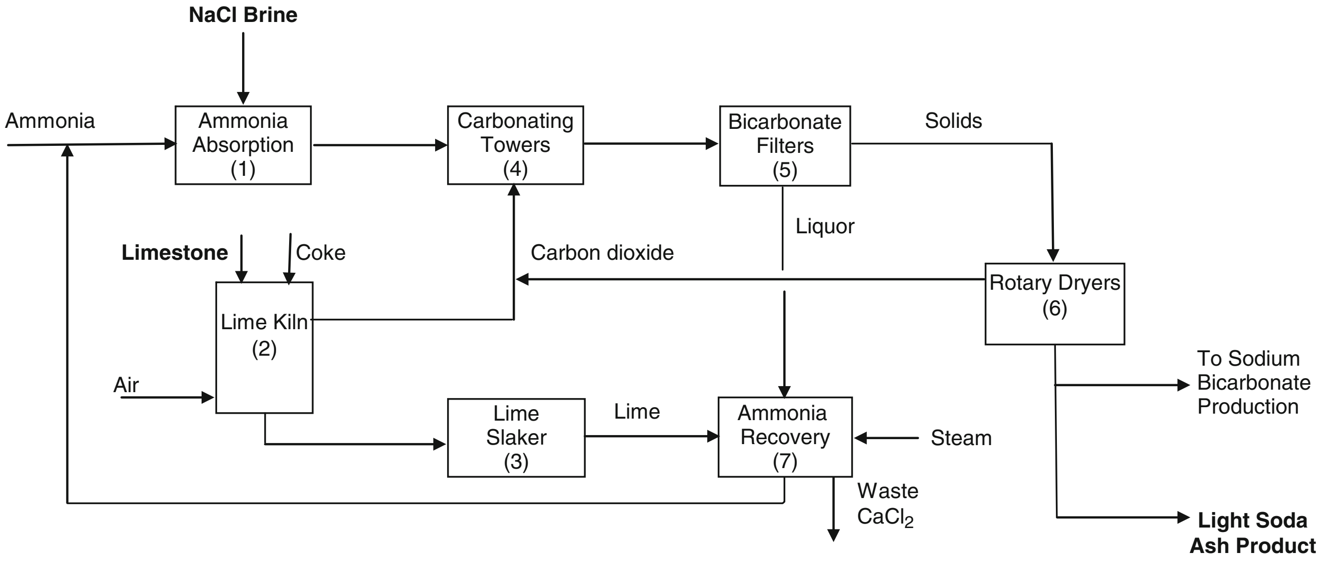


Fig. 26.1 Block flow sheet for Solvay ammonia-soda process for production of soda ash

The overall stoichiometry of the Solvay process is quite simple and involves only salt and limestone:

$$2NaCl + CaCO_3 \rightarrow Na_2CO_3 + CaCl_2$$

However, direct carbonation of NaCl solution is not feasible because of the acidity due to the implied formation of HCl from NaCl and carbonic acid. The ammoniation of the brine in step (1) creates an alkaline solution of NH_4Cl and allows the carbonation to proceed. Ammoniated brine then enters near the top of a carbonating tower (4) and cascades down over a set of disc-and-donut baffles. Carbon dioxide enters both at the bottom of the column and approximately at the midpoint. Off-gases vent from the top, and bicarbonate slurry leaves the bottom.

The carbonation area includes groups of identical towers. The alternating discs and rings promote mixing of the falling liquor with the rising gas stream and resist plugging by the crystallizing solid phase. The lower parts of the towers also include series of heat-exchange bundles. Still, several days' operation as a crystallizing unit tends to plug the internals with sodium bicarbonate. Then, ammoniated brine is passed downward through the column while gas from the kilns provides agitation and heat to dissolve the scale and bring the liquor close to saturation. After adjustment to the desired temperature, this liquor passes to the tops of the other towers in the equipment group.

Mixing 40% CO_2 gas from the kilns with 90% gas from the calcination of bicarbonate gives 60–75% CO_2 as feed to the crystallizing units. The heat evolved from absorption and neutralization of carbonic acid gas and the crystallization of bicarbonate tend to raise the temperature of the liquid in the column by about 25°C. In normal operation, the temperature of the discharged slurry is held at about 27°C by flow of water through cooling tubes. The heat removed from the carbonator is about 75,000 kcal/net metric ton of ash produced, whereas that removed from the crystallizing units is 350,000 kcal/t (all "tons" in this chapter are metric).

Washed bicarbonate is recovered from reactor slurry on rotary vacuum filters (5). The filtrate and wash liquor flow to the distillation unit for recovery of ammonia. The crude solid contains 3–4% NH_4HCO_3 (dry basis) and so is not an item of commerce. It will assay about 80% $NaHCO_3$ and have a water content of 12–15% $NaHCO_3$. Its calcination produces the carbonate while releasing half of the CO_2 content:

$$2NaHCO_3 \rightarrow Na_2CO_3 + H_2O + CO_2$$

This is achieved in rotary dryers (6) with rotating seals and gas-tight feed and discharge mechanisms. The design excludes air and keeps the concentration of CO_2 in the gas as high as 90%. The product of the reaction is about 99.7% Na_2CO_3, containing 0.1–0.15% each of NaCl and water, along with small concentrations of metallic impurities. The heat consumed in the process is about 560,000 kcal/t of soda ash produced.

The second major product of the carbonation of ammoniated brine is ammonium chloride, NH_4Cl. It collects in the liquor from the bicarbonate filters. Ammonia is regenerated (7) by reacting filtrate with milk of lime. The reaction is

$$Ca(OH)_2 + 2NH_4Cl \rightarrow 2NH_3 + CaCl_2 + 2H_2O$$

The resulting solution goes to the top of a distillation column, with steam injected at the bottom. The stripped ammonia is ready for recycle to the process, and the residual calcium chloride liquor normally contains less than 10 ppm NH_3.

The amount of ammonia in circulation is large, and there are a number of gas streams that contain NH_3. Several of these are scrubbed to prevent losses. Thus, the incoming purified brine first washes the ammonia-bearing air pulled through the bicarbonate cake on the vacuum filters. This is accomplished in a packed absorber. The brine then passes through a second packed section, where it absorbs the NH_3 in the gases from the carbonating towers, and finally flows to the main NH_3 absorber. A small amount of makeup ammonia enters the process here, and the absorber also picks up CO_2 and traces of H_2S from the carbonators. Heat released in the absorption process is about 345,000 kcal/t of soda ash produced. It is removed by circulating brine through water-cooled heat exchangers. The ammoniated brine typically is at a temperature of 38°C and contains 260 g/L NaCl, 90 g/L NH_3, 40 g/L CO_2, and 0.1 g/L H_2S.

It is worthy of note that sulfide will maintain a protective film of iron sulfide on the cast iron equipment and minimize contamination of the crystals by iron corrosion products. A solution containing sulfide ion therefore is usually added to the liquor before the ammonia distillation pre-heater. After preheating, the process stream enters a stripper where excess CO_2 is removed.

Byproduct $CaCl_2$ may be recovered from process waste liquor by evaporation. Most often, the solution is clarified and pumped into watercourses whose natural flow is sufficient to provide the dilution needed for disposal.

The normal product is "light" soda ash. This can be reprocessed to give higher bulk density, less tendency to dust, and better flow properties. The latter product is preferred in some applications (e.g., glassmaking). Some soda ash also is used to produce sodium bicarbonate (*infra*) and, by causticization with lime, caustic soda without chlorine as a co-product.

The natural alkaline brines referred to at the beginning of this section are a relatively small source of Na_2CO_3. It is recovered by a process of fractional crystallization that also produces other sodium and potassium salts. Formation of Na_2CO_3 may be promoted by carbonation of the brine.

Sodium Bicarbonate

Further carbonation of a saturated solution of sodium carbonate produces the bicarbonate:

$$Na_2CO_3 + CO_2 + H_2O \rightarrow 2NaHCO_3$$

The Na_2CO_3 may be produced synthetically or mined as a natural ore (*supra*). The relatively insoluble bicarbonate crystals precipitate from solution and are centrifuged and then dried. Drying temperatures below 50°C prevent reversion to carbonate.

Sodium bicarbonate also can be obtained by solution mining of the mineral nahcolite. Wells are drilled in groups or in pairs. Steam or hot water (175–215°C) is pumped down one well to dissolve the ore, and the resulting solution is pumped up from another. The solid product may be obtained by crystallization. In some circumstances, such as those in which the mined solution must be transferred a long distance for processing, direct recovery may not be feasible. The positive temperature coefficient of solubility of $NaHCO_3$ would allow plugging of piping or equipment. In such cases, the bicarbonate first is converted to carbonate, which is reprocessed at the destination.

Nahcolite is often associated with oil shale or with deposits of salt or trona [9]. Reserves in the Piceance Creek basin in northwestern Colorado alone amount to nearly 30 billion tons [10]. This amount would supply the world for about 20,000 years. Total US capacity for $NaHCO_3$ is about 780,000 t/year, two-thirds of it in Wyoming and Colorado. Current production is about 600,000 t/year, with a long-term growth rate of 2–3%/year [11].

As the common name "baking soda" implies, most applications of $NaHCO_3$ (~60%) are in the food industry. The product must be purified as necessary and handled in accordance with all food-grade rules. Refined product accordingly has some use in pharmaceuticals. Sodium bicarbonate's use as a rising agent in baked goods derives from its alkalinity and the fact that CO_2 evolves at the convenient temperature of 70°C. It is also used as a buffer, an antacid, and a cattle feed supplement/rumen pH regulator. A very successful application in the household market is as a deodorizer and air freshener. It has also established a market as a paint remover, replacing the environmentally unacceptable methylene chloride. Some traditional applications now are shared with the sesquicarbonate, and this change should perhaps not be viewed as a true loss of market.

Sodium Sulfate

Sodium sulfate classically was a co-product (with HCl) of the reaction of salt with sulfuric acid (Mannheim process) or SO_2 and air (Hargreaves process). While these processes may survive elsewhere, they are no longer in use in the United States. Instead, all production is by recovery from natural sources or as a byproduct of other operations. US production is approximately evenly split between these two categories [12]. Estimates of worldwide production in 2010 are six million tons from natural sources and 1.5–2 million tons as byproduct. Known reserves, shown in Table 26.8, are widespread and would last hundreds of years at the present rate of consumption. Reserves in China are extensive. Production is concentrated in Jiangsu Province, where annual production of 4.8 million tons is anticipated by 2013 [12].

There are two producers of natural sodium sulfate in the United States, one in Texas and one in California. The former location has the advantage that the brine is located near salt domes and natural gas wells. Byproduct Na_2SO_4 is recovered in 11 plants in 9 states. Sources include battery reclamation and the production of ascorbic acid, cellulose, chromium chemicals, rayon, resorcinol, and silica pigments. Approximate consumption by end use is summarized in Table 26.9.

Table 26.8 World reserves of sodium sulfate

Country	Reserves[a]
United States	860
Spain	180
Mexico	170
Turkey	100
Canada	84
Others[b]	1900
World total	3300

[a] Millions of metric tons

[b] Principally Botswana, China, Egypt, Italy, Mongolia, Romania, and South Africa

Source: US Geological Survey, Mineral Commodity Summary, *Sodium Sulfate*, 2011

Table 26.9 Sodium sulfate market distribution—USA in 2010

Market sector	% of total
Soap and detergents	35
Glass	18
Pulp and paper	15
Textiles	4
Carpet fresheners	4
Others	24

Source: US Geological Survey, Mineral Commodity Summary, *Sodium Sulfate*, 2011.

A frequent advantage of natural products over synthetic and byproduct materials is their higher purity. In one process, natural brine containing about 10% Na_2SO_4 is saturated with NaCl by pumping it through a salt deposit. Cooling the concentrated brine to about −10°C produces quite pure crystals of Glauber's salt ($Na_2SO_4 \cdot 10H_2O$). This is then melted and the water of hydration removed by evaporation driven by a submerged burner. The wet Na_2SO_4 is dried in a rotary kiln until anhydrous. More than 10 t of natural brine and 450,000 kcal are required per ton of product. In a similar process, a different brine is first carbonated and chilled to remove sodium carbonate and borax. Further chilling crystallizes Glauber's salt and some residual borax. After recovery by hydraulic classification, the sulfate fraction is handled as described above to recover the anhydrous salt.

An alternative process when carbonate is also present in a brine (e.g., Searles Lake) is evaporation to produce NaCl and a double salt of Na_2CO_3 and Na_2SO_4. The two solids are separated in a hydraulic classifier, and the sulfate fraction is re-dissolved and re-crystallized as Glauber's salt. Anhydrous Na_2SO_4 can be recovered without the need for thermal decomposition by adding the Glauber's salt to concentrated NaCl brine. The altered solubility characteristic forces the unhydrated sulfate to precipitate. In the first process mentioned above, it was Glauber's salt that precipitated from solution; in the final step here, it is anhydrous Na_2SO_4. The difference is explained by different positions on the graph of mutual solubilities of NaCl and Na_2SO_4. The first process takes place in the low-chloride region, and the process follows a Glauber's salt isotherm. Addition of small amounts of water or salt may be necessary to the material balance in a continuous crystallization process.

Sodium sulfate often is a byproduct of the reaction of sulfuric acid with a sodium salt or of the neutralization of caustic soda. An example is the viscose rayon process, in which cellulose xanthate reacts with H_2SO_4, which also neutralizes free NaOH in the fiber spinning bath. Evaporation of spin-bath liquor produces crystals of Glauber's salt, which are centrifuged off, melted and evaporated to produce anhydrous Na_2SO_4. The sulfate produced actually outweighs the rayon product by 10%.

Changes in markets or regulation can affect the available supply of byproduct Na_2SO_4. Thus, for example, increased on-site generation of chlorine dioxide for bleaching itself produces more Na_2SO_4 byproduct. Opposing this change, the production of sodium dichromate and its byproduct sulfate has declined with mandated recycling of electroplating wastes. Other changes affect demand. The kraft pulping process formerly consumed huge quantities of Na_2SO_4. Changes in the process and tightening of the sulfur balance have caused a major reduction in its use. Next, the use of Na_2SO_4 as an inert filler in detergent formulations has decreased as liquid detergents have become more common and dry products have become more concentrated. Finally, sulfate also is gradually disappearing from glass manufacture, where it has been used as a fluxing and fining agent. The consumption of Na_2SO_4 in the United States therefore is much lower than it was in previous decades. As Table 26.10 shows, it has been dropping irregularly in recent years. Variations have been much smaller in production than in consumption, the difference being picked up by a shifting trade balance. Worldwide production of sodium sulfate continued to increase into the 1990s before similar trends prevailed and total consumption became stagnant [13].

Table 26.10 US production and consumption of sodium sulfate (natural and synthetic)

	2007	2008	2009	2010(est.)
Production	312	319	292	300
Imports for consumption	43	69	77	60
Exports	101	107	140	190
Apparent consumption[a]	254	281	229	170

All data in thousands of metric tons

[a] Production plus imports minus exports

Source: US Geological Survey, Mineral Commodity Summary, *Sodium Sulfate*, 2011

Sodium Sulfides

Sodium sulfide, Na_2S, and sodium hydrosulfide or sulfhydrate, NaHS (often written as NaSH), both are derived from caustic soda and H_2S. Controlled caustic scrubbing of H_2S forms a solution of the half-neutralized NaHS:

$$NaOH + H_2S \rightarrow NaHS + H_2O$$

The NaHS solution is filtered to remove heavy metal sulfides. The clear filtrate may be sold as a 44–46% solution or evaporated in stainless steel equipment to produce a flaked product containing 70–72% NaHS. There is ample capacity for generation of NaHS from waste gases. Approximate capacity in the United States in 2006 was 240,000 t/year, but production was only about 77,000 t [14].

Treating NaHS solution of the proper concentration with solid caustic soda yields a hydrated product of 60–62% Na_2S:

$$NaHS + NaOH \rightarrow Na_2S + H_2O$$

This is sold as flakes or cast solid in drums. High-quality raw materials produce Na_2S suitable for use in dyes, photography, rayon, and leather manufacture. Lower-quality sulfides are obtained by using H_2S byproduct from the manufacture of CS_2 from methane (or other low molecular weight hydrocarbons) and sulfur:

$$CH_4 + 4S \rightarrow CS_2 + 2H_2S$$

The gas from this catalytic reaction is cooled and then scrubbed with caustic soda. The sulfide produced contains small amounts of mercaptans (e.g., CH_3SH), and thus has a very objectionable odor.

The manufacture of $BaCO_3$ from barite ore also yields a sodium sulfide byproduct. The process involves roasting with coal, leaching with water, and treatment with soda ash:

$$BaSO_4 + 4C \rightarrow BaS + 4CO$$

$$BaS + Na_2CO_3 \rightarrow Na_2S + BaCO_3$$

The principal use for sodium sulfide in the United States is de-hairing leather before tanning. In Western Europe it is further used in tanning and in Japan more extensively in synthetic chemistry. Most NaHS in the United States is used in the pulp and paper industry (42%). The market has grown at the expense of the more expensive sodium sulfate [14]. Other uses are in metals and minerals, most especially as a copper mineral flotation agent (30%), and in chemicals and dyes (13%). Leather processing (10%) is a declining market. Again in Japan, NaHS historically has found more downstream uses in chemicals, dyes, wastewater treatment, and leather tanning [15].

A solution of sodium sulfide can dissolve elemental sulfur, which then forms covalently linked chains of sulfur atoms and polysulfide ions S_n^{2-}. Chain lengths vary from three to more than ten [16]. One of the uses of sodium polysulfide is the production of polymers that are insoluble in water, oils, and most organic solvents. The polymers are used widely as sealants. They form by condensation with bifunctional derivatives of hydrocarbons. Ethylene dichloride (EDC) is an example:

$$Na_2S_n + ClCH_2CH_2Cl \rightarrow (CH_2CH_2S_n) + 2NaCl$$

Sodium Thiosulfate

Most sodium thiosulfate, $Na_2S_2O_3$, is obtained as a byproduct in the manufacture of sulfur dyes and Na_2S. In the case of sulfur dyes, byproduct leach liquor contains thiosulfate, which can be recovered by evaporation and crystallization in stainless steel equipment. Much of this is sold as the pentahydrate. Of less importance today is the older soda ash-sulfur process. A liquor containing $NaHSO_3$ (made by absorbing SO_2 into soda ash solution) is heated with powdered sulfur in an agitated stainless steel digestion tank. Evaporative crystallization again yields the pentahydrate. After centrifuging, washing, and drying, this material is packaged in airtight containers, in order to prevent efflorescence.

$Na_2S_2O_3{\cdot}5H_2O$ is the photographer's "hypo." While many photographic markets are declining or growing only slowly because of the advance of digital technology, this remains the principal use of the product. Its value lies in its ability to dissolve water-insoluble silver salts. Other uses include leather processing, water treatment, and oil recovery. It is also used as an active chlorine scavenger, particularly in Western Europe. It has the advantage of an average sulfur valence state of only +2, giving it twice the reducing power of other sulfur-based agents. However, other reactions can occur at low pH [17], and $Na_2S_2O_3$ is fully effective only at pH 11 and higher [18].

Sodium Sulfite

Sodium sulfite (Na_2SO_3) is produced by reacting sulfur dioxide with an alkaline solution, either sodium carbonate or sodium hydroxide. It is also recovered as a byproduct of various processes. Its principal use historically has been in the pulp and paper industry. While improvements in mill efficiency and substantial reductions in sulfur discharges have tended to reduce its use there, the spread of chemithermomechanical processing from Europe has helped to keep this market strong. The pulp industry accounts for about 55% of the market in the United States [19]. Other large uses are in water treatment (20%), the photographic industry (10%), and oil recovery (5%).

Sodium sulfite is a white crystalline or powdered compound. It is readily soluble in water and decomposes on heating. It is a strong reducing agent and so is subject to oxidation; the solid compound is more stable than solutions in this regard. Na_2SO_3 is used in water treatment as an oxygen and free chlorine scavenger and in the photographic industry to protect developer solutions from oxidation. Miscellaneous applications are in the textile industry as a bleaching, desulfurizing, and dechlorinating agent; in the leather trade in the sulfitization of tanning extracts; and in the organic chemical industry as a sulfonation and sulfomethylation agent. It forms adducts with aldehydes and ketones and so plays a role in purifying and isolating those compounds. Other applications include ore flotation, oil recovery, food preservation, dye making, detergent production, and the production of sodium thiosulfate (*supra*).

Yearly production in the United States is about 90,000 t, with an installed capacity of about 245,000 t/year. These figures are inexact largely because they exclude the substantial and variable captive production and use [19].

Sodium Bisulfite

Sodium bisulfite may be supplied as a solution or as a solid. The solution normally contains 23–27 wt.% SO_2 equivalent. The solid product actually is in the form of sodium metabisulfite, which is the anhydride $Na_2S_2O_5$. Commercial metabisulfite is about 98% pure, typically containing 1.5% Na_2SO_3 and 0.5% Na_2SO_4.

Methods of production are all variations on the same theme:

$$Na_2CO_3 + 2SO_2 \rightarrow Na_2S_2O_5 + CO_2$$

In one version, SO_2 is sparged into a stainless steel absorber through which a solution of Na_2CO_3 flows. The product crystallizes when the saturated solution is cooled. Recovery is by centrifugation, and then the crystals are dried. The last step must be rapid, as in a flash dryer, in order to avoid air oxidation to sulfate.

Markets for sodium bisulfite are diverse [20]. They include chemical intermediates, pharmaceuticals, food preservation, water treatment, and dye fixing. Additionally, sodium bisulfite is used in photography, in the manufacture of sodium hydrosulfite (*infra*), and as an active chlorine scavenger in pulp bleaching and chlor-alkali brine dechlorination.

Sodium sulfite and bisulfite are interchangeable in many applications, and some producing plants can switch between the two forms. Their reducing power (and that of SO_2) in all cases depends on the oxidation of S^{+4} to S^{+6}. Proton shifts between HSO_3^- and SO_3^{2-} are rapid [21], and so the form of the anion actually present in solution is determined by the pH.

Sodium Hydrosulfite (Sodium Hyposulfite)

$Na_2S_2O_4$ is variously referred to as sodium hydrosulfite, sodium hyposulfite, and dithionite. It is a powerful reducing agent used principally for the reduction of vat dyes and the brightening of mechanical wood pulps and clays. There are several methods for its production, all of which involve reduction of SO_2 or $NaHSO_3$. Reducing agents include sodium amalgam, sodium formate, and sodium borohydride. Previous editions of this handbook also described two methods for reduction by zinc powder. These are obsolescent because of the environmental problems associated with the release of zinc.

The amalgam process [22] uses the dilute (<0.5%) sodium amalgam from a chlorine cell (*infra*). Reaction of the amalgam with bisulfite solution at pH 5–7 produces the hydrosulfite along with a sodium sulfite co-product:

$$4NaHSO_3 + 2Na(Hg) \rightarrow Na_2S_2O_4 + 2Na_2SO_3 + 2H_2O + (Hg)$$

The parentheses indicate that mercury is not involved in stoichiometric quantities. The mercury returns to the electrolytic cell to form more amalgam. The sulfite is reacted with SO_2 in aqueous solution to regenerate bisulfite for recycle:

$$Na_2SO_3 + SO_2 + H_2O \rightarrow 2NaHSO_3$$

SO_2 normally supplies the sulfur value in the formate process:

$$HCOONa + NaOH + 2SO_2 \rightarrow Na_2S_2O_4 + H_2O + CO_2$$

This reaction requires the use of acidified methanol. The anhydrous sodium hydrosulfite precipitates and is recovered by filtration.

In pulp and paper mills, the hydrosulfite often is produced in situ at low concentrations (5–7%). Here, SO_2 reacts with alkaline sodium borohydride:

$$\begin{aligned} &(NaBH_4 \cdot 3.4NaOH) + 5.6NaOH + 9SO_2 \rightarrow \\ &\qquad \text{(borol solution)} \\ &\qquad 4Na_2S_2O_4 + NaBO_2 + NaHSO_3 + 6H_2O \end{aligned}$$

Finally, hydrosulfite can be produced by the direct electrolysis of a bisulfite solution [23]. The cathode reaction is:

$$2HSO_3^- + 2H^+ + 2e \rightarrow S_2O_4^{2-} + 2H_2O$$

Such a process might be useful for on-site production.

Sodium Silicates

In the sodium silicate family, a series of derivatives is produced by reacting Na_2CO_3 with varying amounts of silica. The products normally are referred to by their ratios of SiO_2 to Na_2O. The custom in the United States, followed here, is to express these in terms of weight. In some other parts of the world, molar ratios are used. Since the formula weights of the two species are nearly equal (60 and 62), the numerical differences are not large but still may confuse the unwary. The normal product is called *water glass* because while solid it actually is a glass, but, unlike lime-soda glass, it is soluble in water.

In the standard process, high-purity sand and soda ash are fused in a large tank furnace similar to the type used for flat glass. The raw materials, often premixed, are introduced at one end of the furnace and the product

Table 26.11 Properties of typical commercial sodium silicate solutions

Ratio, SiO_2/ Na_2O	Wt. % SiO_2	Wt. % Na_2O	Density, g/ mL	Viscosity, cp	pH
3.22	28.7	8.9	1.38	180	11.3
3.22	29.5	9.1	1.41	400	11.3
2.88	31.7	11.0	1.48	960	11.5
2.58	32.1	12.4	1.52	780	11.8
2.40	33.2	13.8	1.56	2100	12.3
2.00	29.4	14.7	1.53	400	12.8
1.80	23.7	13.1	1.42	655	13.1
1.60	26.2	16.4	1.53	280	13.5

Ratio ±0.05; $\%SiO_2 \pm 1.0$; $\%Na_2O \pm 0.2$. Physical properties at 20°C
Source: PQ Corporation, Bulletin 17-2A (2007)

drawn off at the other. The product ratio is set by the raw material mixture:

$$Na_2CO_3 + nSiO_2 \rightarrow Na_2O(SiO_2)_n + CO_2$$

The reaction temperature is about 1,300°C. Large furnaces are regenerative or recuperative, in order to reduce heat loss by way of combustion offgas. Particularly on smaller scale, electrically heated melters may replace the fuel-fired version. Another variation has been the use of a rotary kiln in place of a stationary furnace. As the melt leaves the furnace, a stream of cold water shatters it into fragments, or it is collected and cooled in molds that form a moving chain. The solid material can be dissolved in hot water/superheated steam in tall steel cylinders. Sodium silicates normally are sold in solutions of varying density and viscosity. Compositions and physical properties of typical commercial grades are shown in Table 26.11. Dry solids also are marketed, usually with ratios between 2.0 and 3.25 [24]. Since some solid silicates are hygroscopic, they may be blended with Na_2SO_4 to prevent caking.

Caustic soda is a possible replacement for soda ash in the silicate process:

$$2NaOH + nSiO_2 \rightarrow Na_2O(SiO_2)_n + H_2O$$

Those who can use either NaOH or Na_2CO_3 as a source of alkalinity must continually be aware of the comparative economics of the two, and it is useful to have the option to switch materials when desired. Unfortunately, the use of caustic soda, solid or liquid, is not practicable in the glass furnace. Instead, standard practice is to use an autoclave to dissolve sand in NaOH solution. Results are limited by the solubility of crystalline silica to ratios less than about 2.4.

The buffering capacity of silicate solutions is important to the many applications that depend on their alkalinity (e.g., detergents, oxide bleaching of wood pulp, de-inking of waste paper). Figure 26.2 shows that silicate solutions are superior in this respect to the other salts shown. The increased use of peroxide bleaching of wood pulp and the growing importance of waste paper de-inking have increased sodium silicate demand in the pulp and paper industry. Silicates sequester iron and other metal ions that reduce the efficiency of the peroxide. They also buffer the pH in the alkaline range required for bleaching. In the de-inking process, in addition to its buffering and sequestering action, silicate provides detergency and helps to prevent redeposition.

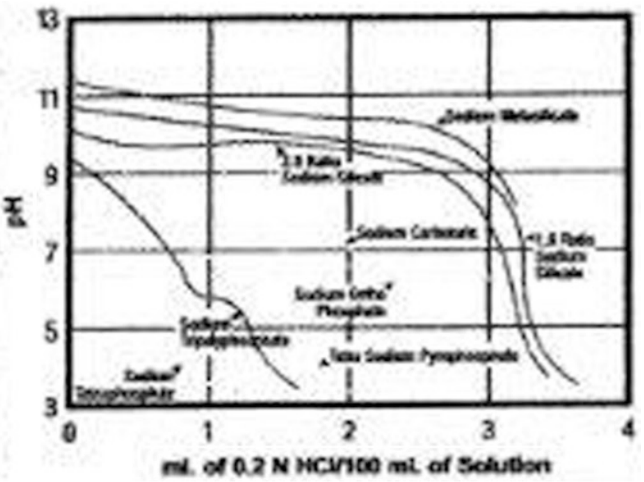


Fig. 26.2 Alkaline buffering capacities of several compounds (Courtesy of PQ Corporation)

While the uses of dissolved silicates are extraordinarily diverse, the most rapid growth in the sodium silicate market is in custom derivatives such as precipitated silicas and zeolite catalysts. Processing of silicate solutions with mineral acids yields hydrated silica, which forms silica gel when dried. Silica gel finds wide application as an adsorbent and is used in the manufacture of catalysts (petroleum cracking and hydrocracking, hydrocarbon isomerization, etc.). Modified processing can yield silica sols, which are useful in water treatment. Zeolites primarily are aluminosilicates, made in many different grades. Properties depend on SiO_2/Al_2O_3 ratio, particle size, and cell structure. Varying the positive counterion and incorporating different metals into the structure give a wide variety of catalytic functions.

Other specialized products are those of higher alkalinity, such as the metasilicate, Na_2SiO_3 (molar ratio 1.0), and the orthosilicate, Na_4SiO_4 (molar ratio 0.5). Such high alkalinity would be corrosive to normal furnace refractories, and so these products are made in special apparatus by adding caustic soda to sil4icates of higher ratio [25]. The metasilicate is supplied in the anhydrous form and as the pentahydrate. Solid products with the SiO_2/Na_2O ratio of orthosilicate may be supplied as the compound or as a mixture of solid metasilicate and anhydrous NaOH beads.

Chlor-Alkali (Chlorine and Caustic Soda)

The most important use of salt in the chemical industry is the production of chlorine and caustic soda (sodium hydroxide, NaOH). Each is among the world's ten largest volume commodity chemicals. In 2009, combined production in the United States was 21.5 million metric tons [26]. This may have been a recessional anomaly, as in each of the preceding 5 years the total was between 22.8 and 24.7 million tons. Global production of chlorine alone in 2008 was about 42 million metric tons, distributed by region as shown in Table 26.12 [27].

Historically, electrolysis of NaCl brine accounts for about 96% of all chlorine production. It consumes about 2% of the electrical power generated in much of the industrial world. One of the difficulties faced by chlorine/caustic soda merchants is the fixed ratio of the amounts of the two products that are generated. Theoretically, nearly 1.13 kg of NaOH accompany each kilogram of chlorine. In the United States, the ratio in the electrolytic industry has hovered around 1.05–1.06. The same ratio applies to the data in Table 26.12. This reflects the fact that some plants produce chlorine without NaOH. Co-products may be Na, Mg, or KOH. There is also some production of chlorine by electrolysis of HCl. At the same time, there is some production of caustic soda without chlorine by reaction of sodium carbonate with lime. These alternative processes are described in standard texts [28, 29].

The chlor-alkali industry in the United States has been consolidating. Some merchants who supplied the pulp and paper industry, for instance, have left the business. Growth has been in the form of large plants integrated forward to produce EDC. The number of producers has been declining. Table 26.13 shows that four of them accounted for more than 80% of the capacity in 2010. Applications of caustic soda are quite diverse. Table 26.14, in which 30% is assigned to the very broad category of "organics," confirms this. The uses of chlorine historically were just as diverse, but social and environmental pressures have removed certain market segments (e.g., pulp bleaching, some chlorinated solvents). As a result, the chlorine market depends more heavily on vinyl chloride monomer (VCM) and its polymers. Table 26.15 shows 43% of the American market in this single use [27].

The heart of the chlor-alkali process is an electrolytic cell in which a nearly saturated solution of purified NaCl is decomposed. The three types of cell now in use are shown schematically in Fig. 26.3. The anode reaction is the same in every case:

$$2\mathrm{Cl}^- \rightarrow \mathrm{Cl}_2 + 2e$$

Simple discharge of the sodium ion cannot occur in an aqueous medium; the favored reaction instead would be

$$2\mathrm{H}_2\mathrm{O} + 2e \rightarrow \mathrm{H}_2 + 2\mathrm{OH}^-$$

Chlorine and hydroxide ion would react immediately to form chloride and hypochlorite if the anolyte and catholyte were

Table 26.12 Worldwide production data for caustic soda and chlorine

Area	Caustic soda production[a]	Chlorine production[a]
Asia	15.7	14.8
Europe	12.2	11.5
North America	11.5	10.9
South America	1.8	1.7
Others[b]	3.6	3.4
Total	45.0	42.3

[a] Millions of metric tons; 2008

[b] Includes India, Middle East, etc.

Courtesy of Consulting Resources Corporation

Table 26.13 Leading producers of chlorine in the United States

Company	% of total capacity
Dow	31.8
Occidental	23.3
PPG	13.6
Olin	13.5
Formosa Plastics	5.9
Georgia Gulf	3.1
Others	8.8

Source: The Chlorine Institute, Inc. [26]

Table 26.14 Caustic soda end use pattern in USA (Courtesy of Consulting Resources Corporation)

Market segment	% of total
Organics	30
Inorganics	19
Pulp and paper	12
Petroleum	6
Soap and detergents	5
Miscellaneous	13
Export	15

Table 26.15 Chlorine consumption pattern in USA (Courtesy of Consulting Resources Corporation)

Market segment	% of total
EDC/VCM	43
Other organics	29
Inorganics	12
Water treatment	4
Miscellaneous	12

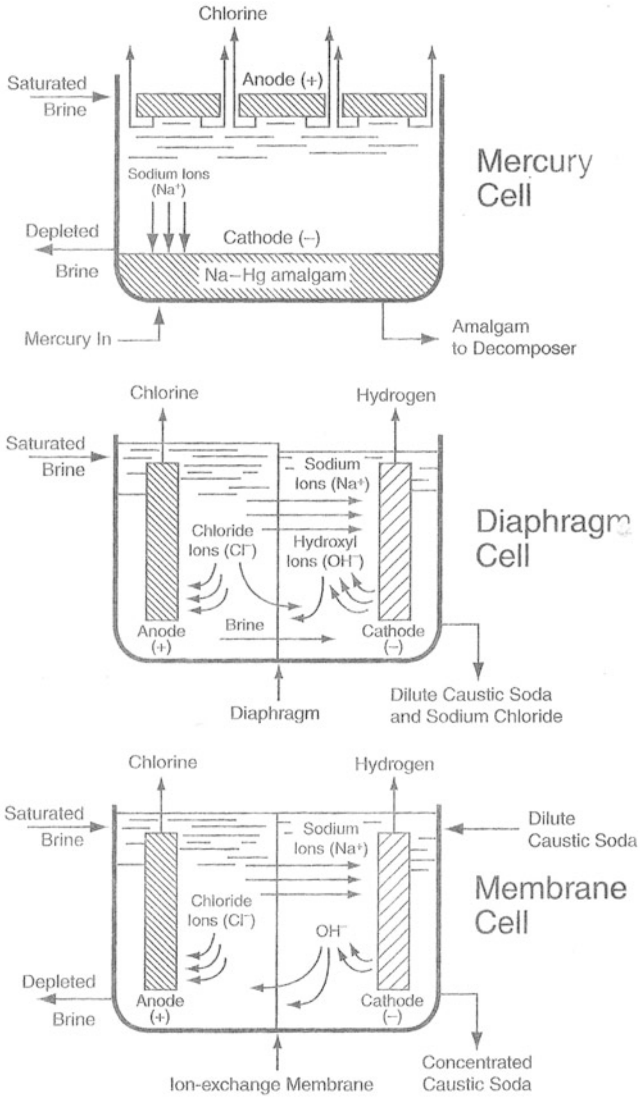


Fig. 26.3 Schematics of mercury, diaphragm, and membrane chlor-alkali cells (Reproduced from Handbook of Chlor-Alkali Technology, vol. I, p. 38. Copyright 2005 Springer Science + Media and reproduced with permission.)

allowed to mix. Chlor-alkali cells must be designed to prevent this mixing. Diaphragm and membrane cells physically separate anolyte and catholyte. Mercury cells rely on a different chemistry. The cathode there is a dilute sodium amalgam flowing over a sloped floor of carbon steel. The sodium ion discharges to Na^0, amalgamated with the mercury. The amalgam that leaves the main part of the cell then flows through a decomposer, or denuder, forming caustic soda by its reaction with purified water:

$$2Na(Hg) + 2H_2O \rightarrow 2NaOH + H_2 + (Hg)$$

The parentheses indicate that mercury is not involved in stoichiometric quantities; the amalgam is in fact quite dilute ($<0.5\%$) in sodium. The flow rate of water, when correctly proportioned to the rate of production, controls the strength of the NaOH solution. The denuded amalgam is recycled to the cell.

In the diaphragm cell, a fibrous mat separates anolyte and catholyte. The feed brine enters the anode side of the cells, and the anolyte flows through the diaphragm into the cathode side. The chlorine formed at the anode passes off as a gas. Not all the chloride ion is oxidized; practical considerations limit its conversion to about 50%. The unreacted chloride ion goes into the catholyte. Hydrogen gas and caustic soda form at the cathode. Figure 26.4 shows the components of a modern diaphragm cell.

Diaphragm-cell liquor contains roughly equal amounts of chloride and hydroxide (molar basis). This is the fundamental disadvantage of the diaphragm cell because extensive processing is required to produce a high-quality NaOH product and recover the unconverted salt for recycle. Another disadvantage, which also goes to explain the continuing gradual replacement of diaphragm cells, is that the fiber used to form the diaphragms classically has been asbestos. Modern cells use a modified form into which a fused fluoropolymer is incorporated. All such formulations are gradually disappearing, largely because of the toxicity of asbestos and increasing restrictions on its use. Synthetic diaphragms free of asbestos are available and are gradually replacing the older types [30].

The technical advantage of membrane cells over diaphragm cells lies in the selectivity of the interelectrode barrier, shown as an ion-exchange membrane in the bottom part of Fig. 26.3. Membranes are sheets of perfluorinated polymer that are designed to permit the flow of water and alkali metal ions from anolyte to catholyte and to block the flow of anions in either direction. This means that chloride ions do not pass into the catholyte to any great extent, and so their concentration in the final product caustic usually is measured in tens of parts per million. To a lesser extent, the passage of hydroxide from catholyte to anolyte also is restricted. The hydroxide flow is a few percent of the total ionic flux through a properly performing membrane. The efficiency of its rejection is determined by the nature of the polymer facing the catholyte. For best efficiency, the ion-exchange group in the polymer will be based on a carboxylic acid. Carboxylate polymers have the disadvantage of low electrical conductivity. Perfluorosulfonate polymers are more conductive but operate at lower current efficiency. The typical membrane therefore has composite construction. It may also include PTFE fibers for physical reinforcement and a surface coating that promotes the release of gas bubbles when the gaps between cell components are small. Figure 26.5 shows how these features are combined in a commercial membrane.

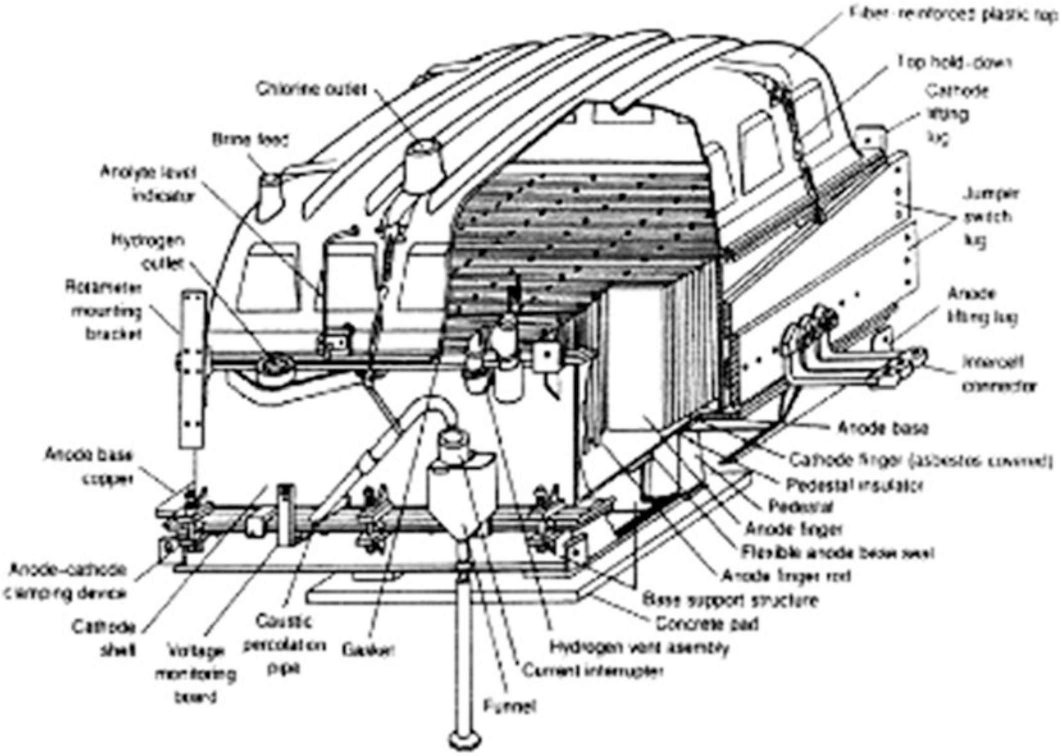


Fig. 26.4 Cut view of ELTECH diaphragm cell (Courtesy of ELTECH Systems Corporation)

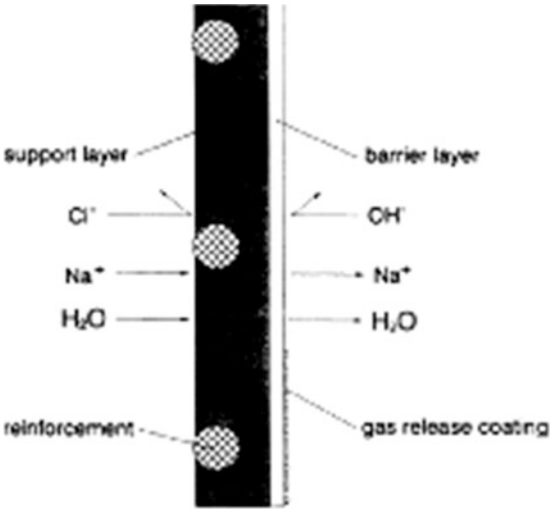


Fig. 26.5 Chlor-alkali membrane cross-section

Before the commercial advent of the membrane cell, diaphragm and mercury technologies competed primarily on the bases of energy consumption and product quality. Mercury cells consume more electrical energy, and this is their chief economic drawback. On the other hand, they operate at much higher current densities and produce caustic soda at commercial strength without extensive processing. The liquor produced in a diaphragm cell is much weaker and has a high concentration of dissolved salt. Expensive multi-effect crystallizing evaporators are required, with large steam consumption and the need to recover and recycle salt. Removal of the dissolved salt is not perfect, and 50% diaphragm-cell caustic soda typically contains 1% residual NaCl. Mercury-cell caustic, often referred to as "rayon-grade," therefore was preferred for some applications and often could be sold at a premium price. The cells are sensitive to a different set of brine impurities, i.e., those that can interfere with amalgam chemistry. Their defining disadvantage, however, is simply the presence and inevitable release of small amounts of the highly toxic mercury. Societal and regulatory pressures against the use of mercury continue to increase. The use of mercury cells has been in steady decline, and their days are numbered.

Membrane cells generally consume less electrical energy than do diaphragm cells. Evaporators are required, but they are simpler in construction and consume much less steam than their diaphragm-cell counterparts. The caustic product is somewhat less pure than mercury-cell caustic, and the stringent brine purity requirements add to the cost and complexity of the process. The advantages outweigh the disadvantages, and the membrane cell is now the technology of choice. It is the universal choice for new installations and frequently is used in conversion of plants from the older technologies. A typical installation of new bipolar membrane cells is shown in Fig. 26.6. Figure 26.7 shows the

Fig. 26.6 Membrane-cell installation at BorsodChem Zrt., Kazinbarcika, Hungary (Courtesy of Chlorine Engineers Corporation, Ltd.)

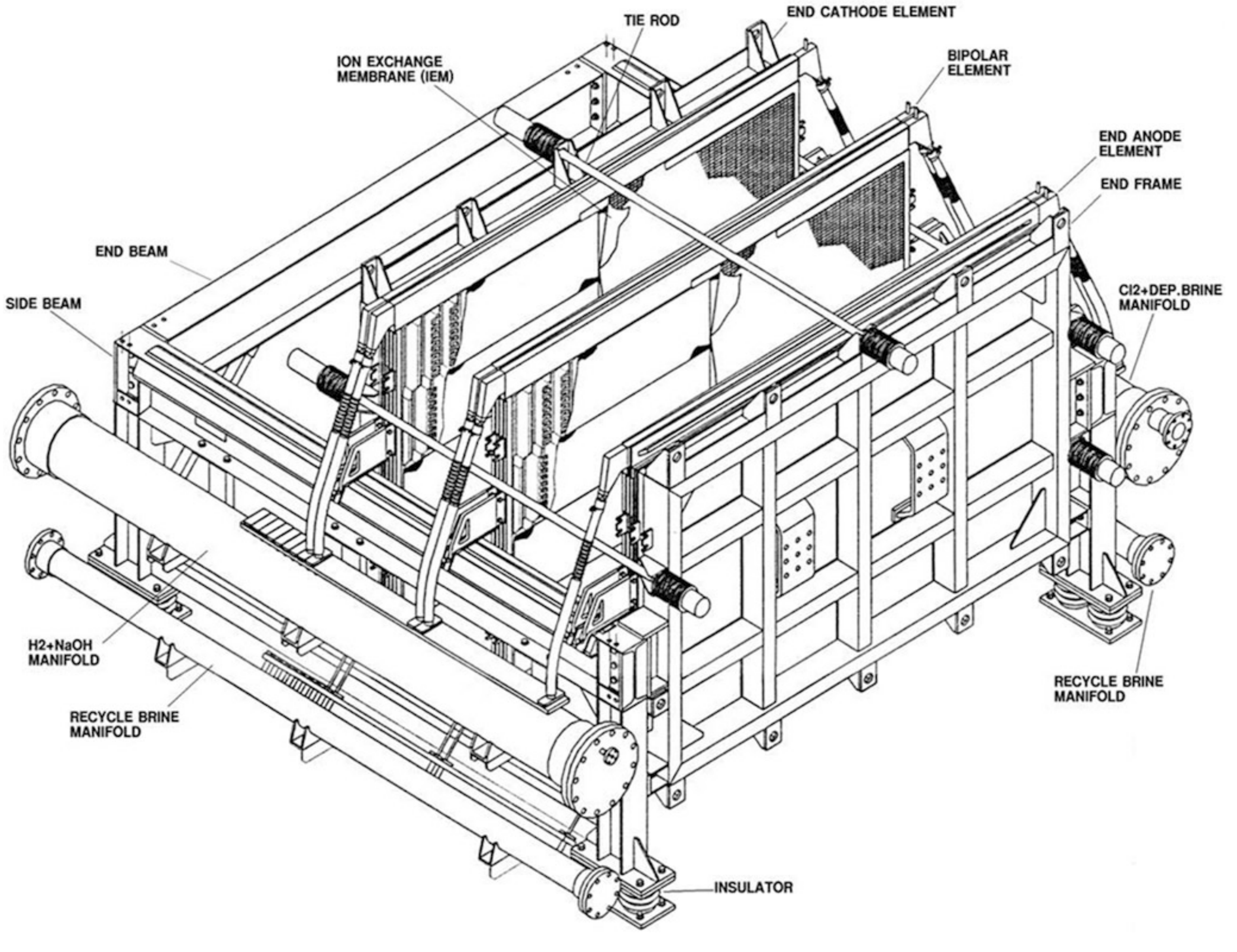


Fig. 26.7 Overview of BiTAC® bipolar membrane electrolyzer (Courtesy of Chlorine Engineers Corporation, Ltd)

assembly of a typical state-of-the-art electrolyzer. Cells such as these can operate at 3 V or less at membrane current densities of 6 kA/m^2 (Table 26.16).

The United States has resisted the change to membrane-cell technology longer than most countries. Most US capacity (>70%) is on the Gulf Coast, where several factors continue to favor diaphragm cells. There are extensive salt deposits, and solution-mined brine is readily available. Diaphragm cells can operate without penalty when salt is supplied as concentrated brine, which Table 26.5 shows to be the cheapest form. Fuel is plentiful and relatively cheap, so the large evaporative demand of the diaphragm cell is less of a liability. Also, the balance between electrical and thermal demand of a diaphragm plant is better suited to the operation of a cogeneration system. Finally, continued improvement in cell components has reduced energy consumption and increased on-line time between maintenance shutdowns [31]. In 2008, 67% of US chlorine capacity was in diaphragm cells. Membrane cells provided 26% and mercury cells, expected to phase out within the current decade, 4%. The situation in Europe was different. Mercury cells had been favored, but there, too, they are being phased out. Membrane-cell capacity became more than 50% of Europe's total in 2011 [32].

Table 26.16 Components of membrane electrolyzer voltage (Courtesy of Chlorine Engineers Corporation, Ltd)

Component	Voltage	% of total
Decomposition voltage	2.26	76.3
Anodic overpotential	0.06	2.0
Cathodic overpotential	0.07	2.4
IR losses[a]	0.57	19.3
Total	2.96[b]	100.0

[a] Include gas effects and IR drops through solution, membrane, and structure

[b] 32% NaOH, 90 °C, 6 kA/m^2

Membrane cells provide more compact installations, and conversion from one of the older technologies frequently has allowed an increase in capacity within an existing footprint. Evolution of the technology has continually increased the maximum practicable current density. Any new project requires an economic balance between the lower capital cost associated with increasing current density and the resulting higher operating voltage that increases energy cost. The recent trend in the industry is to higher current densities and little improvement in energy consumption. It is interesting to note that this parallels the history of development of the mercury cell many years ago [33].

Figure 26.8 shows the elements of the electrolyte processing required with each type of cell. In every case, NaCl brine is the major feed material. The first step in its treatment is the removal of hardness elements and heavy metals by chemical precipitation. The solids that form are removed by settling and one or two stages of filtration. The resulting brine, with a few ppm of hardness remaining, is suitable for use in diaphragm or mercury cells. Membrane

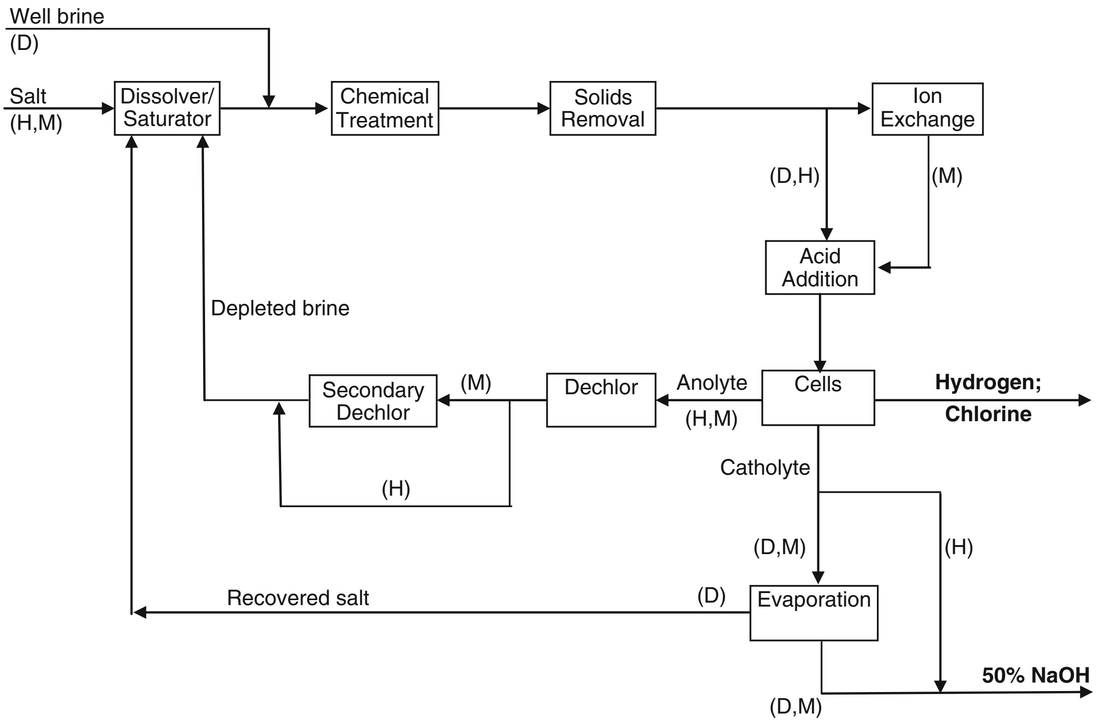


Fig. 26.8 Chlor-alkali electrolyte processing key: D, diaphragm cells; H, mercury cells; M, membrane cells

cells require additional treatment with a chelating ion-exchange resin to reduce hardness to a low ppb level. The treated brine will be alkaline and will contain dissolved CO_2. The addition of acid has several advantages including the removal of the CO_2, improving the quality of the cell gas, and scavenging of some of the OH^- that leaks back from the catholyte in some cells.

The cells produce hydrogen and chlorine gases at their electrodes. The quality of the gases does not depend greatly on the type of cell. The hydrogen is quite pure, and extensive processing is required only for special applications. At the anode, several reactions compete with chlorine formation. Impurity concentrations in the chlorine gas will amount to several percent.

The processing of hydrogen depends on its end use. It is frequently used for its fuel value. All that is required then is to deliver it at a useful pressure and with an acceptable water content. Membrane cells can operate under modest positive pressure. This at least relieves the problem of ingress of air and may also eliminate the need for compression. Other cells may require a blower. The cooling duty depends strongly on the type of cell. Latent heat being a large fraction of the total, process duty correlates with the amount of water accompanying the hydrogen. With mercury cells, the duty is quite small. Membrane cells are the intermediate case. The hydrogen is saturated with water over 30–35% NaOH. In diaphragm cells, the catholyte is more dilute in NaOH, and so the water content of the hydrogen is higher. To choose single values for comparison, with unity assigned to membrane-cell hydrogen, one might say that the cooling duty with mercury-cell hydrogen is about 0.1 and with diaphragm-cell hydrogen about 3 or 4 [34].

Chlorine may be used on site, transferred by pipeline, or liquefied for export. Broadly, about half the chlorine produced is consumed as the liquid. The full process then includes cooling, drying, compression, and liquefaction. Modern practice is to cool the gas, incidentally removing most of its water, in vertical shell-and-tube exchangers with titanium on the process side. Drying is by countercurrent contact with concentrated sulfuric acid [35]. As more water is removed during the cooling process, the cost of supply and disposal of sulfuric acid goes down. Design, however, does not involve simply an economic balance; there are technical limits on the amount of water that can safely be removed. Below 10°C, a solid hydrate of chlorine forms (approximate composition $Cl_2{\cdot}7.2H_2O$). This can deposit in equipment and cut off flow. Upon shutdown, it will melt, causing very rapid corrosion. Next, the titanium used in the coolers is suitable for service in wet chlorine because a resistant oxide film forms by hydrolysis of the metal. The film must be continuously renewed to prevent contact and rapid reaction of the metal with Cl_2. Cooling must not deplete the water content of the gas to a level where it cannot maintain the TiO_2 film.

Dry gas can be handled in ferrous-metal equipment. Centrifugal compressors are standard in large plants; liquefaction is by indirect contact with boiling refrigerant [36]. Centrifugal compressors in chlorine service are restricted in compression ratio, because temperature must be limited in order to prevent combustive reaction between the gas and the metal of the compressor. Compression then becomes a multistage process, with gas coolers operating between stages. In older and smaller plants, liquid-ring compressors are a frequent choice. The sealing liquid is concentrated sulfuric acid. These compressors have much lower energy efficiency, but the circulating acid removes much of the heat of compression, and so the compression ratio can be higher without producing excessive temperature.

The final pressure on the chlorine normally is 2–10 atm. It is determined by downstream needs if the chlorine is to be processed immediately or by a balance between refrigeration and chlorine compression costs when the gas is to be liquefied.

The impurities present in chlorine are mostly permanent gases that do not condense in the liquefaction system. They leave as a vent stream, carrying chlorine at a partial pressure equal to its vapor pressure at liquefaction temperature. The extent of liquefaction therefore increases at higher pressure and lower temperature. It, too, may be a multistage operation. Since deeper liquefaction requires higher severity, energy can be saved by liquefying most of the chlorine under relatively mild conditions and then increasing the severity to reach the desired recovery.

The presence of chlorine in the uncondensed tail gas is a problem that must be dealt with. It can be destroyed by conversion to an innocuous waste, converted to another useful product, or recovered as the element. There are many processes for treating tail gas. The paper by Silver [37] remains a good summary. Destruction most often is by scrubbing with an alkaline liquor. This absorbs the chlorine from the gas and first converts it to hypochlorite, as in the equation above. Some of this will react to form chlorate (*infra*), and the rest can be chemically or catalytically reduced to chloride.

Controlled absorption with precautions not to decompose the hypochlorite is one example of conversion to a useful product. Production of hypochlorite bleach in this way is discussed below. Another option is the production of hydrochloric acid (*infra*). This can be an especially attractive operation as it upgrades some of the hydrogen from the cells and provides a material for internal consumption in brine treatment and dechlorination. Still another option is production of ferrous and ferric chlorides from the metal.

For recovery as the element, absorption/desorption processes suggest themselves. Chlorine would be absorbed from the tail gas at high pressure and low temperature and then released under opposite conditions as a nearly pure vapor.

The ideal characteristics of a solvent have been enumerated [38]. Several plants have operated with carbon tetrachloride as solvent. Environmental restrictions on CCl_4 have put a stop to this practice, and some recovery units have been scrapped. The usual choice for their replacement has been enhanced liquefaction. This might take the form of a tertiary stage where the gas after further compression is exposed again to the lowest temperature in the liquefaction process.

Anolyte, which is a separate stream only with mercury and membrane cells, will contain dissolved chlorine. Most of this can be removed by adding acid to reverse its hydrolysis and reducing the operating pressure. It can join the chlorine from the cells for processing. A small amount of chlorine will remain in the brine. With mercury cells, this is tolerable and may even be beneficial by keeping mercury in solution and preventing deposits of the metal in the brine system. In the membrane-cell process, the residual Cl_2 is not tolerable. A second stage of dechlorination then is necessary. Here, Cl_2 and OCl^- are chemically or catalytically reduced to Cl^-. The depleted and dechlorinated brine then returns to the brine feed process.

Differences in caustic soda processing are profound, reflecting the differences among the various cell products:

1. Mercury cells: NaOH at sales or use concentration, low levels of impurities
2. Membrane cells: 30–35% NaOH, some anionic impurities (Cl^-, ClO_3^-, SO_4^{2-})
3. Diaphragm cells: 10–12% NaOH, ~15% NaCl, other anionic impurities

Mercury-cell caustic requires very little processing. The evaporation of membrane-cell caustic is a straightforward process in which the amount of water to be removed is roughly equal to the amount of contained NaOH. A typical evaporation system will have two or three effects. The nature of diaphragm-cell caustic complicates the evaporation process. The combined presence of NaCl and NaOH will exceed the limits of solubility well before reaching the desired concentration of 48–50% NaOH. NaCl then begins to drop out of solution. Evaporation of about 85% of the water releases about 95% of the salt as crystals that form a concentrated slurry in NaOH solution. The salt, recovered on centrifuges, returns to the brine system for another pass through the cells. The solution is then cooled, releasing more dissolved salt, and is at its final concentration. The newly crystallized salt is removed from the solution by another set of centrifuges. Filtration of the product NaOH solution, which in the 50% grade still contains about 1% dissolved NaCl, normally is necessary to remove the last traces of solid salt.

Most diaphragm-cell caustic is sold with its residual dissolved NaCl. This has been a marked disadvantage in some applications. The less attractive composition of the product, the complexity and cost of an evaporation plant, and the increased energy requirement all favor the replacement of diaphragm with membrane cells. The evaporative load is about five times as great as that in a membrane-cell evaporator. As a result, diaphragm-cell evaporators generally have three or four effects.

A small fraction of the diaphragm-cell caustic output is purified to remove anionic impurities, principally Cl^-. This involves extraction by anhydrous ammonia in a column at high pressure. Extensive processing removes traces of ammonia from the purified caustic and reconstitutes the ammonia for recycle. The byproduct of this process is a stream of caustic soda with an elevated concentration of impurities. It may become a waste or may be used as a source of alkalinity.

In some applications, the water accompanying the caustic is undesirable or unacceptable in further processing. The cost of shipping it along with the caustic is always undesirable. Approximately 5% of the NaOH produced therefore is concentrated beyond 50%. Simple concentrating evaporators can produce 70–73% NaOH in one stage. More complex processing gives products that are 98–99.5% solids. Typically, this involves evaporation in vertical units heated by a molten salt followed by vacuum flashing to remove further traces of water. The molten product may be merely cast in drums. More elaborate processes include the formation of flakes on a cooled roll equipped with doctor knife and the formation of nearly spherical particles by atomization into a stream of dry, relatively cool air.

All the products of a chlor-alkali plant are hazardous, and the hazards are diverse. It would be unwise to provide an inadequate summary here. The interested reader may consult the voluminous safety literature, the guides and publications of industry associations [39, 40], and the standard references already cited here [41, 42].

Hydrochloric Acid

More than 93% of the hydrogen chloride used in the United States is obtained as a byproduct. The chief source is the chlorination of hydrocarbons. Other important sources are isocyanate and fluorocarbon production and the thermal oxidation of chlorinated organic wastes. In substitution chlorination of paraffinic hydrocarbons, HCl is generated in the chlorination reaction. Chlorine never enters the organic molecule during the preparation of isocyanates. Its function is that of a carbonyl carrier, being used first to produce phosgene. Phosgenation of an amino compound then produces two molecules of byproduct HCl.

After addition chlorination of olefins, HCl may be generated by a pyrolysis that leaves a monochlorinated olefin product. In the classical process, this is so in the case of

VCM, the principal derivative of chlorine (*supra*). Addition of Cl_2 to ethylene produces ethylene dichloride. VCM results when a molecule of HCl is split off. If all VCM were produced this way, there would be a serious glut of HCl. This fact provided much of the incentive for development of the oxychlorination process in which HCl is oxidized as it forms to regenerate chlorine:

$$4HCl + O_2 \rightarrow 2Cl_2 + 2H_2O$$

Ideally, all of the chlorine value goes into the product, with no net formation of HCl.

Most synthetic HCl is the product of combustion of chlorine with hydrogen. This is a convenient way to upgrade the tail gas from chlorine liquefaction (*supra*). HCl is recovered as the acid, and much of it is recycled for internal use. Accurate data on hydrogen chloride production and consumption are difficult to obtain because of the large numbers of producers and consumers, the variety of grades, and the confidentiality of information relating to captive use of HCl.

Hydrogen chloride is marketed both as anhydrous HCl and as hydrochloric acid, typically a 32% solution. HCl normally is absorbed in a falling film column followed by a packed tails tower that vents inerts into the atmosphere. The weak acid from the tails tower flows to the primary absorber. The heat of absorption of HCl is about 400 kcal/kg and requires that the primary absorber be water-cooled. With the addition of a ceramic burner to produce HCl from H_2 and Cl_2, similar apparatus becomes an HCl synthesis unit.

Acid made by burning hydrogen in chlorine tail gas is quite pure and requires no further treatment except, in some cases, for the removal of traces of dissolved chlorine. Byproduct acid requires purification in packed scrubbers in which the parent hydrocarbon is circulated in the scrubbing tower before flowing to the chlorinator. Frequently, the absorption efficiency is improved by cooling the scrubbing liquor with a refrigerant.

Bromine and Sodium Bromide

Bromine has long been produced by the displacement of bromides from seawater. While this is an inexhaustible source (10^{14} metric tons Br), operation is expensive because of the low concentration of bromide (about 65 ppm [43]), and the process no longer operates in the United States. Rather, underground brines found in Arkansas at depths of about 2,400 m and containing up to 0.6% Br are the favored source. There are two US producers with combined capacity of 278,000 t/year [44].

The United States is the world's largest producer of bromine (about one-third of the total), but China and the Dead Sea countries account for a large and growing share.

Table 26.17 US production and consumption of bromine

	2007	2008	2009	2010
Production	243	w[a]	w[a]	w[a]
Exports	8.6	9.6	8.1	8.0
Imports	32.2	41.2	34.2	45.0
Consumption[b]	277	w[a]	w[a]	w[a]

Source: US Geological Survey Mineral Commodity Summary, *Bromine*, 2011

All quantities in thousands of metric tons bromine equivalent

[a] Quantities withheld; fewer than three domestic producers

[b] Apparent consumption = production + imports–exports

Table 26.18 Worldwide production and reserves of bromine, 2010

Country	Production (est.)	Reserves
United States	–	11,000
Israel	185	(Dead Sea)
China	150	130[a]
Jordan	85	(Dead Sea)
Japan	20	NR
Ukraine	4	400
Azerbaijan	3.5	300
Spain	0.1	1,400
Turkmenistan	0.15	700
Others	3	(from bitterns)
Total non-US	450	Large

Sources: US Geological Survey: Mineral Commodity Summary, *Bromine*, 2011; advanced release of 2010 Minerals Yearbook

All quantities in thousands of metric tons bromine equivalent

[a] China has a total of 3.5 million metric tons of base reserves of uncertain quality and value (reported elsewhere)

Table 26.17 shows that production and consumption rates in the United States are hard to obtain because data are not reported when there are fewer than three active producers. Table 26.18 shows worldwide production and reserves. Non-US production in 2010 was about 10% higher than the average of the previous 4 years. Bitterns of potash production from Dead Sea brines, used in Israel and Jordan, are very rich in bromide, containing 1.0–1.2% Br (w/v). The total amount of bromine available there is about 10^9 metric tons. Some countries produce relatively small amounts of bromine by recovering it from solar salt or potash bitterns.

There are four major steps in the recovery of bromine from brine: (1) oxidation of Br^- by injection of chlorine, (2) stripping of elemental bromine from solution, (3) recovery of bromine from vapor, and (4) purification. The first step depends on the equilibrium between chlorine and bromine:

$$Cl_2 + 2Br^- \leftrightarrow 2Cl^- + Br_2$$

In the second step, elemental bromine is released by contact of the solution with air or steam. There is an economic balance here between the cost of steam and the difficulty of recovering bromine from air in the third step.

Generally, steam is favored when processing brines and air when processing the much more dilute seawater. The use of air requires the bromine to be trapped in an alkaline or reducing solution to concentrate it.

Operating a steam-based process under vacuum allows bromine to be stripped at lower temperature. Claimed benefits are increases in steam economy and plant capacity and reductions in chlorine loss, spent brine treating requirements, maintenance costs, and atmospheric emissions [45, 46].

Primary uses of bromine and bromine chemicals are in flame retardants, drilling fluids, water treatment, and pesticides. The once-prominent market in petroleum additives has declined with the phase-out of lead-based antiknocks, which consumed ethylene bromide. Similarly, certain brominated flame retardants are being withdrawn from the market [44]. A growing application is the use of bromine to bond with mercury in flue gases. In some cases, it replaces chlorine as a scavenger. The mercuric bromide that forms is relatively easy to remove in flue-gas scrubbers. One process uses fine powdered carbon impregnated with about 10% bromine [47].

Sodium bromide sometimes is recovered as a byproduct of other operations, and it can be prepared by adding excess bromine to a solution of NaOH. After evaporation to dryness, treatment with a reducing agent converts sodium bromate to NaBr. It is used as an intermediate in production of other chemicals, often by double decomposition with organic chlorides. Bromides are the proverbial sedatives, and other uses are as algaecides and photographic chemicals. Their major use has been in drilling fluids, where the high density of concentrated aqueous solutions (up to 1.55 g/cm^3) is useful in well completions. NaBr is the form of supply when using bromine as a biocide. The active form is released by activators such as hypochlorites and chlorinated isocyanurates.

Recovery of bromine from byproduct solutions and HBr emitted in organic brominations is the source of roughly one-third of total production. Incineration of plastics that contain brominated flame retardants also provides bromine for recovery [48].

Bleaches and Disinfectants

The oxidative and biocidal properties of chlorine have been widely exploited from the beginnings of the chemical industry. Textile manufacturers and the paper industry have long recognized the value of bleach in their operations. The extensive use of chlorine and its derivatives to bleach wood pulp is described in Chap. 28.

Active chlorine compounds also find wide use as disinfectants and in laundry bleaching. The spread of the great killer diseases cholera and typhoid was checked by the addition of small amounts of chlorine bleach to wastewater and drinking water. Today, potable water disinfection accounts for only a few percent of the world's chlorine consumption (see Table 26.15) but remains perhaps its most important and most beneficial use. The use of chlorination is widespread only in the developed world, and an estimated one-third of deaths and 80% of all diseases in developing countries are due to consumption of contaminated water [49].

When chlorine is added to water, it hydrolyzes to form hydrochloric and hypochlorous acids:

$$Cl_2 + H_2O \rightarrow H^+ + Cl^- + HOCl$$

HOCl is a strong disinfectant that is not ionized at low pH. Rather than transport and handle liquid chlorine, both hazardous undertakings, treatment plant operators may use solutions of sodium hypochlorite. This is produced, along with an equal amount of salt, by absorption of chlorine into a solution of caustic soda:

$$Cl_2 + 2NaOH \rightarrow NaCl + NaOCl + H_2O$$

The strength of the solution may vary to suit the application. The use of 20% NaOH gives a 16% solution of NaOCl, which is regarded as commercial strength. At high pH, hypochlorite is completely ionized. It is less effective in this form, and, at lower pH at the point of application, it may convert to the more active HOCl.

Before the commercialization of chlorine liquefaction, bleaching compounds were the only transportable form of chlorine. The earliest practice was to absorb chlorine in hydrated lime, $Ca(OH)_2$. This forms chloride of lime, $CaOCl_2$, which is easily transported as a solid. Dissolved in water, it forms equal moles of $CaCl_2$ and $Ca(OCl)_2$:

$$2CaOCl_2(\text{dissolved}) \rightarrow CaCl_2 + Ca(OCl)_2$$

Chlorination of milk of lime or of a slurry of lime produces the same mixture, in which calcium hypochlorite is the active ingredient [50]. If sodium hypochlorite is separately produced or if a mixture of lime and caustic soda is used, the lower solubility of $Ca(OCl)_2$ allows double decomposition of NaOCl and $CaCl_2$, and all the calcium value can be converted to bleach. The overall reaction becomes, ideally:

$$2Cl_2 + Ca(OH)_2 + 2NaOH \rightarrow 2NaCl + 2H_2O + Ca(OCl)_2 \downarrow$$

$Ca(OCl)_2$ has characteristics similar to those of NaOCl, but it is approximately neutral and in its granular form has a high available chlorine content. The chloride of lime as produced in the past contained only about 35% available chlorine.

The term "available chlorine" describes the bleaching capacity of a substance or solution referred to an equal weight of elemental chlorine. The chloride ion that results from hydrolysis of Cl_2 is not a bleaching agent, but the hypochlorite ion can abstract two electrons (the chlorine atom is in the +1 oxidation state, and $Cl^{1+}O^{2-}$ is a more informative notation). The net result of hydrolysis and bleaching is as if each chlorine atom picked up one electron in oxidation of the material being treated. A hypochlorite group, it will be seen, then has all the bleaching capacity of a molecule of chlorine. Pure calcium hypochlorite, noted above, therefore contains $2 \times 100 \times 71/143 = 99.3\%$ available chlorine. The commercial product actually is the dihydrate $Ca(OCl)_2{\cdot}2H_2O$. Its strength varies with the purity of the lime used in its preparation, the purity of the mixed crystals formed in precipitation, the degree of contamination by unreacted lime, and the final processing conditions. Centrifugation, granulation, and air drying normally yield a product with 65–70% available chlorine. This is the form packaged for sale.

Chlorine/hypochlorite has largely been displaced from wood pulp bleaching in the move to elemental chlorine-free (ECF) bleaching. It still is widely used in household bleaches because of its ease of manufacture and handling, as well as its low cost. Many plants are dedicated to the production of bleach and operate on purchased chlorine and caustic soda. Others are attached to chlor-alkali plants (*supra*) and frequently operate on liquefaction plant tail gas. Commercial soda bleaches vary in their NaOCl content up to 15–16%, the higher NaOCl contents requiring higher excess concentrations of caustic to maintain stability. Household bleach is relatively dilute and normally contains 5–6% available Cl_2. Lacking the additives found in many consumer products, bleach has few residual effects, and the Centers for Disease Control and Prevention recommend its use for all sanitizing purposes [51].

The conventional expression "trade %" is simply the available chlorine content in weight/vol.%. Thus,

$$(\text{trade \% available } Cl_2) = (\text{weight \% available } Cl_2) \times (\text{specific gravity})$$

Since pure sodium hypochlorite has 95.3% available chlorine, one also can write

$$(\text{trade \% available } Cl_2) = (\text{gpl NaOCl}) \times 0.0953$$

Bleach also can be produced by the absorption of chlorine from liquefaction tail gas. Even the liquor from an emergency vent scrubber may be a useful bleach. Maintaining the quality of the bleach often is at cross purposes with the function of an emergency scrubber. Design and operating technique can overcome some of the difficulties. Some key points as well as the problems of decomposition of hypochlorite and the use of filtration and proper materials of construction to improve product stability are discussed elsewhere [52].

A much weaker bleach results from electrolysis of seawater in undivided cells. The electrode products, NaOH and Cl_2, react as noted above to produce hypochlorite. As a low-level bleach and disinfectant, it can serve as a utility-type water and sewage treating agent. It is a natural choice for marine sanitation and as an algaecide in seawater evaporation plants.

Dilute sodium hypochlorite is approved for sanitization of food processing equipment. A concentration of 200 ppm available chlorine is acceptable without washing if equipment is adequately drained [53]. More concentrated solutions may be used along with proper rinsing protocols.

The insoluble components in lime-based products are especially objectionable in laundry and dishwashing uses. Sodium compounds are preferred, and sometimes in spite of their greater cost, lithium compounds. LiOCl can be prepared by absorbing chlorine in LiOH or by adding LiCl to soda bleach. In the latter case, it is the insolubility of NaCl in concentrated solution that drives the double decomposition:

$$NaOCl + LiCl \rightarrow LiOCl + NaCl \downarrow$$

The demand for chlorine dioxide was an early beneficiary of the move away from chlorine bleaching of wood pulp. For many years, ClO_2 had been used late in many pulp-bleaching sequences. It is a strong bleaching agent that can remove certain compounds that resist bleaching by chlorine. Its use thus increases the brightness of the final product. It also produces fewer waste products that require disposal [54] and in particular produces much less adsorbable organic halide and no dioxin [55]. The use of more ClO_2 in earlier stages permitted the reduction or elimination of chlorine bleaching.

The shift to ClO_2 increased the market for sodium chlorate. Chlorine dioxide is very unstable and cannot be safely and economically transported. Instead, it is produced at the point of use by treating a solution of $NaClO_3$ and NaCl with acid and a reducing agent. There are several different types of commercial chlorine dioxide generator in operation, classified by acid medium (sulfuric, hydrochloric) and reducing agent (sulfur dioxide, methanol, sodium chloride, hydrochloric acid). Most processes generate some elemental chlorine as a byproduct. The use of sulfuric acid or SO_2 also produces sodium sulfate or waste acid for disposal. As an example of the chemistry, consider HCl, which serves both as acid and as reducing agent:

$$NaClO_3 + 2HCl \rightarrow ClO_2 + {}^1\!/_2 Cl_2 + NaCl + H_2O$$

Generators must be operated with care to avoid ClO_2 concentrations above 10%, which can lead to explosive decomposition. The reactor offgas therefore is diluted with air, and the ClO_2 is immediately absorbed in water. The molar ratio of ClO_2 to Cl_2 formed is typically 2:1. If this level of Cl_2 is objectionable, the gas is passed through a tower in contact with circulating chilled water. This dissolves all the ClO_2 but only about 25% of the Cl_2, raising the molar ratio to $8ClO_2{:}1Cl_2$. The remaining Cl_2 can be removed by scrubbing with alkali.

Since the oxidation state of chlorine in ClO_2 is +4, it abstracts five electrons as it is reduced to chloride. A chlorine molecule, as we have seen, transfers only two. The available chlorine content of ClO_2 therefore is $5/2 \times 100 \times 71/67.5 = 263\%$.

Sodium Chlorate

Sodium chlorate is produced by the electrolysis of NaCl brine in a cell similar to a chlor-alkali cell but lacking means to separate the products of the two electrodes. The chlorine and caustic soda produced in the cell therefore react immediately, producing sodium hypochlorite:

$$Cl_2 + 2NaOH \rightarrow NaOCl + NaCl + H_2O$$

The cell liquor is kept hot (80–95°C) in a vessel that provides enough residence time for the hypochlorite to disproportionate into chlorate and chloride:

$$3NaOCl \rightarrow NaClO_3 + 2NaCl$$

The overall reaction, requiring 6 F/mole, becomes

$$NaCl + 3H_2O \rightarrow NaClO_3 + 3H_2$$

Optimum cell conditions are 80–90°C and pH 6–6.5. The liquor contains 550–750 gpl $NaClO_3$ and 100–120 gpl NaCl [56]. Coated titanium anodes and steel cathodes are used, and small amounts of dichromate are added to the cell liquor to prevent the cathodic reduction of chlorate or hypochlorite. Resulting current efficiencies are above 90%, and energy consumption is in the vicinity of 4,500 kWh/metric ton.

The adoption of permanent metal anodes permitted an increase in sodium chlorate and a reduction in salt concentrations in the liquor without loss of current efficiency. This in turn allowed direct recovery of sodium chlorate by evaporative crystallization [57]. This was an advance on the process described in previous editions of this handbook. The earlier process yielded solid salt when the liquor was evaporated. After removal of the salt, cooling of the liquor produced $NaClO_3$ crystals.

Nearly all the sodium chlorate made in North America is used in the on-site manufacture of chlorine dioxide for pulp bleaching. The surge in demand due to the industry's adoption of ECF bleaching has run its course. The rest is used to produce other chlorates, perchlorates, and chlorites; in water treatment; and in the hydrometallurgical recovery of uranium. Sodium chlorate also is a nonselective herbicide. While this market might tend to increase along with GDP, there is also pressure to convert to total chlorine-free (TCF) bleaching. The market for $NaClO_3$ has been growing faster where conditions are more favorable. In Asia and South America, the abundance of faster-growing pulps presents more opportunity. Russia, with 22% of the world's forest and relatively cheap electricity, also has seen some growth. Where electricity is more expensive, the higher cost of $NaClO_3$ has allowed H_2O_2 to penetrate the bleaching-chemical market. In Europe and Japan, where ClO_2 accounts for only 90 and 84% of consumption, respectively, the use of sodium chlorate as a herbicide is more common. However, this market is threatened by a European Union directive to withdraw sodium chlorate products from the market [58].

Potassium chlorate can in some circumstances be produced electrolytically from KCl. It is more common to produce it from $NaClO_3$ by double decomposition:

$$KCl + NaClO_3 \rightarrow KClO_3 + NaCl$$

After the reaction is carried out in solution, crystallization yields solid $KClO_3$. $KClO_3$ is used in matches, pyrotechnic products, explosives, cosmetics, and pharmaceuticals [59].

Bromates are produced electrolytically, like chlorate, or by bromination of the respective carbonates. Bromates are used in small quantities, but they are important in the bread-making industry for maturing flour and conditioning dough. They also are used in permanent-wave lotions and in the manufacture of dyes.

Both chlorates and bromates have substantial oxidizing power, and care must be taken in their storage and handling. Contact with reducing agents or combustibles can be disastrous.

References

1. www.saltinstitute.org/Uses-benefits. Accessed 21 Nov 2011
2. Kaufmann DW (ed) (1960) Sodium chloride. ACS Monograph 145, Reinhold, New York, pp. 562–566
3. Sedivy VM (2009) Global NEST J. 11(I)41
4. Feldman SR (2006) Sodium chloride. In: Kogel JE, Trivedi NC, Barker JM, Krukowski ST (eds) Industrial minerals and rocks, 7th edn. Society of Mining Engineers, Littleton
5. Sedivy VM (2009) Proceedings, Ninth international symposium on salt, Beijing, vol. 2, pp. 1385–1402

6. United States Geological Survey (2011) Mineral commodity summary. In: Soda ash. USGS, Denver
7. United States Department of the Interior, brochures and websites, 2010, 2011.
8. O'Brien TF, Bommaraju TV, Hine F (2005) Handbook of chlor-alkali technology, vol V., pp 1391–92
9. Brownfield ME et al (2009) Fact Sheet 2009–3011, United States Geological Survey, Denver
10. Dyni JR (1974) Rocky mountain association of geologists guidebook, Colorado Geological Survey, Denver, pp. 111–122
11. www.icis.com, Chemical profiles: sodium bicarbonate, 2007
12. United States Geological Survey (2011) Mineral commodity summary. In: Sodium sulfate. USGS, Denver
13. United States Geological Survey (2006) Statistical compendium, minerals
14. www.chemexpo.com, Chemical profiles: sodium sulfhydrate, 2007
15. Chemical economics handbook, 780.4000, SRI International, Menlo Park, 1999
16. www.britannica.com/EBchecked/topic/469121/polysulfide, 2011-11-21.
17. Bean SL (1997) Thiosulfates. In: Kirk-othmer encyclopedia of chemical technology, 4th edn, vol. 24, Wiley, New York, p. 51
18. Water treatment uses and applications—dechlorination. Allied Corporation Product Bulletin, Syracuse, 1985
19. www.chemexpo.com, Chemical profiles: sodium sulfite, 2008
20. Product brochure, BASF Corporation, 2006
21. Bengston S, Bjerle I (1975) Chem Eng Sci 30:1429
22. MacMullin RB (1950) Chem Eng Progr 46(9):440
23. U.S. Patent 4,793,906
24. PQ sodium silicates: liquids and solids. Brochure 17–103. PQ Corporation, Valley Forge, 2004
25. Product stewardship summary: sodium metasilicate. Occidental Chemical Corporation, Dallas, 2008
26. North American chlor-alkali industry plants and production data report. The Chlorine Institute, Washington, 2010
27. Shamel RE (2010) President, Consulting Resources Corp., Lexington, personal communication, 2010-06-22
28. Schmittinger P (ed) (2000) Chlorine: principles and industrial practice. Wiley, New York, p. 123
29. O'Brien, Bommaraju, Hine, op. cit, vol. V, pp. 1389–1392
30. Schmittinger, op. cit., pp. 58–59
31. Florkiewicz TF (2001) Advanced diaphragm cell technology. In: Moorhouse J (ed) Modern chlor-alkali technology, vol 8. Blackwell Science, Oxford, pp 165–72
32. www.eurochlor.org, 2011-09-03
33. Collins JH, Entwisle JH (1980) Development and operation of high-current density mercury cells. In: Coulter MO (ed) Modern chlor-alkali technology. Ellis Horwood, Chichester
34. O'Brien, Bommaraju, and Hine, op. cit, vol. III, p. 937
35. O'Brien TF, White IF (1995) Process engineering aspects of chlorine cooling and drying. In: Curry RW (ed) Modern chlor-alkali technology, vol 6. Royal Society of Chemistry, Cambridge, pp 70–81
36. O'Brien TF, White IF (1998) Process engineering considerations in chlorine compression and liquefaction. In: Sealey S (ed) Modern chlor-alkali technology, vol 7. Royal Society of Chemistry, Cambridge, pp 202–213
37. Silver MM (1982) Chlorine Tailgas and Snift disposal systems. 25th Chlorine institute plant operations seminar, Atlanta
38. O'Brien, Bommaraju, and Hine, op. cit., vol. III, p. 885
39. Euro Chlor, Brussels, www.eurochlor.org.
40. The Chlorine Institute, Washington, www.chlorineinstitute.org.
41. Schmittinger, op. cit., pp. 155, 156
42. O'Brien, Bommaraju, and Hine, op. cit, vols. II, III, and V
43. Castro P, Huber ME (2007) Marine biology. McGraw-Hill, New York
44. United States Geological Survey, Minerals yearbook 2010 (advanced release), *Bromine*, 2011
45. U.S. Patent 4,719,096
46. U.S. Patent 4,725,425
47. Albemarle Corporation product literature, 2010
48. Frim, R., and Ukeles, S. D., Mining Engineering 59, No. 6, 22 (2007)
49. Schmittinger, op. cit., p. 160.
50. Sconce JW (1972) Chlorine: its manufacture, properties and uses, ACS Monograph No. 154, Robert E. Krieger, Huntington, NY, p. 180
51. www.marylandpublicschools.org. General sanitation guidelines, 2011-11-21
52. O'Brien, Bommaraju, and Hine, op. cit., vol. V, pp. 1378–1386
53. U. S. Code of Federal Regulations, Title 21, Part 178
54. Reeve DW, Earl PF (1989) Pulp and Paper Can 90:4
55. Fleming BL et al (1988) Pulp and Paper Can 89:12
56. Schmittinger, op. cit., p. 172
57. U.S. Patent 3,883,406
58. Schlag S, Mori H (2008) Sodium chlorate. In: Chemical economics handbook, SRI, Menlo Park
59. Schmittinger, op. cit., p. 173

Industrial Gases

27

Steven J. Cooke

Overview

Industrial gases may actually be used as gases, liquids, or cryogenic liquids. Industrial users generally accept them as those gases used primarily in their pure form in large quantities. Most of the gases we consider to be industrial gases have been in use for many years. Processes for the cryogenic separation of the air gases were developed as early as 1895, with commercial production of oxygen beginning in 1902. Nitrous oxide was used as an anesthetic as early as 1799. Carbon dioxide had been identified as a specific substance by 1608. Methane has been used as an energy source since the 1700s. Other gaseous compounds commonly used today for specific manufacturing processes (e.g., electronics/semiconductors, plastics) are discussed in other chapters of this Handbook related to those processes.

The reactivity of gases can be summarized into three classes: oxidizers, inert gases, and flammable gases. The gases that fall into the category of inert are nitrogen, argon, helium, and carbon dioxide. The oxidizers are oxygen, chlorine, and nitrous oxide. Acetylene, liquefied natural gas (LNG), and hydrogen are the flammable gases. These elements and compounds by no means encompass all gases considered to be industrial gases, but they represent the major gases used and produced in industry (see Table 27.1) [1].

The cryogenic air separation process, in which air is liquefied and separated into its major constituents (oxygen, nitrogen, argon) by the use of cryogenic technology, is the major source of nitrogen, oxygen, and argon produced for industry, as well as the noble gases krypton, neon, and xenon. Table 27.2 gives the cryogenic boiling point temperatures and concentrations for some of the gases present in the atmosphere.

The cryogenic air separation process starts with air entering the plant through air filters, where it is compressed and cooled. The air is passed through heat exchangers for further cooling and for removal of water vapor and carbon dioxide by liquid condensation. Solid adsorbents, such as molecular sieves, silicas, and aluminas, can also be used to remove the water and carbon dioxide by adsorption before separation of the air. When the air reaches a temperature of −338°F, it is sufficiently free of water and carbon dioxide and ready to be passed into the distillation column [6].

Separation occurs by distillation, which is the physical partition of compounds by the differences in their boiling points (see Fig. 27.5 below in the section "Argon"). The distillation method used is referred to as "double column," as it contains two separate distillation columns operated at two different pressures. The cooled air is fed to the bottom of the high-pressure column. Air is distilled in this column where the higher boiling temperature fluid oxygen stays at the bottom and the lower boiling temperature nitrogen boils to the top as a distillate product. The crude liquid oxygen from the high-pressure column is flashed into the low-pressure column as the two columns are thermally linked by a common reboiler/condenser.

Vapor is withdrawn from the middle of the low-pressure column and is passed as feed to the crude argon column. The vapor in that column is condensed and taken off as liquid crude argon. Pure nitrogen vapor is extracted from the top of the low-pressure distillation column, and liquid oxygen is taken from the bottom of this column. Figure 27.1 shows an air separation plant with a molecular sieve front-end cleanup and liquid storage tanks.

Alternatives to the cryogenic separation process include the membrane separation process and adsorption processes. The adsorption processes include pressure swing adsorption (PSA) [15] and vacuum swing adsorption (VSA). These methods are non-cryogenic and produce a vapor product only. This reduces the cost of production considerably when the local use of gas-phase product is the primary objective.

S.J. Cooke (✉)
Process Systems Consulting, Charlotte, NC, USA
e-mail: scooke@sprynet.com

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_27,

Table 27.1 Properties of Industrail Gases

Gas Property		Helium-4 (He)	Para-hydrogen ($p-H_2$)	Normal Hudrogen ($n-H_2$)	Air	Nitrogen (N_2)	Oxygen (O_2)	Argon (Ar)	Carbon Dioxide (CO_2)	Acetylene ($(C_2)H_2)$)	Methane (CH_4)
Atomic/ molecular weight	4.00	2.016	2.016	2.016	28.975	28.01	31.9988	39.95	44.01	26.04	16.042
Normal boiling point	°F	−452.1	−423.2	−423.0	−317.8	−320.4	−297.3	−302.6	−109.33[c]	−103.4 (10 psig)	−258.68
										−118[c]	
Triple point	°F	−455.75	−434.8	−434.55	–	−346.0	−361.8	−308.8	−69.9	−116	−296.5
	psia	0.730[a]	1.021	1.045	–1.81	0.0216	9.99	60.4	17.7	1.69	
Critical point	°F	−450.3	−400.31	−399.93	−221.1	−232.4	−181.4	−188.1	87.9	96.8	−115.78
	psia	33.0	187.5	190.8	547	493	731.4	711.5	1070.6	907	673.1
Gas density at NTP	Lb/cf	0.0103	0.00521	0.00521	0.07493	0.072	0.08279	0.103	0.1144	0.0678	0.042235
Gas density STP	Lb/cf										
Vapor density at NBP	Lf/cf	1.0543[b]	0.084	0083	0.2805	0.2879	0.2795	0.3606	0.1462	–	0.1134
Liquid density at NBP	Lb/cf	7.802	4.42	4.43	54.56	50.48	71.23	87.02	73.5[d]	24.0 (70°F)	26.57
Specific heat at NTP	BTU/ tb-°F	1.24	3.555	3.425	0.241	0.249	0.2197	0.125	0.203	0.383	0.5271
Specific heat ratio at NTP	1.66	1.38	1.42	1.40	1.41	1.40	1.67	1.304	1.26	1.307	
Latent heat of vaporization at NBP	BTU/ lb	8.72	191.6	191.7	88.2	85.6	91.7	69.7	245.5[c] (C)	264	219.22
Ref: 8 *CGA Handbook*, 1999	p. 408	p.416	p.415	p.234	p.528	p.555	p.262	p.296	p.225	p.488	

NTP = 14.696 psia and 70°F; STP = 14.696 psia and 32°F; Lb/cf = pound per cubic foot.
[a] Lower lambda point; [b] NIST Technical Note 631, "Thermophysical Properties of Helium 4.", Nov. 1972; [c] Sublimation point; [d] Triple point. CGA References from NIST Technical Notes 1025 (1980), 1048 91982), 1079 (1985), 361, Monograph 168 (1981).

Table 27.2 Cryogenic Gases: Boiling Point and Concentration in Air

	Normal Boiling Point (°F)	Dry Air Concentration (% v/v)
Argon	−302.6	0.93
Nitrogen	−320.5	78.1
Oxygen	−297.3	20.94
Helium	−452.1	0.0005
Hydrogen	−422.99	0.00005
Neon	−410.7	0.0018
Krypton	−244.0	0.00011
Xenon	−163.0	0.000008

The membrane separation process is used to produce a number of gases for industrial use. Many membrane technologies have been developed over the last 10 years, including various polymer and pure carbon membranes [13, 17]. A common polymer membrane may consist of a microporous substructure of cellulose acetate and a thin layer of dense cellulose acetate (active layer) on the upper surface, resembling a sponge covered by a plastic film. The membrane is cast into a supporting cloth for added mechanical strength. The active layer serves as the separating barrier

Fig. 27.1 Air separation plant with a front-end cleanup, cold box tower and liquid storage tanks. (Courtesy Air Liquide.)

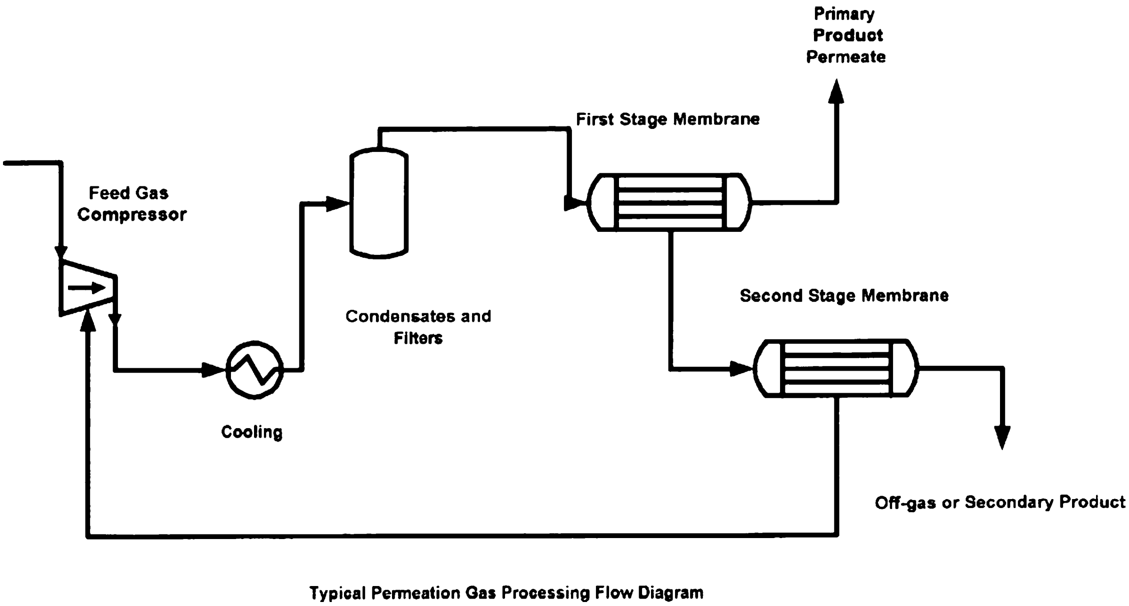


Fig. 27.2 Typical permeation gas processing flow diagram

and, because of its thinness, provides very high transport rates. The membrane selectively allows the smaller molecular-sized components to pass through the membrane substructure, thereby accomplishing the separation process. Carbon and zeolite-based membranes depend on the relative molecular size of the gases to effect the separation of the gases. A typical membrane separation flow system is shown in Fig. 27.2.

The VSA system is a non-cryogenic process that makes use of pressure that is less than atmospheric (vacuum) and adsorbents to separate nitrogen, and carbon dioxide from air. The PSA system is a similar process for producing a particular gas by passing a pressurized gas stream through a set of adsorbent beds. The impurities are selectively adsorbed using a molecular sieve or other selective adsorbents to trap the undesired components. In the production of oxygen using the VSA process, the VSA system separates air at ambient temperature by passing it through a column of adsorbent. The adsorbent, which is an inorganic crystal material, selectively adsorbs nitrogen molecules on its

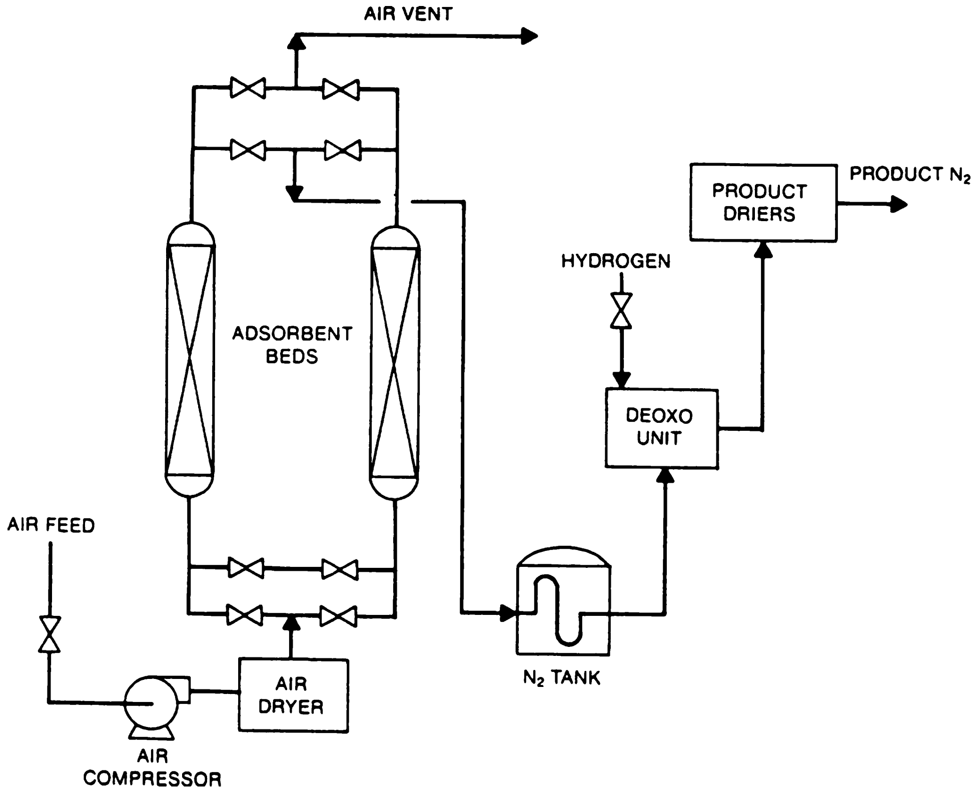


Fig. 27.3 Process diagram for nitrogen PSA system. (Courtesy Air Products and Chemicals, Inc.)

surface, while the oxygen molecules pass on through the adsorbent column.

The final oxygen product exits the adsorber column. The nitrogen adsorbed is released when the adsorber column is reactivated by use of differential pressure from a vacuum. Once the adsorber column is reactivated, the cycle is ready to be repeated. In the PSA system, compressed air is passed through a column of chemical adsorbent, which is a carbon-based material that essentially traps the oxygen molecules and allows the nitrogen molecule to pass through to the use point. Figure 27.3 shows the process flow of a PSA system, which can produce 98–99.5% pure nitrogen without the need for downstream purification. These adsorption systems usually are located at customer sites. See Fig. 27.4 for an example of an on-site PSA.

Small-volume users of the industrial gases are supplied by high-pressure gas cylinders filled with the desired product. These cylinders then are loaded onto flatbed trucks for delivery. For larger-volume gas transportation, long-length gas cylinders are permanently mounted on trailer frames. These high-pressure-cylinder vehicles commonly are referred to as tube trailers. The cylinders are constructed of alloy steels, and are designed, built, tested, and maintained in accordance with US Department of Transportation (DOT) regulations in the United States. Similar regulatory agencies in other countries define the requirements for safe transport and storage of compressed gases (e.g., Transport Canada, Ministère des Transports (France), Department for Transport (UK), and other national regulators). The typical small-volume cylinders are constructed to a DOT 3A or 3AA specification, and typically have rated service pressures from 2,015 to 2,640 psig and water volumes from 0.39 to 1.75 ft^3. The larger cylinders often have similar pressure ratings but water volumes of approximately 75 ft^3.

The containers used for transporting and storing the majority of cryogenic liquids are insulated cryogenic tankers, liquid cylinders (sometimes referred to as Dewars), and stationary storage tanks. These containers are similar in design to a Thermos® bottle. There is an inner vessel, which contains the cryogenic product, and an outer vessel. The space between the two vessels is under a vacuum and usually contains a wrapped insulation material such as aluminized Mylar, or it may contain a powdered insulation such as expanded perlite. The liquid cylinders, usually 160–250 L in capacity, are constructed to a DOT 4 L specification and operate with pressure up to 235 psig.

The cryogenic tank trucks may be constructed to a DOT MC-338 specification, and range in capacity from 3,000 to 14,000 gal, usually depending on the product. The operating pressures may range from 25 to 150 psig, and the vehicles often are equipped with cryogenic pumps used to deliver the

Fig. 27.4 An on-site nitrogen PSA supply system. The large vessels are absorbers and the smaller ones are dryers. (Courtesy Cirmac International, BY, Netherlands.)

product to higher-pressure storage tanks. The stationary storage tanks are designed and built to ASME standards, typically are constructed for pressures of 250 psig or less, and may range in capacity from 500 to 70,000 gal [3].

Nitrogen

Nitrogen makes up 78% of the earth's atmosphere. It is a colorless, tasteless, nontoxic, and relatively inert gas. Nitrogen does not support combustion nor does it support respiration; but it reacts with the more active metals such as lithium and magnesium to form nitrides, and at high temperatures it will combine with hydrogen, oxygen, and other elements. As a cryogenic liquid, nitrogen is nonmagnetic, stable against mechanical shock, and free from toxic or irritant vapor.

Nitrogen usually is produced by cryogenic air separation, liquefaction of atmospheric air, and removal of the nitrogen by distillation, as described earlier. Non-cryogenic adsorption technology, including PSA and VSA systems, economically provides nitrogen for requirements between 4,000 and 30,000 standard cubic feet per hour (SCFH) [13]. The development of more efficient membrane separation systems has happened in the last 10 years. The production of gaseous nitrogen for refinery instrumentation and vessel purging needs has shifted more to the use of membrane systems on site rather than cryogenic liquid in bulk storage due to significant cost savings.

Nitrogen has many applications, both as a gas and as a liquid. Gaseous nitrogen is widely used for inerting and pressurizing systems. Applications include the agitation of color film solutions in photographic processing; blanketing and deaeration of oxygen-sensitive liquids, volatile liquid chemicals, and nonferrous metals; and in the production of semiconductor electronic components, primary metals heat treating, and brazing of copper tubing. It is used to pressurize aircraft tires and emergency gas bottles to operate landing gear; and often to purge and pressurize pipelines, reactor vessels, or storage tanks throughout the petrochemical industry [1].

Gaseous and liquid nitrogen have applications in the foods industry in the areas of food processing and packaging. Liquid nitrogen is used for rapid freezing of expensive or highly perishable foods, such as shrimp and other meat products, as well as to refrigerate foods in long-distance hauling. A gaseous nitrogen atmosphere can help to retard the spoilage of some food products.

Liquid nitrogen is used in the cold-trapping of materials such as carbon dioxide and volatile organic carbons (VOCs) from gas streams, as a coolant for electronic equipment, for pulverizing plastics or rubber material, for deflashing of rubber tires, and for simulating the conditions of outer space. It is used to freeze liquids in pipelines during repairs to the pipeline. It can be used to generate very high-pressure gaseous nitrogen (15,000 psig) by liquid pumping. Liquid nitrogen is frequently used in medicine to remove skin blemishes in dermatology, and to preserve whole blood, livestock sperm, and other biological specimens.

Liquid nitrogen gas is transported in liquid cylinders, nonpressurized Dewars, cryogenic tankers, and trailers. Gaseous nitrogen is also transported in high-pressure gas cylinders and tube trailers. Both forms of nitrogen also are transported via pipelines when the sources are at economical locations and distances from the applications.

Oxygen

This gas, which occupies approximately 21% of the earth's atmosphere at sea level, has a number of very important uses. Oxygen is a colorless, odorless, and tasteless gas that is essential to the support of life. All elements except the inert gases combine directly with oxygen to form oxides. Oxygen is nonflammable, but is an oxidizer that readily supports combustion.

A large percentage of the oxygen produced commercially is made by cryogenic air separation. It may also be produced by PSA and membrane systems, as described in the production of nitrogen. Oxygen also may be made by the most common decomposition reaction, electrolysis of water.

Small quantities of pure oxygen may also be made by other chemical reactions. These reactions produce not only oxygen but by-products as well, and they are not usually as economical an alternative when compared with the other technologies.

Applications in enhanced combustion and life-support systems provide the majority of the demand for oxygen. It is used extensively in medical applications for therapeutic purposes, for resuscitation in asphyxia, and with other gases in anesthesia. Also it is used in high-altitude flying and deep-sea diving, and it is used for life-support and as a fuel oxidizer in the US space program [1].

For industrial applications, oxygen is in large demand for use with acetylene, hydrogen, and other fuel gases for purposes such as metal cutting, welding, hardening, and dehydrating. Oxygen enhances the capability of steel and iron furnaces. Oxygen is used in the production of synthesis gas–hydrogen–carbon monoxide mixtures from coal, natural gas, or liquid fuels; the synthesis gases are used to make gasoline, methanol, and ammonia. Oxygen may be used for the cracking of methane or natural gas by partial oxidation to produce acetylene, and in the production of nitric acid, ethylene, and other basic chemical feedstocks.

One of the largest commercial applications for oxygen is the Basic Oxygen Furnace (BOF) for refining iron and making steel. Another common application is its use in oxyfuel burners in the steel industry. It is used in electric arc furnaces, which melt scrap metals by passing an electric current through large graphite electrodes.

Oxygen enrichment is a process whereby oxygen is added to an air stream to improve a plant's efficiency. Combustion results in the generation of heat and gaseous by-products, usually carbon monoxide, carbon dioxide, and water vapor. Oxygen enrichment decreases the amount of nitrogen in the air stream, improving the combustion reactions while reducing the formation of undesired by-product gases. Oxygen enrichment is used in pulp mills to increase lime production from existing equipment and eliminate the need to purchase lime from other industries. It is also used in pulp mills to reduce airborne sulfur emissions to comply with federal environment regulations.

Gaseous oxygen is transported in high-pressure gas cylinders, tube trailers, and portable cylinder modules. Liquid oxygen is transported as a cryogenic fluid in liquid cylinders and insulated cryogenic tankers. Customer on-site storage is usually accomplished in large insulated cryogenic tanks.

Argon

Argon, a relatively scarce gas, is colorless, odorless, tasteless, and nontoxic. It forms no known chemical compound; thus it is extremely inert. Representing only 0.93% of the earth's atmosphere, it is the most valuable of the commodity air gases. Argon is utilized in a variety of applications. In the manufacture of lighting devices, it is used to fill lamp bulbs and to produce colors in display tubes for lighted signs.

Argon provides an inert atmosphere for the production of semiconductors and specialty metals, such as titanium and zirconium. It is often used to provide a protective shield for the growing of silicon and germanium crystals. Some metal welding and cutting applications require an inert gas shield, and argon gas often is used to fill this requirement [1].

Argon is manufactured in oxygen–nitrogen plants by means of fractional distillation after the liquefaction of air as described previously, in the cryogenic air separation process. In the distillation process where air is separated into oxygen and nitrogen, a stream of gas is withdrawn from the low-pressure column to enter the crude argon column. The argon content of the gas, initially 10%, is increased to 95% by subjecting the gas to several stages of distillation. This "crude" argon is further purified to yield the final pure argon product. Figure 27.5 shows a typical air separation process flow diagram. Argon is most economically shipped as a liquid. Liquid argon commonly is transported in liquid cylinders, cryogenic trailers, and portable tanks. Argon can be shipped in high-pressure gas cylinders and high-pressure cylinder trailers.

Hydrogen

Hydrogen, the lightest element known, is extremely flammable and is found in 0.00005% concentration in the air (0.5 parts-per-million). It has received notoriety in the late 1990s into the twenty-first century as a potential replacement for hydrocarbon fuels. This is because the actual combustion of hydrogen results only in the emission of water and the generation of heat which can be used in any of the existing forms of energy generation. However, as the subsequent discussion of hydrogen generation illustrates, there is much more to the carbon cycle of hydrogen use than the final combustion step [2].

The hydrogen molecule exists in two distinct forms: ortho and para, named according to their types of nuclear spins. *Ortho*-hydrogen molecules have a parallel spin, and *para*-hydrogen molecules have an anti-parallel spin. There is no difference in the chemical or transport properties of these forms, but there is a difference in the thermodynamic properties. *Para*-hydrogen is the form preferred for rocket fuels. Hydrogen consists of about three parts ortho and one part para as a gas at room temperature, which is referred to as normal or equilibrium hydrogen. The concentration of the para form increases with decreasing temperature until, for the liquid, the para concentration is nearly 100%. If hydrogen should

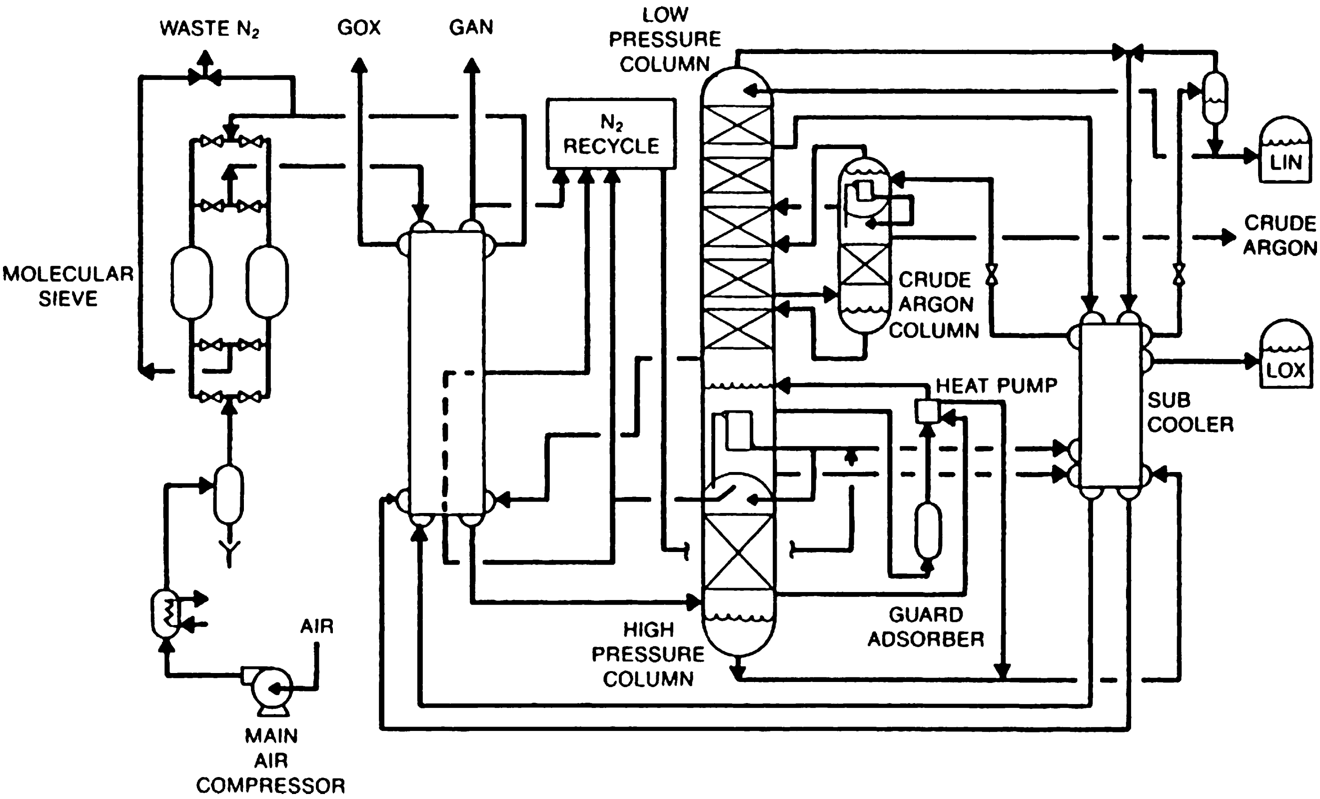


Fig. 27.5 Flow diagram for an air separation plant. (Courtesy Air Products and Chemicals, Inc.)

be cooled and liquefied rapidly, the relative three-to-one concentration of ortho to para would not immediately change [3].

Conversion to the para form takes place at a relatively slow rate and is accompanied by the release of heat. For each pound of rapidly cooled hydrogen that changes to the para form, enough heat is liberated to vaporize approximately 1.5 lb (0.68 kg) of liquid hydrogen. However, if a catalyst is used in the liquefaction cycle, para-hydrogen can be produced directly without loss from self-generated heat.

Hydrogen can be produced by steam methane reformation (the most widely used process) or by off-gas purification. Hydrogen can be produced at normal purity (95–98%) or at ultra-high purity (99.99%). The first step to obtain pure hydrogen in the steam reformer process when using a natural gas feedstock is to perform a pretreatment. There are small amounts of sulfur in natural gas that must be removed. Sulfur removal is done by preheating the natural gas to 700°F, and then sending it through a catalytic reactor. The catalyst inside the reactor adsorbs the sulfur. The second step is steam methane reforming. In this procedure, natural gas and steam are sent through a set of reformer tubes that are packed with a reforming catalyst. A furnace inside the reformer heats the reaction. The methane and steam react, and the results are 45% hydrogen, 40% steam, 8% carbon monoxide, and 7% carbon dioxide. The third step to obtain normal purity hydrogen is gas shifting.

Carbon monoxide, hydrogen, and steam are sent through a shift converter where the gas is heated between 400 and 700°F. The carbon monoxide and steam react to yield hydrogen and carbon dioxide. To produce pure hydrogen, the carbon dioxide must be removed. The gas passes through a carbon dioxide removal system, which contains a chemical solvent that selectively absorbs the carbon dioxide as the gas passes through the solvent [18]. Heat then is added to the solvent to discharge the carbon dioxide. The regenerated solvent is returned to the system to continue the removal of carbon dioxide (see also Chap. 22 for a detailed description of the water shift process).

To obtain ultra-pure hydrogen, a PSA [15] system may be used. In this process, the hydrogen, steam, and carbon dioxide are introduced to the system. At high pressure, the gas is passed through an adsorbent bed, and the adsorbent picks up the impurities and allows hydrogen to pass through the bed. In the regeneration process, the pressure is lowered to purge out the impurities. Then the hydrogen is used to rinse the adsorbent and repressurize the vessel. The efficiency of the hydrogen purification system is optimized at an operating pressure that allows use of a catalyst tube in the reformer. There are four other distinct processes by which hydrogen may be produced under the off-gas purification method: PSA, membrane separation, catalytic purification, and cryogenic separation. The cryogenic separation process uses

Fig. 27.6 A typical facility where liquid hydrogen is produced and stored. (Courtesy Air Liquide Canada.)

cryogenic principles to condense and separate the impurities from hydrogen. Impure hydrogen is introduced to the cryogenic system or "cold box." The impurities are condensed out, and pure hydrogen is obtained. Figure 27.6 shows a typical facility where liquid hydrogen is stored and produced. Membrane separation makes use of two layers of membrane material; the first layer is nonporous, and the second layer is porous [17, 19, 20]. A gas mixture under increased pressure is subjected to the surface, which allows a smaller molecule such as hydrogen to permeate faster than a larger molecule such as carbon dioxide.

The catalytic purification process requires the use of many catalysts to convert impurities into removable compounds. Hydrogen, chlorides, oxygen, and carbon oxides are passed through several catalytic reactors. In the first reactor, the chlorides are absorbed. In the second reactor, oxygen and hydrogen react to obtain water/steam, which is easily removed. In the third reactor, the carbon oxides react with hydrogen to yield methane, which would be considered an acceptable impurity in this particular hydrogen product stream. Some other sources of hydrogen are the dissociation of ammonia, steam reforming of light hydrocarbons, and by-product streams of chemical and petroleum processing. The "steam–iron" process is another method used to make hydrogen. Passing steam over heated, spongy iron reduces the steam to hydrogen with additional formation of iron oxide. There are several varieties of this process. The water–gas reaction is a reaction of steam with incandescent coke or coal; it is a source of hydrogen with carbon monoxide as an additional product. In the catalytic version of this reaction, excess steam breaks down to form more hydrogen while oxidizing the carbon monoxide to form carbon dioxide. Hydrogen frequently is obtained as a by-product of cracking operations using petroleum liquids or vapors as feedstock. Also hydrogen may be produced through the electrolysis of water.

Nearly all hydrogen production in the United States today is by steam reformation of natural gas. This, however, releases fossil carbon dioxide in the process and trades one relatively clean fuel for another, with associated energy loss, so does little to meet national energy needs. For high purity needs, a small amount of hydrogen is produced by electrolysis, but this again is only as good as the energy source used to produce the electricity used. Electrolysis can electrochemically split water into hydrogen and oxygen in essentially the reverse of the reaction in a fuel cell. To make sense for large-scale use, this process must use an inexpensive source of electricity. Because wind energy is currently the lowest cost renewable energy, it is the leading candidate. It is also a variable source that would benefit from being able to produce hydrogen when its electricity is not needed and to add fuel-cell generation when electricity demand exceeds what the wind turbines can provide. The combination also benefits because electrolyzers require direct current and wind turbine power is produced as direct current before conversion back to alternating current suitable for the electric grid [4].

In recent years, considerable research and interest has been given to the potential use of hydrogen as a nonpolluting energy source, in particular, the use of hydrogen in chemical fuel cells, in which the gas reacts with the oxygen from ambient air to directly produce an electric current. Although this has an immediate appeal for environmentalists, it remains to be seen whether the total production and use cycle can really achieve a lower environmental impact than the existing energy production methods.

Based on current production methods, which involve steam reforming of natural gas, hydrogen cuts carbon dioxide emissions for each kilometer driven by up to 30% compared with conventional petrol and diesel cars. A study sponsored by a European Union Commission and completed in 2010 shows that the 95% reduction target can be achieved by 2050 with the help of various hydrogen production methods, with an increasing share of renewable sources in the mix.

In its “balanced” scenario, the study assumes that fuel-cell vehicles will account for 25% of all cars on the road by 2050. This requires an investment of around three billion euros in a Europe-wide hydrogen refueling infrastructure by 2020. Expressed as a percentage of the total cost of ownership, this corresponds to a mere 5% [5].

Although the actual use of hydrogen in a fuel cell produces only energy and water as a by-product, it is in the actual production of hydrogen as previously discussed that the environmental effects of carbon dioxide and carbon monoxide releases are realized. With the current ability to use fuel cells with existing fuel sources such as methane and methanol, the pursuit of a purely hydrogen fuel cell may not be the environmental solution that is often implied by its proponents. However, the increased efficiency and use of renewable resources (including solar) to generate hydrogen may overcome this barrier.

It is estimated that the global hydrogen production capacity required to replace the entire current fossil fuel-based energy system would need to be of the order of about 2,500 Tg H_2 yr^{-1} [7]. The United States used about 177 billion gallons of gasoline for transportation in 2009 [8].Using one type of commercial electrolyzer, one can make 0.38 kg of gaseous hydrogen per gallon of water. Conversion of the current US light-duty fleet (some 230 million vehicles) to fuel-cell vehicles would require about 110 billion gallons of water/year to supply the needed hydrogen. For comparison, the United States uses about 300 billion gallons of water/year for the production of gasoline [9], about three times the amount needed for hydrogen. However, that amount is used and recycled, not consumed as it is in the hydrogen production process. Domestic personal water use in the United States is about 4,800 billion gallons/year.

For an estimate of the amount of water needed for hydrogen-powered fuel-cell vehicles, assume a vehicle fuel economy of 60 miles per kg of H_2 (the Honda FCX Clarity gets 72), that vehicle miles traveled $= 2.6 \times 10^{12}$ miles/year, and that you can make 0.38 kg of H_2 from 1 gal of water using Proton Energy System’s H6m Electrolyzer. Total water required for the US fleet = (2.6×10^{12} miles/year)(1 kg of H_2/60 miles)(1 gal H_2O/0.38 kg of H2) $= 1.1 \times 10^{11}$ gal of H_2O/year. This represents the water used directly for fuel. If one considers all water uses along the chain; for example, from construction of wind farms to the electrolysis systems (life cycle assessment), then the total water use would be in the range of 3.3×10^{11} gal H_2O/year.

For a detailed description of the electrolytic process of hydrogen generation, also see Chap. 22. The energy required to produce hydrogen at atmospheric pressure via electrolysis (assuming 1.23 V) is about 32.9 kWh/kg. A kilogram is about 2.2 lb. For 1 mole (2 g) of hydrogen the energy is about 0.066 kWh/mole. Compressing or liquefying the hydrogen would take additional energy. One company produces hydrogen through electrolysis at about 7,000 psi at an energy usage of about 60 kWh/kg Hydrogen.

Because a Watt is Voltage × Current, this is equivalent to Power × Rate × Time. The power in this case is the voltage required to split water into hydrogen and oxygen (1.23 V at 25°C). The rate is the current flow and relates directly to how fast hydrogen is produced. Time, of course, is how long the reaction runs. It turns out that voltage and current flow are interrelated. To run the water splitting reaction at a higher rate (generating more hydrogen in a given time), more voltage must be applied (similar to pushing down on the accelerator of a car; more gas is used to make the car go faster.) For commercial electrolysis systems that operate at about 1 A/cm^2, a voltage of 1.75 is required. This translates into about 46.8 kW-hr/kg, which corresponds to an energy efficiency of 70%.

Lowering the voltage for electrolysis, which will increase the energy efficiency of the process, is an important area for research [11].

The one possible solution to effective renewable energy from water may be in the current research centered around catalytic thermal decomposition of water using solar energy. In this process, first developed by the Solar Hydrogen Energy

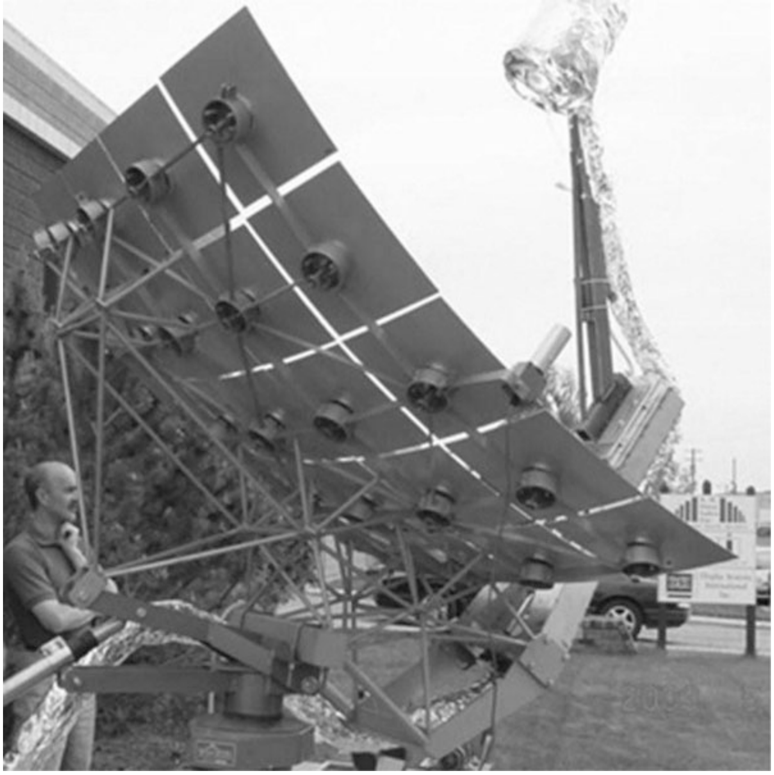


Fig. 27.7 Solar Hydrogen Catalytic Generator—small-scale prototype. (Courtesy of SHEC LABS.)

Corporation (SHEC) in Canada and now pursued in multiple laboratories, including the Global Photonic Energy Corporation in the United States, water is catalytically broken down into hydrogen and oxygen. The unique application of solar energy as the driving force, enhanced by the catalytic system, bypasses the usual thermodynamic barrier of standard electrolysis, which requires more energy (from a fossil fuel source) to generate the hydrogen than the energy that can be obtained from the hydrogen in a fuel cell (Fig. 27.7).

Other options proposed, such as the use of ammonia as a fuel, are still based on the production of hydrogen as the primary feedstock for the alternative fuels [12].

There are many applications for hydrogen. It is used to make fertilizers, as the active ingredient in detergents, and in the manufacture of polyurethanes. It finds application in the production of semiconductor devices, in the process of refining crude oil, in food processing for edible oils and shortening, and as fuel for the space shuttle. The hydrogenation of edible oils in soybeans, fish, cottonseed, and corn produces solids used in shortening and other foods. Many alcohols also are produced by the hydrogenation of the corresponding acids and aldehydes. It is possible to use hydrogen in low concentrations as a physiologically inert gas for a breathing atmosphere where multiple gases are involved. Hydrogen is an important gas in several semiconductor-manufacturing stages. It is used as a protective atmosphere in silicon crystal growth, as a reactant and carrier gas during semiconductor chip manufacturing, and as a reducing atmosphere for sintering and the bonding of leads to the finished chip [1].

The world economy currently consumes about 42 million tons of hydrogen per year. About 60% of this becomes feedstock for ammonia production and subsequent use in fertilizer (ORNL, 2003). Petroleum refining consumes another 23%, chiefly to remove sulfur and to upgrade the heavier fractions into more valuable products. Another 9% is used to manufacture methanol (ORNL, 2003), and the remainder goes for chemical, metallurgical, and space purposes (Holt 2003) [14].

Some recent worldwide hydrogen production totals are shown below [14]:

Origin	Amount in billions (Nm_3/year)	Percent
Natural gas	240	48
Oil	150	30
Coal	90	18
Electrolysis	20	4
Total	500	100

Each year, the United States uses more than nine million tons (about 90 billion normal cubic meters, 3.2 trillion standard cubic feet) of hydrogen, 7.5 million tons of which are consumed at the place of manufacture. The remaining 1.5 million tons are considered to be “merchant” hydrogen, or hydrogen that is sold. Today, most of this hydrogen is used as a chemical, rather than a fuel, in a variety of commercial applications [14]:

- Commercial fixation of nitrogen from the air to produce ammonia for fertilizer (about two-thirds of commercial hydrogen is used for this)
- Hydrogenation of fats and oils, in which vegetable oils are changed from liquids to solids; shortening is an example of a hydrogenated oil
- Methanol production, in hydrodealkylation, hydrocracking, and hydrodesulphurization
- Welding
- Hydrochloric acid production
- Metallic ore reduction
- Cryogenics and the study of superconductivity (liquid hydrogen)
- Preventing oxidation in the manufacturing of semiconductors
- Cooling turbines (hydrogen transfers heat very well)
- Hydrogen's main use as a fuel is in the space program. Today hydrogen fuels both the main engine of the Space Shuttle and the onboard fuel cells that provide the Shuttle's electric power.

Merchant hydrogen will remain the largest and fastest growing product in the US refinery chemical market. Advances will be driven by tightening sulfur standards for diesel fuels. Such environmental regulations promote the use of hydrotreating as a means of removing sulfur and other contaminants. As of 2010, diesel fuel must meet a 15 part-per-million (ppm) sulfur limit, and going forward, significant sulfur reductions are also expected for heating oil. Hydrocracking represents another growth application for merchant hydrogen, as US refineries continue to expand their hydrocracking capacity in efforts to boost gasoline and diesel fuel yields [16].

In the metals industry, hydrogen is used for applications such as annealing metals. It is also used as a fuel in underwater oxy-hydrogen torches and with oxy-hydrogen welding and cutting systems. Hydrogen serves as a nonoxidizing shield alone or with other gases in furnace brazing and in welding aluminum, magnesium, and lead. The temperature in the oxy-hydrogen flames is about 4,000°F, which is suitable for low-temperature welding and brazing. Oxy-hydrogen flames are used in the fabrication of quartz and glass, and hydrogen is used as a protective atmosphere in the float glass process whereby molten glass is floated on a pool of liquid tin. Atomic hydrogen welding is suitable for very thin stock and can be used with virtually all nonferrous alloys. In the process, an arc with a temperature of about 11,000°F is maintained between two nonconsumable metal electrodes. Molecular hydrogen fed into the arc is transformed into atomic hydrogen, which sends heat from the arc to the weld zone. At the surface of the welding area the atomic hydrogen recombines with molecular hydrogen, with the release of heat.

Liquid hydrogen is very important as a fuel for powering missiles and rockets. It is used in laboratory research on the properties of some materials at cryogenic temperatures, work often associated with studies of superconductivity. Liquid hydrogen can be used as a fuel, for either propulsion or heating.

Hydrogen gas is shipped in high-pressure gas cylinders and high-pressure cylinder trailers. Liquid hydrogen is shipped in insulated portable containers and in cryogenic tankers. The normal liquid hydrogen supply system utilizes a cryogenic stationary storage tank rated at 150 psig, just below the critical pressure of hydrogen. A cryogenic pumping station, which has a liquid hydrogen storage tank, a high-pressure cryogenic pump capable of supplying liquid hydrogen at 3,000 psig, and high-pressure gas storage tubes, often is used to supply large quantities at pressure above 150 psig.

Helium

Helium was discovered during a solar eclipse as an orange line in the spectrum of the sun's atmosphere. Physically, helium is colorless, odorless, and tasteless; and it is the second lightest element, second only to hydrogen. Helium is chemically inert and has only a 5 ppm concentration in the atmosphere (0.0005%).

Small amounts of helium are extracted from the atmosphere by fractionation methods, but not commercially because of the small amount of helium in the atmosphere. A number of natural gas wells contain helium, which can be recovered by a liquefaction and stripping process. Natural gas containing at least 0.2% helium has been found in the American Southwest, where the natural gas fields are the major US source of helium. Those helium-rich fields are within 250 miles of Amarillo, Texas; other helium-bearing fields have been found in Saskatchewan, Canada, and in areas near the Black Sea.

Helium cannot be synthesized; so conservation and cleanup recycle systems for "spent" gas are important means of preserving the earth's helium resources. One of helium's first uses was as a nonflammable replacement for hydrogen for inflation of lighter-than-air aircraft. Today liquid helium is vitally important in cryogenic research, as it is the only known substance to remain fluid at temperatures near absolute zero. It has a unique use as a refrigerant in cryogenics and is the only nuclear reactor coolant that does not become radioactive. Liquid helium is used extensively with superconducting magnets; in the medical field, it is used to cool the superconducting magnets for magnetic resonance imaging (MRI). It is used to purge and pressurize NASA vehicle liquid hydrogen tanks because it is the only gas that remains a vapor at liquid hydrogen temperatures.

Helium has a variety of other uses in welding and lighting. It is used as an inert gas shield in arc welding, for filling cold weather fluorescent lamps, and to trace leaks in refrigeration and other closed systems. Also helium is used for specialized purging and pressurizing applications. Helium can be used in place of or together with nitrogen as the inert portion of the atmosphere in sealed environments such as space vehicles, deep-sea submergence vehicles, and diving suits. The use of helium shield gas in a plasma arc furnace is a popular method for melting precious metal ores and scrap. Helium is very important for use in lasers and fiber-optic production, as well as in the production of germanium and silicon crystals for semiconductors.

Gaseous helium is stored and transported in high-pressure gas cylinders, high-pressure gas trailers, and portable cylinder modules. Liquid helium is transported in liquid cylinders, portable tanks, and cryogenic tankers. Because of the extremely low temperature of liquid helium and the cost of producing it, special design considerations must be given to storage containers in order to maintain the helium in the liquid phase. Some containers are designed to include a sacrificial liquid nitrogen shield that intercepts heat before it is transferred into the liquid helium.

Carbon Dioxide

The carbon dioxide market in the United States is served by gaseous, liquid, and solid carbon dioxide. The consumption of carbon dioxide gas is approximately 25 million tons per year with the enhanced oil recovery and urea production markets being the primary consumers. Liquid and solid consumption is about 7.5 million tons per year, serving primarily the food and beverage processing industries [21]. Carbon dioxide is used extensively in a variety of areas. It is used for pressurizing, and as a source of the bubbles and the acidic taste in soft drinks. Carbon dioxide is used to fill a type of fire extinguisher that literally depends on the inertness of the compound. Carbon dioxide is also used in freezing specialty and quality foods. Supercritical carbon dioxide (a dense, high pressure, single-phase form) is finding many new applications in pharmaceutical processing, plastics recycling, dry cleaning, flavor and fragrance extractions, and other solvent-based processes [1].

More specifically, each physical phase of carbon dioxide has uses in many areas. Gaseous carbon dioxide is used for pH control in water treatment and as a growth stimulant for plant life. It has been used successfully as a grain storage fumigant against pests and as a wide area mosquito lure to enhance chemical spraying effectiveness. Alone and in combination with other gases, carbon dioxide in food packaging extends the freshness and shelf-life of many products. It is second only to argon as a welding shield gas, and is used for flammable tank purging and inert blanketing of reactive liquids. Liquid carbon dioxide has multiple applications as a rapid, controllable refrigerant. It is used in one case as an expendable refrigerant for low-temperature testing of aviation, missiles, and electronic components. Carbon dioxide is also used in controlling chemical reactions and for stimulation of oil and gas wells. It is used extensively in food chilling and freezing applications, both in processing and in transportation.

Solid carbon dioxide, commonly called “dry ice,” is used extensively as a refrigerant for dairy products, meat products, and other frozen foods while in transit. It is also used as a cooling agent in many industrial processes such as grinding heat-sensitive materials, cold-treating metals, shrink-fitting machinery parts, in vacuum cold traps, in cryo-surgery, and specimen preservation and storage. It is used as a residue-free abrasive cleaner for many industrial equipment cleaning applications through the use of special air-driven particle blasting machines.

Carbon dioxide is usually nonreactive and nontoxic. At normal atmospheric pressure and temperature, it is colorless and odorless. Carbon dioxide will not burn, nor will it support combustion. Unrefined carbon dioxide gas is typically obtained from the combustion reaction of coal, coke, natural gas, and other carboniferous fuels. Other major sources are ethanol fermentation plants and the development of landfill gases as renewable sources.

The gas obtained is liquefied and purified by several different processes to a purity of about 99.99%. The major source of carbon dioxide is as a by-product of steam–methane reforming. The resulting reformer syngas may utilize one of the following three processes in the production of carbon dioxide. Recovery of pure carbon dioxide from reformer syngas has most often been accomplished through absorption by liquid solvents. The solvents used are monoethanolamine (MEA), diethanolamine (DEA), and methyl diethanolamine (MDEA).

A hot potassium carbonate solution may also be used in some applications (see Fig. 27.8 for the process flow diagram). In this process, the syngas stream from the reformer is sent to the stripper column where the carbon dioxide is absorbed by the solvent. The residual gas stream components are discharged from the stripper column. The carbon dioxide in solvent is sent to the regeneration column where the carbon dioxide is regenerated and discharged for further purification by compression dehydration and

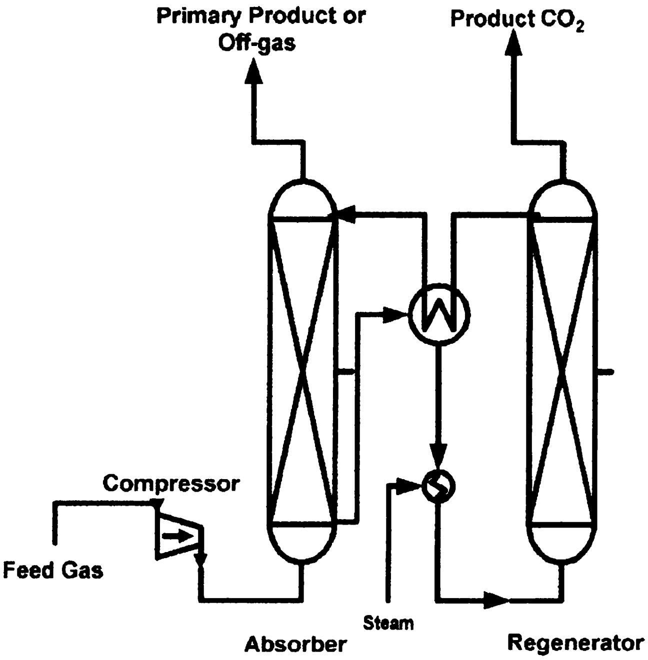


Fig. 27.8 CO_2 recovery by solvent adsorption process. (Courtesy Process Systems Consulting, Inc.)

refrigeration steps. The solvent is returned to the stripper column. The final product is essentially 99.99% pure.

Carbon dioxide may be manufactured at 95% purity by the membrane separation process from a similar gas or from natural gas [18, 19, 22]. In the recovery of high-purity carbon dioxide from the syngas or natural gas, and before any separations can occur, a pretreatment step removes any contaminants and any condensate liquids that may be present. The feed gas then enters the membrane system at 300–1,000 psig. The permeate gas is delivered at pressures between 50 and 150 psig and at a carbon dioxide purity of about 95%. The high-pressure residual gas containing 10–30% carbon dioxide enters a conventional treatment process such as MEA or DEA for a final cleanup. The carbon dioxide off-gas from the conventional treatment process is then elevated to the same pressure as that of the permeate gas from the membrane systems.

The combined streams may be further compressed to higher pressures, such as 2,000 psig for applications in carbon dioxide flooding in enhanced oil recovery. The third process used in the production of carbon dioxide is PSA. The feed gas usually contains approximately 20% carbon dioxide, 70% hydrogen, and the remainder methane, carbon monoxide, nitrogen, and water. The feed gas is typically under a pressure of 125–400 psig at temperatures of 80–120°F. The carbon dioxide and water are strongly adsorbed in the adsorb beds and the residual gas stream is depressurized for further recovery. The adsorber vessel is then evacuated through vacuum blowers where the carbon dioxide, which has been adsorbed by the bed, is released at purities of essentially 99%.

Carbon dioxide produced from ethanol fermentation plants or landfill gas may be recovered with similar processes. Unique to the fermentation plant is the ability to recover the carbon dioxide directly from the ethanol distillation tower, followed by a secondary water wash. Final purification and liquefaction stages then follow the normal process flow. Landfill gas recovery is unique in requiring essentially the removal of the methane and trace impurities [10, 23]. Several processes exist to provide two gas products (via membrane separation), a liquid methane and gaseous carbon dioxide stream (via the CryoFuel® system), or a gaseous methane and liquid carbon dioxide stream (via the Acrion® process). The Acrion® system is unique in that it uses the carbon dioxide as part of the purification process instead of relying on additional chemical treatments. Figure 27.9 shows the process flow diagram for the Acrion® system. Liquid carbon dioxide is normally stored in a foam-insulated tank that has a refrigeration unit to maintain the tank temperature and pressure.

Liquefied Natural Gas

(For Methane, or Compressed Natural Gas, See Chap. 20, Methane)

The primary advantage of LNG over gaseous natural gas is the cost and space savings due to the tremendous change in the product volume. Over 600 standard cubic feet of natural gas occupy only 1 ft^3 of space in its liquid form at −260°F and atmospheric pressure. This significant volume reduction reduces the storage and transport volume, and allows shipment of natural gas in areas where pipelines were never feasible.

The primary vehicle for transporting LNG is the cryogenic tanker. It also is transported in specially designed ships. LNG comes primarily from natural underground reservoirs. Significant quantities are now being recovered from both landfill sources and bioreactor sources around the world. LNG is composed predominantly of methane, which may contain minor quantities of ethane, propane, nitrogen, helium, and other components normally found in natural gas.

In the past the liquefaction of natural gas used a classic cascade cycle. The process required 120,000 hp for liquefaction of over 150 million standard cubic feet (mmscf) per day. Provisions are made for some of these cycles to use seawater for cooling. Later, baseload LNG plants utilized mixed refrigerant cycles, such as Air Products and Chemicals, Inc.'s propane precooled mixed refrigerant system. Baseload plant capacities range from about 70 mmscf/day to about 350 mmscf/day of LNG. Baseload plants move LNG from remote sites by ship to populated

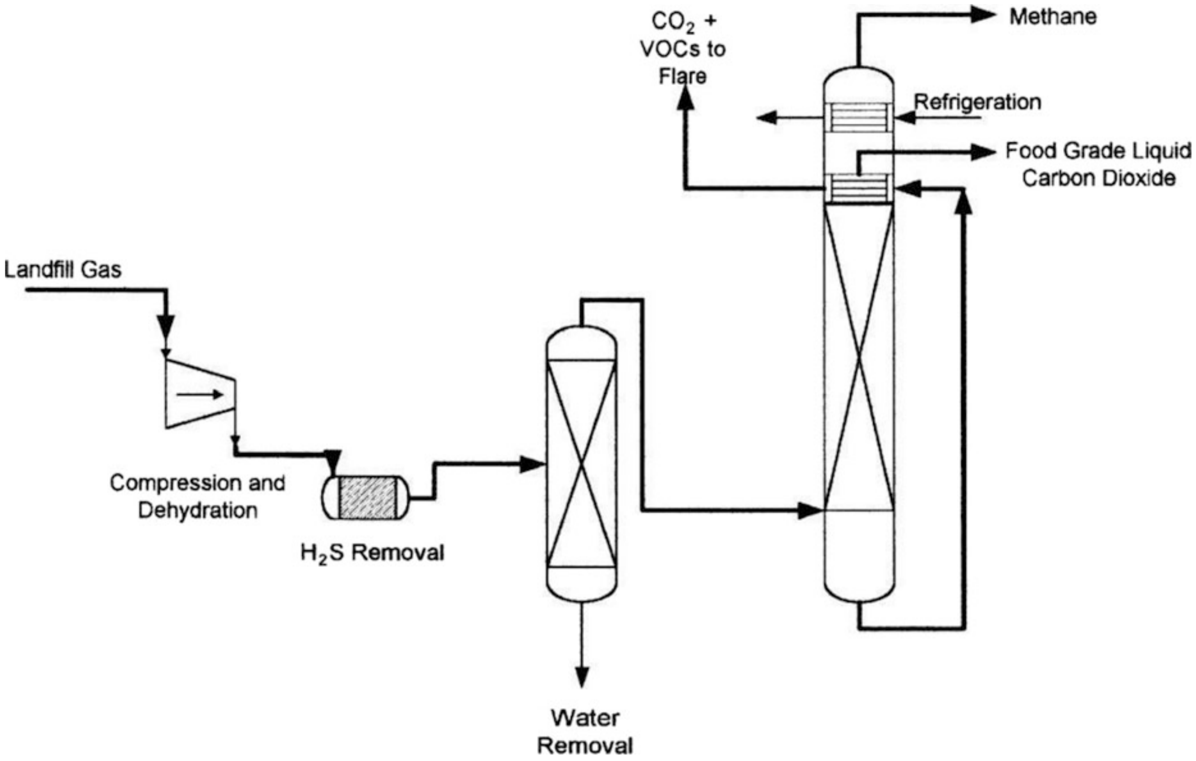


Fig. 27.9 Acrion CO_2 wash process flow diagram. (Courtesy Acrion Technologies, Inc.)

areas. For example, Indonesia and the United Arab Emirates supply LNG to Japan for electric power generation [10].

At baseload LNG production plants, natural gas enters the plant at near 85°F and 600 psig. The gas is treated to remove carbon dioxide and any sulfur compounds, and it is cooled with high-level propane refrigerant to remove as much water as possible. Dehydration to a dew point of about −100°F is done by carrying out regeneration with dried outlet gas by way of a high-pressure steam heater and a regeneration gas compressor. The dry, treated gas is passed through a stationary bed of activated carbon for removal of mercury, and is cooled with two successively lower levels of propane refrigeration. A scrub column flows to a fractionation train, which consists of a deethanizer, a depropanizer, and a debutanizer where relatively pure ethane and propane are produced. Excess ethane, propane, and butane from the fractionation process are reinjected into the main gas stream just prior to liquefaction. The overhead from the scrub column flows to the main heat exchanger for liquefaction to LNG. Figure 27.10 shows the process flow for an LNG plant [24]. For years LNG has been used for peakshaving purposes, which is the storage of excess capacity of LNG to be revaporized for the coldest days of the year [25]. In recent years LNG has been used in heating and vehicle propulsion. Also LNG is used to produce carbon black, which has applications in the manufacture of rubber products and printing ink. The burning of high-purity LNG (methane) is done to make carbon black for particular use in electronic devices.

Acetylene

Acetylene (C_2H_2) is a colorless and flammable gas. At 100% purity, acetylene is odorless, but at commercial purity it has a garlic-like odor. Acetylene can be liquefied and solidified with ease, although in both cases it explodes with extreme violence when ignited. Acetylene can be inhaled in rather high concentrations without chronic, harmful effects. In fact, it has been used as an anesthetic. However, it is a simple asphyxiant if present in concentrations high enough to deprive the lungs of oxygen and produce suffocation [1].

A primary method for the manufacture of acetylene is to react calcium carbide, the principal raw material, with water to produce acetylene with calcium hydroxide as a by-product:

$$CaC_2 + 2H_2O \rightarrow C_2H_2 + Ca(OH)_2$$

Acetylene also is manufactured by thermal or arc cracking of hydrocarbons and by a process utilizing the partial

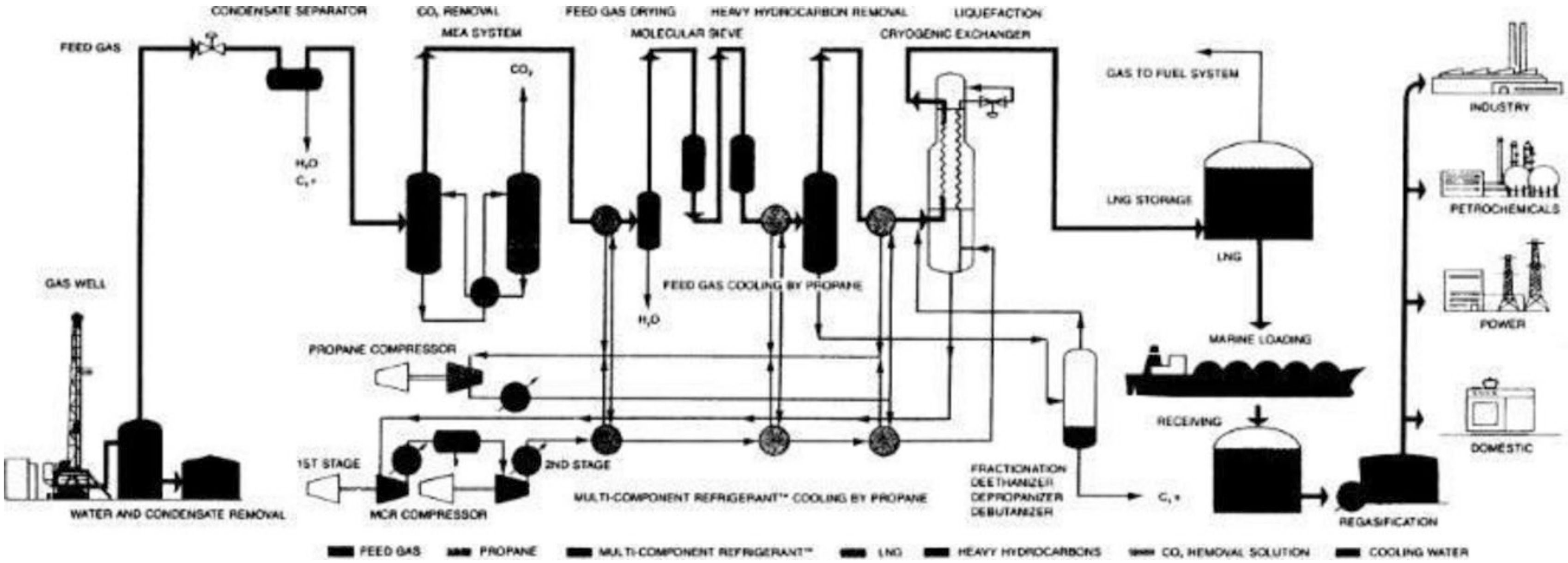


Fig. 27.10 Schematic of process flow for a baseload LNG plant. (Courtesy Air Products and Chemicals, Inc.)

combustion of methane with oxygen. The acetylene produced from calcium carbide is basically pure, with only water and air as impurities. Eighty percent of the acetylene produced is used for chemical synthesis.

Acetylene is an important raw material for a series of organic compounds such as acetaldehyde, acetic acid, acetic anhydride, acetone, and vinyl chloride. Those compounds are used to manufacture a large group of products, including plastics, synthetic rubber, dyestuffs, solvents, and pharmaceuticals, and in the manufacture of carbon black. The other 20% is used for oxy-acetylene welding, cutting, heat treating, lighting purposes, buoys, and beacons [1].

In the primary method for producing acetylene, calcium carbide is fed into a generator containing water. The resulting wet acetylene gas at 120–140°F is piped to a cooler condenser where it is cooled to 80–90°F and the water vapor is condensed. This is the first step in the drying process. The gas then goes to a low-pressure dryer filled with calcium chloride. The calcium chloride absorbs the moisture from the acetylene gas, which then enters a compressor, where the pressure is increased to 200–375 psig. An oil separator filters out excess impurities and oil. The gas next goes to a high-pressure dryer, where all remaining water is removed. The gas is then pumped into cylinders for shipment. The by-product, calcium hydroxide, is sent to a storage tank, where water is decanted off the top, and is recovered.

The only method of distributing acetylene, other than pipeline, is by means of portable steel tanks containing a porous solid filler saturated with acetone or other suitable solvent, in which the acetylene is dissolved under pressure. Acetylene alone is not handled at pressures higher than 30 psig because of its tendency to decompose explosively. Dissolved in acetone it may be stored at pressures of 150–250 psig.

Nitrous Oxide

Nitrous oxide (N_2O), commonly known as laughing gas, is a colorless and virtually odorless, tasteless, nontoxic gas. Nitrous oxide is an oxidizer that will support combustion, but it is a nonflammable gas. It is water-soluble; but when it is dissolved, its acidity does not change. The primary commercial method for producing nitrous oxide is by thermal decomposition of ammonium nitrate, with nitrous oxide and water in the primary reaction:

$$NH_4NO_3 \rightarrow N_2O + 2H_2O$$

The impurities formed are mostly high oxides of nitrogen and are highly toxic. After the water is removed in a condenser, the gas is washed in a solution of potassium dichromate to remove nitric oxide, in caustic to remove nitric acid, and finally in water. Nitrous oxide is an inhalant type of anesthetic or analgesic gas. It serves as a propellant for various aerosol products, particularly with foods such as whipped cream; and it is used as an oxidizing agent in blow-torches, atomic spectroscopy instruments, and in the manufacture of various compounds, both organic and inorganic. It also serves as fuel oxidant in rocket fuel and high-performance vehicles, and it is used as part of the working fluid in hypersonic wind tunnels that have been investigated recently. It has applications in cryosurgery similar to nitrogen [1].

Nitrous oxide is contained and transported in its liquid phase in high-pressure gas cylinders or in liquid cylinders. It is transported as a liquefied compressed gas under high pressure in cylinders and at lower pressures and reduced temperatures in refrigerated cargo tanks and insulated portable tanks. Nitrous oxide is stored in a foam-insulated tank accompanied by a refrigeration unit, which is similar to the tank used for carbon dioxide.

References

1. Compressed Gas Association (1999) Handbook of compressed gases, 4th edn. Springer, New York
2. Kreith F, West R (2004) Fallacies of a hydrogen economy: a critical analysis of hydrogen production and utilization. J Energy Resour Technol 126:249. doi:10.1115/1.1834851
3. Compressed Gas Association (1966) Handbook of compressed gases, 4th edn. Springer, New York
4. National Renewable Energy Laboratory, U.S. Dept. of Energy, Office of Energy Efficiency and Renewable Energy, http://www.nrel.gov/learning/eds_hydro_production.html Accessed 2010-12-13.
5. NOW GmbH (2010) A portfolio of power-trains for Europe: a fact-based analysis. P.11, European Fuel Cells and Hydrogen Joint Undertaking. NOW GmbH, Berlin
6. Vance R, Weinstock H (1969) Applications of cryogenic technology. Tinnon-Brown, Los Angeles
7. Turner JA (2004) Sustainable hydrogen production. Science 305(5686):972–974
8. U.S. Energy Information Administration. Annual Energy Review (Figure 5.13b and Table 5.13c, p.153 2009). http://www.eia.gov Accessed 2010-10-15
9. Derwent R, Simmonds P, O'Doherty S, Manning A, Collins W, Stevenson D (2006) Global environmental impacts of the hydrogen economy. Int J Nuclear Hydrogen Production and Application 1(1):65
10. Speir GA (1978) Indonesia's Badak LNG project sets new records. Pipeline & Gas J. (205): 22–24, 26, 28
11. This is a life cycle analysis (M. Mann and M. Whitaker, unpublished data). National Hydrogen Association, http://www.hydrogenassociation.org/general/.
12. NH_3 _ The Other Hydrogen. Norm Olsen, PE, John Holbrook, PhD, Paper presented at the National Hydrogen Association Annual Conference, March 30, 2009. Accessed from http://www.energy.iastate.edu/Renewable/ammonia/index.htm2010-12-03.
13. Castro RP, Cohen Y, Monbouquette YG, in *Proceedings of the 1994 National Conference on Environmental Engineering*, p. 82, Boulder, CO, July 1994.

14. U.S. Department of Energy http://www.eere.energy.gov/ Accessed 2010-10-15
15. *Hydrocarbon Processing*, Gulf Publishing Co., Houston, April 1998.pp. 107–108
16. June 2010 HP In: Brief hydrocarbon processing, http://www.hydrocarbonprocessing.com/ Accessed 2010-06-15
17. Depuis GE et al (1993) Hydrogen membrane recovery estimates. In: Hydrocarbon processing. Gulf, Houston, pp. 61–64.
18. *Hydrocarbon Processing*, pp. 120–124, Gulf Publishing Co., Houston, April 1998
19. Zolandz RW, and Fleming GK in *Membrane Handbook*, WSW Ho, and KK Sirkan (Eds.), Chapter 3, Van Nostrand Reinhold, New York, 1992
20. Humphrey JL, Keller GE (1997) *II: separation Process Technology*, Chapter 5, McGraw-Hill, New York,
21. Marshall TL, Garvey MD (2000) Special report. CryoGas Internat, November
22. Singh R (1998) Industrial membrane separation processes. Chemtech 28(4):33–44
23. Thorneloe, Cosulick, Pacey, and Roqueta (1997) Landfill gas utilization—Survey of United States Projects. Presented at the Solid waste association of North America's twentieth annual international landfill gas symposium, Monterey, CA, 25–27 March 1997. *Conference Proceedings*. EPA-ORD
24. *Hydrocarbon Processing*, p. 112, Gulf Publishing Co., Houston, April 1998.
25. *Hydrocarbon Processing*, p. 120–124, Gulf Publishing Co., Houston, April 1998.

28

Wood and Wood Products

Yuan-Zong Lai

Introduction

Wood along with the outer portion of tree trunk (bark) [1] (Fig. 28.1) represents the bulk of forest biomass materials and has been utilized by humans since antiquity. Trees provided a source of many products required by early humans such as food, medicine, fuel, and tools. For example, the bark of the willow tree, when chewed, was used as a painkiller in early Greece and was the precursor of the present-day aspirin. The leaf extracts and nuts from Gingo trees have been used in traditional Chinese medicines for thousands of years. Wood served as the primary fuel in the United States until about the turn of the nineteenth century, and even today over one-half of the wood now harvested in the world is used for heating fuel.

Throughout history different types of wood have served many purposes. The tough, strong, and durable white oak, for example, was a well-proven raw material for ships, bridges, cooperage, barn timbers, farm implements, railroad ties, fence posts, flooring, paneling, and other products. In contrast, woods such as black walnut and cherry became primarily cabinet woods. Hickory was manufactured into tough, hard, resilient tool handles. Black locust was used for barn timbers and treenails. What the early artisan learned by trial and error became the basis for intelligent decisions concerning which species was best suited to a given purpose, as well as what characteristics should be looked for in selecting trees for different applications. It was known that wood from trees grown in certain locations was stronger, more durable, and more easily worked with tools than wood from the same species grown in other locations. Modern wood quality research has substantiated that location and growth conditions significantly affect wood properties [2, 3].

In the United States more than 100 kinds of wood [1, 3] are available to the prospective user, but it is very unlikely that all are available in any one locality. About 60 native woods are of major commercial importance. Another 30 wood types are commonly imported in the form of logs, cants, lumber, and veneer for industrial uses, the building trades, and crafts.

The wood from many different species of trees is generally divided into two major categories according to the botanical classification of seed plants. The softwoods are derived from the gymnosperms (naked seeds), generally characterized as cone-bearing trees with persistent needle-like foliage; the hardwoods are from the angiosperms (encased seeds), which are generally the deciduous broad-leaved trees. The majority of the trees in the former classification fall into the division Coniferophyta and thus are often referred to as conifers. However, woods are often designated based on their physical appearance or wood structure (Table 28.1). Thus, softwoods are also called conifers, evergreen, or nonporous wood while hardwood are called broadleaves, deciduous or porous wood.

It is important to note that the terms "hardwood" and "softwood" have no direct relation to the hardness or softness of the wood. In fact, hardwood trees such as cottonwood, aspen, and balsa have softer wood than the western white pines and true firs; and certain softwoods, such as longleaf pine and Douglas fir, produce wood that is much harder than that of basswood or yellow poplar. Although typical softwood has needlelike foliage, there are exceptions. For example, Gingo tree is softwood in wood structure and has the appearance just like a typical deciduous hardwood.

The importance of wood as a raw material supplying fiber, energy, and chemicals is similar in magnitude to its use as a solid material. Lumber, plywood, and reconstituted boards consume about one-half of the timber harvest; usage for fiber, chemicals, and fuels accounts for the remaining half. Fiber production consumes about 25% of the total timber use. Although the relative value of wood as a source

---

Y.-Z. Lai (✉)
Department of Paper and Bioprocess Engineering, SUNY College of Environmental Science and Forestry, Syracuse, NY, USA
e-mail: yzlai@syr.edu

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_28, © Springer Science+Business Media New York 2012

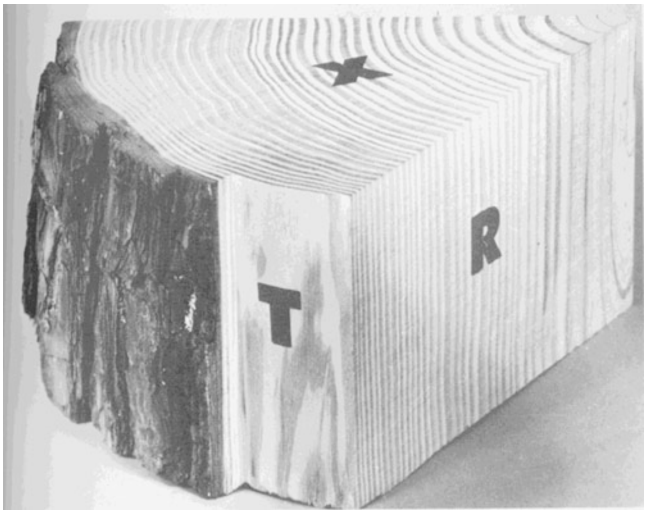


Fig. 28.1 Three different planes of wood and their usual designation: X = transverse surface or cross-section; R = radical surface; T = tangential surface (from Core et al. [1], Copyright © 1976, with permission of the copyright owner)

Table 28.1 Terms used to designate softwood and hardwood

	Softwood	Hardwood
Plant origin	Gymnosperms Conifers	Angiosperms
Leaf type	Needlelike Evergreen	Broadleaves Deciduous
Cell type	Nonporous	Porous

of energy and chemicals has varied considerably through the decades, it has become a major emphasis in recent years to cope with the increasing global demand of energy and the soaring price of fossil fuels. These have stimulated a great deal of interest and developments in the production of biofuels and bioproducts under the umbrella of biorefinery [4, 5]. The many uses for wood are summarized in Fig. 28.2. This chapter discusses the chemical and physical nature of wood, the technology for conversion of wood to pulp fibers for making paper and related products, to wood-based composites, to fuels for energy, and to chemicals for industry and consumers.

Chemical Composition of Wood

Wood is composed of mainly three major polymers (cellulose, hemicelluloses, and lignin) and some extraneous substances called extractives. Table 28.2 lists the variation of these components in softwood and hardwood for five wood species each [6]. The combined polysaccharide components (cellulose plus hemicelluloses) often termed as holocellulose account for about 75% of the wood materials. Major uses of these polysaccharide components have been in the form of fibers for making paper, paperboard, and other fiber products by the paper industry. Recently, the potential utilization of woody polysaccharides as feedstock for conversion to biofuels and chemicals has been examined quite extensively [4, 5].

Higher plants commonly are referred to as lignocellulosic materials because of the typical joint occurrence in them of lignin and cellulose. Lignin being a phenolic polymer is often considered as the gluing or encrusting substance of wood cells and adds mechanical strength or stiffness to the tree and wood. Lignin usually plays a negative role in the chemical utilization of woody materials and must be modified, partially degraded, or completely removed depending on the end uses of the final products. These lignin modification or delignification reactions constitute the commercial pulping and bleaching operations in producing papermaking fibers.

Cellulose

Cellulose is the key component occurring in the cell wall. It is a linear polymer of β (1→ 4)-D-glucan (Fig. 28.3) with a degree of polymerization (DP) of approximately 10,000. Cellulose molecules, because of their unique structure, have a strong tendency to form extensive hydrogen bondings. The molecules could align together in a highly ordered fashion to form a crystalline region, whereas the less ordered molecules being noncrystalline are usually termed amorphous materials.

The cellulose molecules in nature are organized into *elementary fibrils*. Each one has an average width of 3.5 nm and contains about 36 parallel cellulose molecules held together by hydrogen bonding. They are then aggregated into strands called *microfibrils*, which are 5–30 nm in width and contain both the crystalline and amorphous regions. *Microfibrils* are further organized into fibrils and then fibers.

The proportion of the ordered and disordered regions of cellulose (crystallinity) varies considerably with its origin. The crystallinity for wood pulps is about 60% and higher for cotton cellulose (73%). Although the crystalline cellulose contributes to the strength of cellulose polymer, it is a key factor causing the recalcitrance of woody biomass to saccharification for the biofuel production.

Hemicelluloses

Closely associated with cellulose in the wood structure and paper products are the other polysaccharides called hemicelluloses, which consist of both glucose and nonglucose units. They are amorphous and slightly branched polymers.

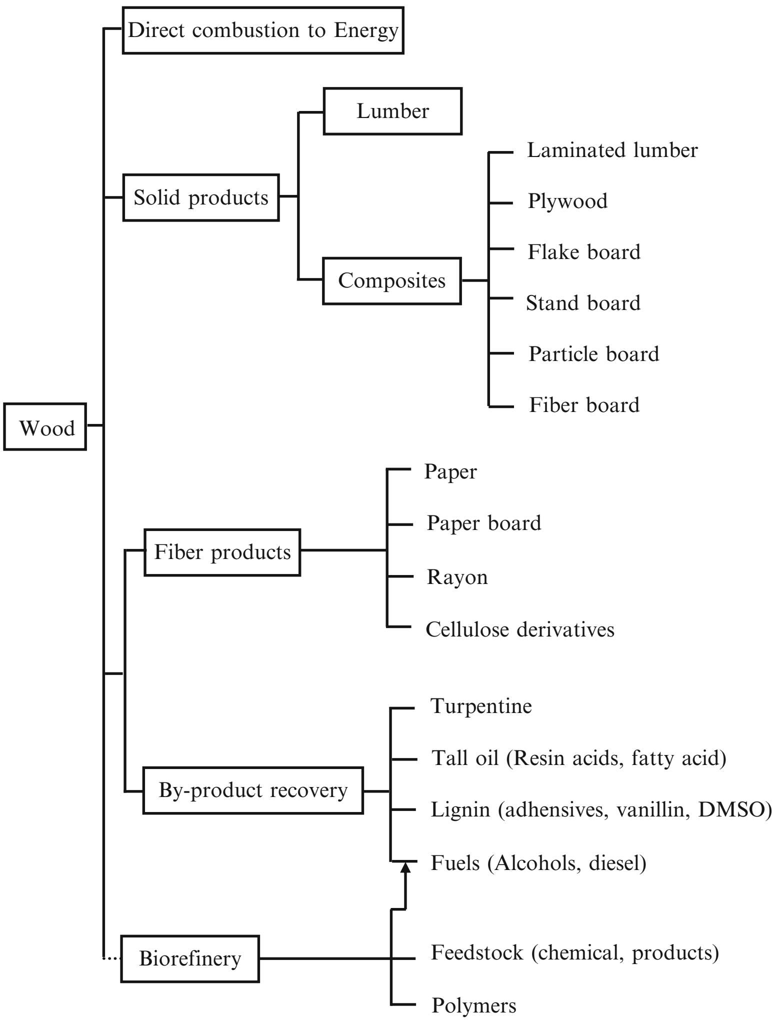


Fig. 28.2 Major uses of wood (DMSO stands for dimethyl sulfoxide)

Table 28.2 Chemical composition of wood [6]

Component	Softwood (%)	Hardwood (%)
Cellulose	42	42
Hemicelluloses		
Xylan	10	20–35
Glucomannan	20	3–4
Lignin	25–35	18–25
Extractives	2–5	2–4

They are often labeled as the matrix material of wood. Of the hemicelluloses, the xylan (polymer of xylose) is dominant in hardwoods, whereas glucomannan is the major type occurring in softwoods (Table 28.2). The DP of hemicelluloses being in the range of 100–200 is much less than that of cellulose.

Xylan. The content of xylan in hardwood (19–35%) is much higher than that of softwood (10–14%). Basically, xylan is a linear polymer of β (1→ 4)-D-xylose containing some 4-O-methyl α-D-glucuronic acid attached to the C_3-position of xylose unit. In hardwood, some of the C_2- or C_3-position of xylose also contains acetyl groups (Fig. 28.4). In case of softwood, the xylan does not contain any acetyl group and instead, some of the C_2-position of xylose units are linked to an α-L-arabinofuranose unit. The presence of these branched units does have an impact in the pulping and bleaching operations.

Glucomannan. Softwood has a significantly higher content of glucomannan (12–18%) than hardwood (3–4%). The structure of glucomannan in hardwood is quite simple being a linear molecule of β (1→ 4)-D-glucose and mannose in a molar ratio of 1:1.5 (Fig. 28.5). In case of softwood, the glucomannan contains some α-D-galactose units linked to the C_2-position of the glucose or mannose unit. Also, some of the glucose or mannose units contain an acetyl group.

Lignin

Lignin is a three-dimensional polymer derived from the enzyme-initiated dehydrogenative polymerization of three

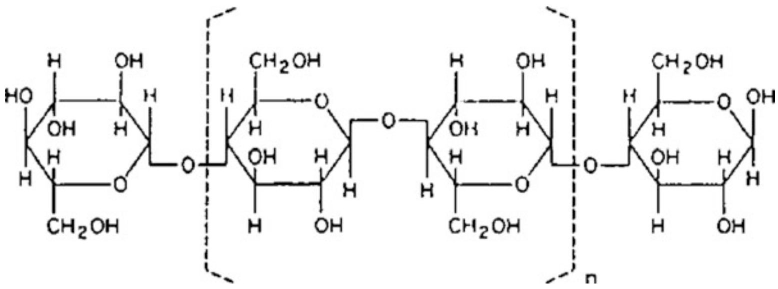


Fig. 28.3 Structure of cellulose

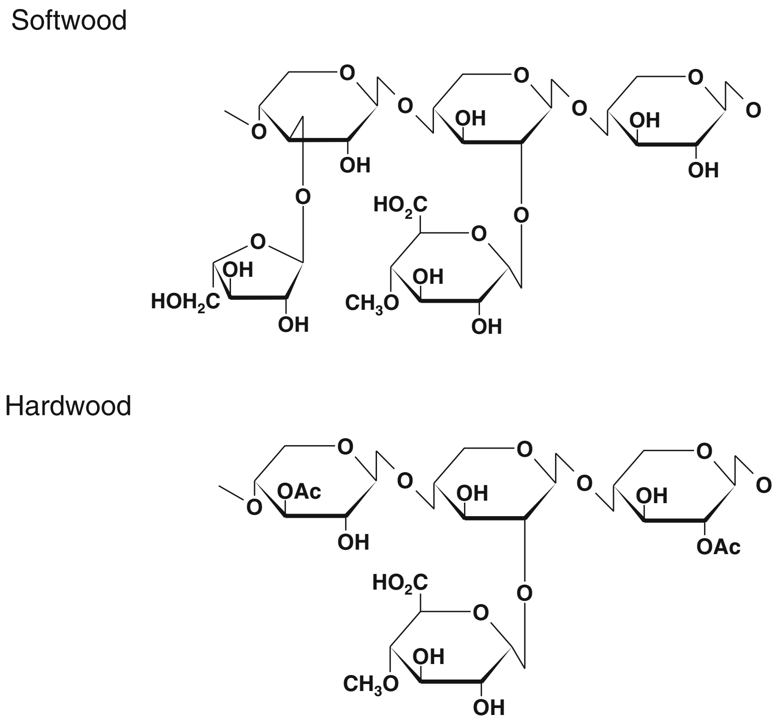


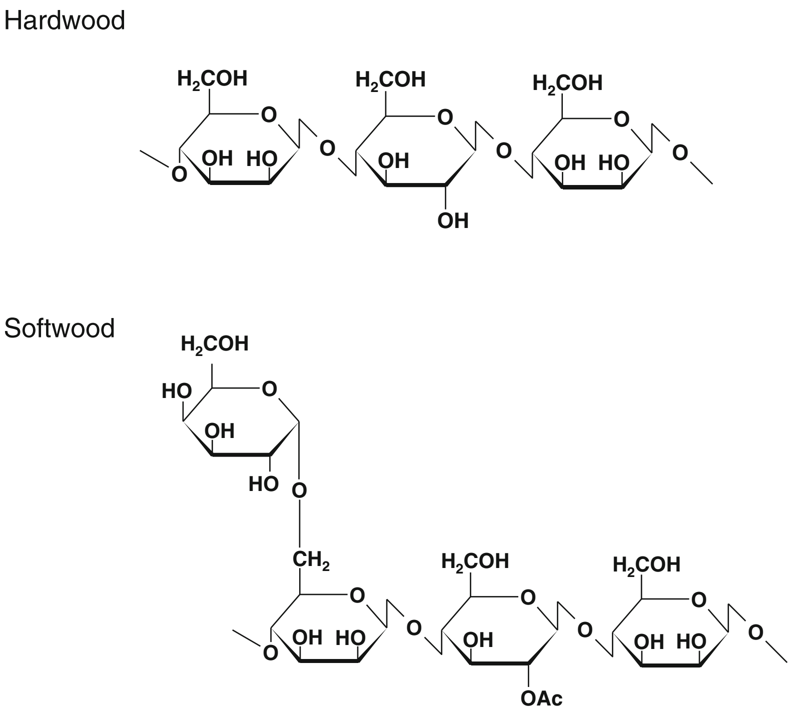
Fig. 28.4 Structure of xylan in wood

major precursors: coniferyl (H), sinapyl (S), and *p*-coumaryl (H) alcohol (Fig. 28.6). Softwood lignin is derived mainly from the G-type precursor (95%) with some H unit (5%) and is generally referred as a guaiacyl (G) type lignin. Whereas hardwood lignin is derived from a mixture of G and S units and its ratio differs considerably among different wood species. The presence of the S-type unit makes the lignin more reactive and renders the hardwood being easier to pulp or degrade. Thus, extensive efforts are continued to breed trees having a high content of the S-type lignin.

Although the approximate contents of major lignin linkages are reasonably well understood now, the chemical structure of lignin, unlike cellulose or hemicelluloses, still cannot be precisely defined. Figure 28.7 illustrates a segment of lignin structure to include all major types of linkages established for softwood. As indicated, lignin contains mainly the linkages of ether and carbon–carbon types. Since carbon–carbon linkages are very resistant to chemical attack, the degradation or fragmentation of lignin is largely limited to the cleavages of ether units at the α- and β-positions. This is the reason that the commercial pulping process generally requires drastic conditions for the lignin removal (delignification reaction) being practiced in the paper industry.

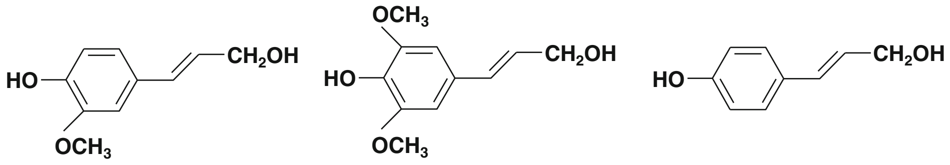
Extractives

Wood contains a variety of minor components that can be extracted by neutral solvents like organic solvents or water. Those soluble in organic solvents are generally classified into wood resins (terpenes, resin acids), fats, unsaponifiable (steroids), and phenolic components (lignan, stilbenes, and tropolones). Those water-soluble extractives include pectin, starch, and polyphenols such as tannins. A specific example for most of these extractives is shown in Fig. 28.8. Although extractives are present in small quantity, they play a significant role in determining the utilization for a number of wood species.


Hardwood

Softwood

Fig. 28.5 Structure of glucomannan in wood



1
Trans-Coniferyl alcohol
(Guaiacyl type: G)

2
Trans-Sinapyl alcohol
(Syringyl type: S)

3
Trans-*p*-Coumaryl alchol
(*p*-Hydroxyphenol type: H)

Fig. 28.6 Lignin precursors

Wood Structure

The tree trunk is consisted of mainly wood (xylem) with bark (phloem) around (Fig. 28.1). Cambium, an active tissue lies between the xylem and phloem, is responsible for tree growth by adding new tissues inward (sapwood) and outward (inner bark). Unlike the sapwood that is continued to accumulate annual growth, the outer bark would often rupture and be placed by the inner bark as the trunk gets bigger.

The physical nature of new tissues generated depends on the growing conditions like seasonal weather. The alternate dark and light bands showing on the cross-section of a softwood tree (Fig. 28.1) reflect the variation of seasonal growth. This pattern constitutes the so-called annual ring. The lighter band corresponds to the growth from the early growing season (springwood or early wood). The darker band termed summerwood or latewood is derived from the later season.

Wood Cells

Wood is the aggregate of countless elongated cells [1]. The bulk of vascular tissues in softwood (90–95%) are in the form of fiber tracheids (Fig. 28.9e). These fibers provide the function of liquid conduction and physical strength for the tree. The remaining wood tissue (5–10%) is consisted of *parenchyma* cells.

One of the functions for parenchyma cells is to convert sugars into storage starch to serve as food reserves for the

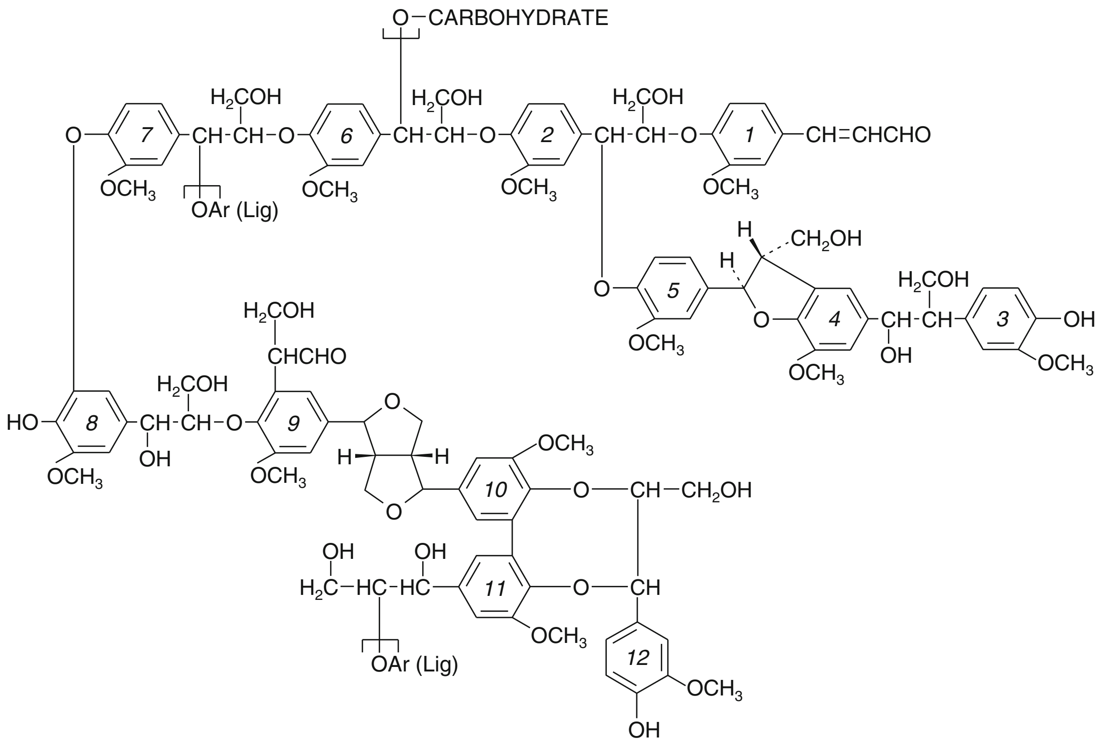


Fig. 28.7 Major types of linkages occurred in softwood lignin

tree and to reconvert starch to sugars when they are needed for growth. The starch and sugar content of sapwood is one reason why most wood-damaging organisms prefer dead sapwood over heartwood, and why sapwood is never quite as durable as heartwood, except in living trees whose sapwood cells fight for the attacking organisms.

Hardwood trees evolved after the softwood and have specialized water conduction cells called vessels (Fig. 28.9a–c). Each vessel consists of numerous drum shaped vessel members with open ends, stacked one above the other to form a continuous vessel pipeline from root tips to the leaves. The physical support for hardwood trees is provided by the libriform fiber, which along with the vessel element is much shorter than the fiber tracheids present in softwood wood (Table 28.3 and Fig. 28.9). The vessels in oaks and some other broad-leaved trees are large enough to be seen with the unaided eye.

Distinct differences between the earlywood (springwood) and latewood (summerwood) are further reflected from the scanning electron micrographs. For red pine (Fig. 28.10a) and white pine (Fig. 28.10b), the wood cell of earlywood is significantly larger in size with a thinner wall as compared to that of latewood. The transition was quite abrupt in case of red pine (Fig. 28.10a) while a gradual change was found in white pine (Fig. 28.10b).

Figure 28.11 is the photomicrography of hardwood cross-sections showing the distinct feature of vessels. For black ash as in oak, the vessels in the earlywood are much wider than those in latewood (Fig. 28.11a). These trees form the distinct growth rings and are generally referred as the ring-porous woods. In case of yellow birch, the vessels are essentially the same size throughout the growth ring (Fig. 28.11c) and these are called the diffuse-porous wood. Also, between these two extremes, a semi-ring-porous pattern was observed for wood like persimmon (Fig. 28.11b). Since coniferous woods lack vessels, they often are referred as the nonporous wood.

The proportion of different wood cells [7] varies considerably among hardwood species (Table 28.3). For example, the proportion of vessel could range from 12% for black ash to 55% for sweetgum. Whereas, the percent fiber content by volume could range from 30% in sweetgum to 70% in black ash wood. Thus, the quality and utilization of hardwood vary considerably with wood species.

Virtually all cavities of wood cells are interconnected, as mentioned earlier for the vessel members of hardwoods. Tiny openings in the cell walls of overlapping ends of tracheids provide vertical passageways for water rising from cell to cell in conifers. Thin walls of ray cells are perforated for the radial transport of sugars, and tangential connections provide for growing cells. Hence, wood is permeable in all three of its cardinal directions. The longitudinal orientation of most cells, however, causes liquids and gases to flow and diffuse many times faster in the longitudinal or fiber direction than in the two transverse directions.

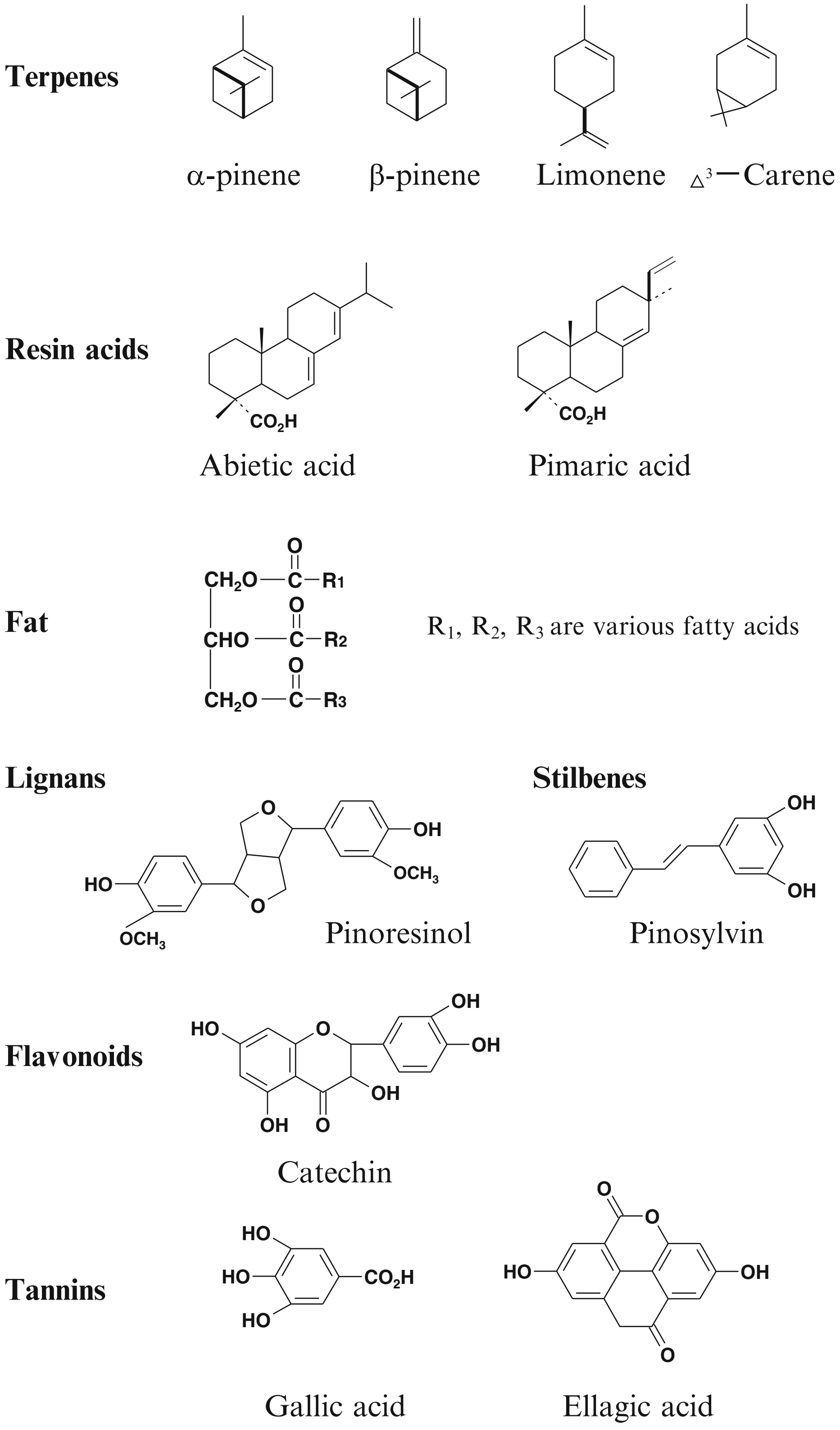


Fig. 28.8 Examples of some wood extractives

This explains why stacked firewood dries predominantly at the exposed ends. The radial orientation of rays allows liquids and gases to move radially up to twice as rapidly as tangentially [8].

As a porous material, wood contains air-filled or water-filled cell cavities surrounded by cell walls. The cell wall substance is remarkably uniform; it differs very little even between conifers and broad-leaves trees and has a constant

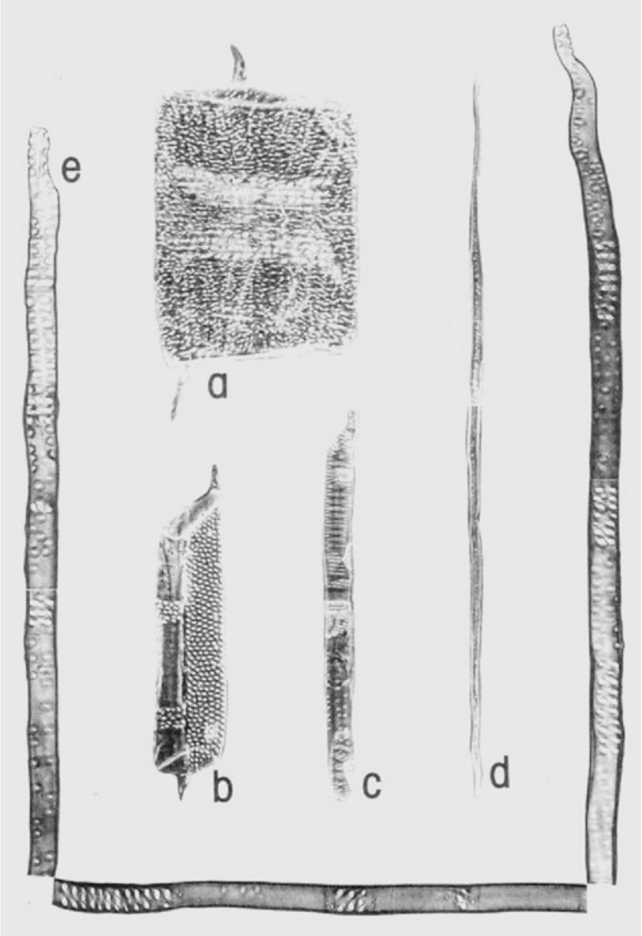


Fig. 28.9 Photomicrographs of major hardwood and softwood cell types illustrating the comparative size and shape of various elements [1]. Note especially how much longer the softwood tracheid (**e**) is than the vessel segments (**a**–**c**). Also shown is the libriform fiber (**d**) from hardwood. To give an indication of scale, the softwood tracheid is approximately 3.5 mm in length (from Core et al. [1], Copyright © 1976, with permission of the copyright owner)

Table 28.3 Proportion and length of wood cells[a]

Type of cells	*Softwood (%)*	*Hardwood (%)*
Tracheids		
Proportion (%)	90–95	–
Length (mm)	3–6	–
Libriform fibers		
Proportion (%)	–	35–69
Length (mm)	–	0.8–1.9
Vessels		
Proportion (%)	–	12–56
Length (mm)	–	0.2–1.1
Rays		
Proportion (%)	5–10	6–20

[a] Taken from the range for 12 softwood and 13 hardwood species [7]

density of around 1.5 g/cm^3. The density of wood thus depends on the proportion of the cell wall volume to the volume of the bulk wood.

Wood density is *the* indicator of nearly all wood properties and of wood quality. Relatively dense species are strong, as one might expect from their large cell wall mass. Wood species also vary in features not related to density, such as color, percentage and arrangement of cell types, and minor chemical constituents. Therefore, even woods of the same density are likely to differ, with the properties varying more than the densities.

The widespread use of wood is due to its many advantages, including its very high strength-to-mass ratio, good toughness, low cost and availability, workability, attractive appearance, and good thermal and electrical insulating properties. However, there also are a number of disadvantages of wood, which have limited its usage for many applications; these include its lack of plasticity and fusibility, shrinkage and swelling properties, deterioration due to fungi and insects, and variability. Fortunately, most of its negative properties can be corrected by proper treatment and modification, as described later in this chapter.

Cell Wall Composite

Wood cells or fibers are joined together by an intercellular layer or so-called middle lamella (ML) being very rich in lignin content (Fig. 28.12). This is the general contention that lignin functions as a natural adhesive to glue the cells (fibers) together. Each cell wall consists of an outer primary wall (P) and a secondary wall (S). The S wall is comprised of three sublayers: outer (S1), middle (S2), and inner (S3) layers. Inside the S3 layer, there is a rough or warty surface, generally termed as warty (W) layer.

One of the interesting features in this cell wall structure is the orientation of microfibrils (aggregates of cellulose molecules), which vary among different layers. As indicated, the microfibril orientation is random in the P wall, almost perpendicular to the long axis of the cell in both S1 and S3 layer, and approximately parallel to the cell axis in the S2 layer. Thus, each cell structure may be considered as a mini cellulose-reinforced composite with the hemicelluloses and lignin being functioned as fillers.

The cellulose polymers in nature, as mentioned earlier, are laid down uniformly with chains paralleling each other. The long-chain molecules associated strongly through extensive hydrogen bondings to form so-called microfibrils. The microfibrils contain regions with perfect cellulose chain alignments known as micelles or microcrystallites (Fig. 28.13).

However, the structure of the microfibrils is not completely uniform in terms of the alignment of the cellulose macromolecules. The regions of nonuniformity between the micelles in the microfibrils are called amorphous regions.

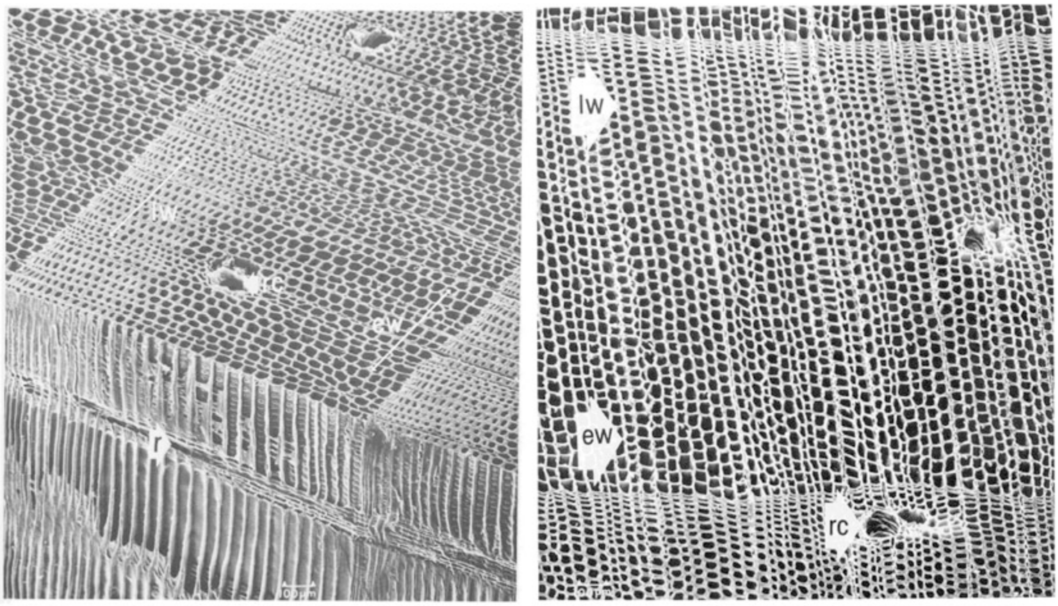


Fig. 28.10 (**a**) The scanning electron micrograph of red pine showing abrupt change in tracheid size from earlywood (ew) to latewood (lw) (from Core et al. [1], Copyright © 1976, with permission of the copyright owner). (**b**) The scanning electron micrograph of eastern white pine showing a gradual change in tracheid size from earlywood (ew) to latewood (lw) (from Core et al. [1], Copyright © 1976, with permission of the copyright owner)

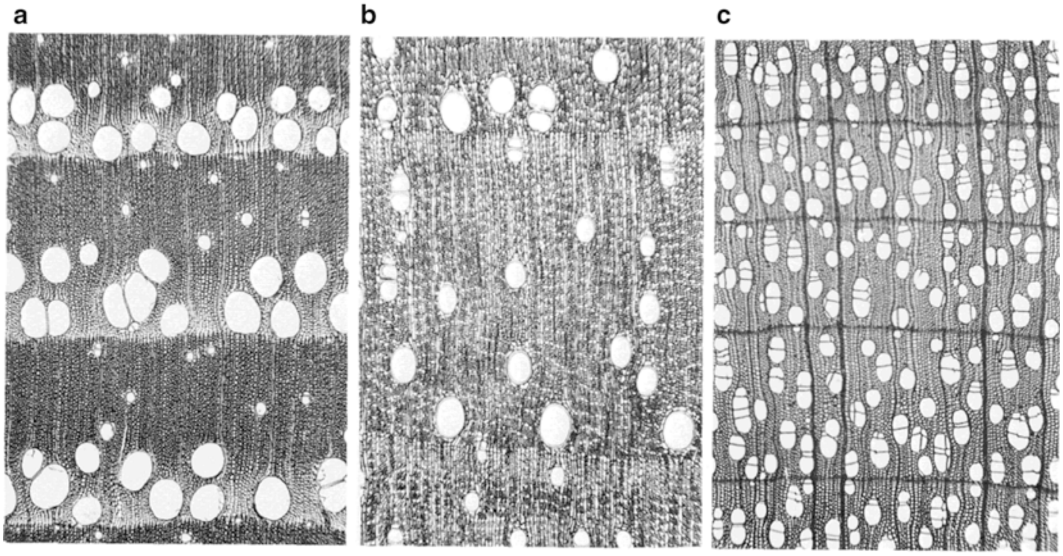


Fig. 28.11 Examples of hardwood groupings based on the pore size changes across the growth increment: (**a**) ring-porous (black ash); (**b**) semi-ring-porous (persimmon); (**c**) diffuse-porous (yellow birch) (from Core et al. [1], Copyright © 1976, with permission of the copyright owner)

Thus, the cellulose microfibril has a crystalline–amorphous character. Water molecules enter the amorphous regions and swell the microfibrils; ultimately this is the mechanism by which fibers and wood swell in moist or wet environments.

The final fiber cell wall structure is essentially layers of the microfibrils or macrofibrils aligned in several different directions, as shown in Fig. 28.12. The microfibrils that make up the wood fiber are visible under the scanning electron microscope, which has a greater magnification than the light microscope.

The entity holding the fibers together, the middle lamella, is almost entirely lignin (90%), as mentioned earlier. For the

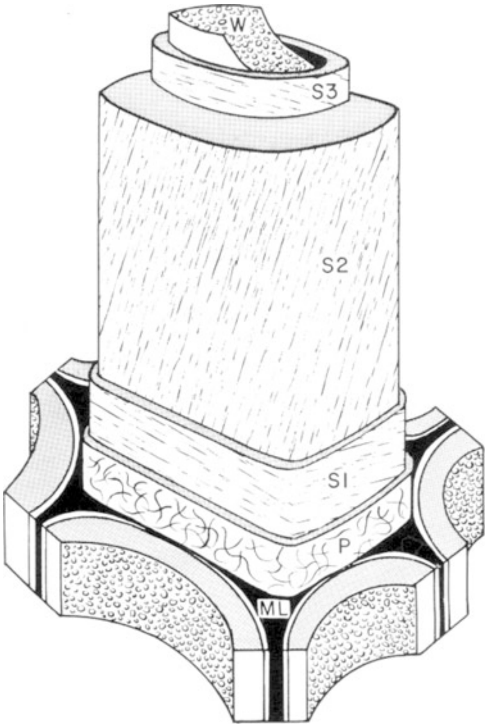


Fig. 28.12 A simplified cell wall structure showing the middle lamella (P), the primary wall (P); the outer (S1), middle (S2), and inner (S1) layers of the secondary wall; and the warty (W) layer (from Core et al. [1], Copyright © 1976, with permission of the copyright owner)

cellulose fibers to be separated, the middle lamella lignin must be chemically removed, a process that also removes most of the hemicelluloses, or must be mechanically degraded to free the fibers for papermaking. A paper sheet then can be formed from the separated cellulose fibers by depositing them from a water slurry onto a wire screen. The water drains away and the fibers collapse, leaving a fiber mat that derives its main strength from reassociation of the fibers through many hydrogen bondings—the same type of bondings that gives mechanical integrity to the fibers [9].

The long fibers from softwoods (Table 28.3) usually are preferred in papermaking for products that must resist tearing, such as grocery bags, whereas the shorter hardwood fibers give improved opacity, or covering power, and printability to the final paper sheet. The type of pulping process also affects the pulp properties, as described in a later section.

Pulp and Paper

The concept of making paper from the fibers derived from lignocellulosic materials—an integrated system of fiber separation (pulping) and re-forming of the fibers into a mat (papermaking)—is attributed to T'sai Lun, a court official in southeast China in AD 105. The first fibers were obtained from old hemp rags and ramie fishnets, but shortly thereafter, the inner bark fibers from paper mulberry trees also were utilized for papermaking. Bamboo was used as a source of fiber several centuries later. The rags were macerated into a pulp in water with a mortar and pestle; then, after dilution in a vat, the pulp was formed into a wet mat on a bamboo frame equipped with a cloth screen to drain the free water. The mat was dried in the sun. The invention was based on the need for a writing material to replace the expensive silk and inconvenient bamboo strips. The invention was a closely guarded secret for many centuries but eventually filtered west, reaching Europe in 1151 and the United States (Pennsylvania) in 1690 [9].

The importance of the pulp and paper industry to the American economy is exemplified by the growth rate in the use of paper and paper products. New uses continually are being found for paper, and these developments together with a rising standard of living have resulted in a constant increase in the per capita consumption of paper. The industry has shown constant growth since about the middle of the nineteenth century, when processes were invented for the production of fiber from wood, replacing reliance on cotton rags as a source of fiber (Fig. 28.14). The relative paper consumption by grade is shown in Fig. 28.15, and differences in the economies of developed and less developed countries are demonstrated in Fig. 28.16 by the wide disparity in per capita consumption of paper and paperboard.

In 2003, a total of 328 million metric tons of paper and paperboard was produced in hundreds of separate mills. This paper and board was made from 170 million metric tons of pulp, produced primarily from wood in over 200 pulp mills, supplemented by recycled paper [10].

The five leading states for the fine paper production are Wisconsin, Maine, Washington, Alabama, and Louisiana. Only a small amount of paper is made from rags or other fiber sources (agricultural residues); so most wood pulp is produced from those areas of the country that are heavily forested. With the exception of Maine and Wisconsin, pulp production is concentrated in the southern and northwestern sections of the United States. Georgia and Washington are the top pulp-producing states. Although a high percentage of pulp is converted into paper or board at the same plant site, a significant portion (classified as "market pulp") is sold in bales to mills in other sections of the country for subsequent manufacture into paper. Thus, New York and Michigan are relatively important in papermaking, but they rank quite low as pulp producers. However, both New York and Michigan are large users of waste paper that is reprocessed into usable fiber.

Raw materials for the pulp and paper industry can be classified as fibrous and nonfibrous. Wood accounts for over 95% of the fibrous raw material (other than waste paper) in the United States. Cotton and linen rags, cotton

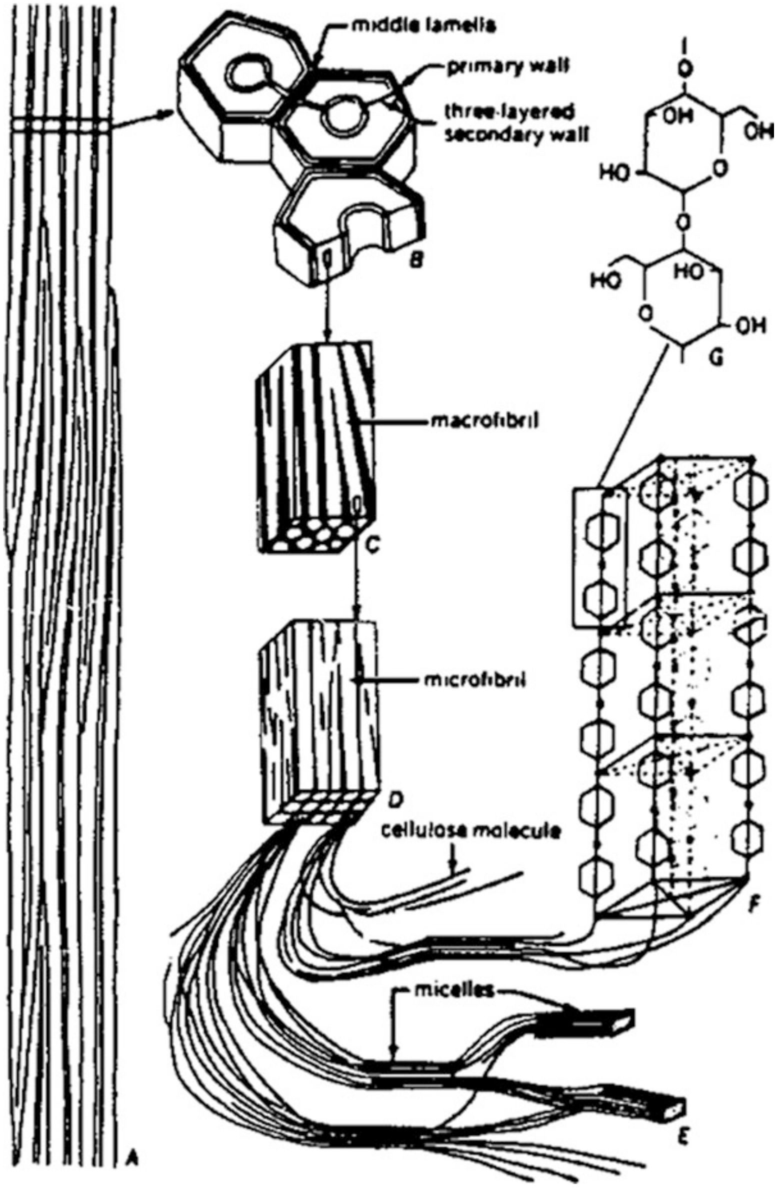


Fig. 28.13 Detailed structure of cell walls. *A*, strand of fiber cells. *B*, cross-sections of fiber cells showing gross layering: a layer of primary wall and three layers of secondary wall. *C*, fragment from the middle layer of secondary wall showing macrofibrils (white) of cellulose and interfibrillar spaces (black) which are filled with noncellulosic materials. *D*, fragment of macrofibril showing microfibrils (*white*), which may be seen in the electron micrographs. The space among microfibrils (*black*) is filled with noncellulosic materials. *E*, structure of microfibrils: Chainlike molecules of cellulose, which in some parts of microfibrils are orderly arranged. These parts are the micelles. *F*, fragment of a micelle showing parts of chainlike cellulose molecules arranged in a space lattice. *G*, two glucose residues connected by an oxygen atom—a fragment of a cellulose molecule (from Esau K (ed) (1977) Anatomy of seed plants, 2nd edn, Courtesy of John Wiley & Sons, Copyright © John Wiley & Sons, New York; with permission of the copyright owner)

linters, cereal straws, esparto, hemp, jute, flax, bagasse, and bamboo also are used and in some countries are the major source of papermaking fibers.

Wood is converted into pulp by mechanical, chemical, or semichemical processes. Sulfite and kraft (sulfate) are the common chemical processes, and neutral sulfite is the principal semichemical process (neutral sulfite semichemical [NSSC]). Coniferous wood species (softwoods) are the most desirable, but the deciduous, broad-leaved species (hardwoods) have gained rapidly in their usage and constitute about 25% of pulpwood.

Table 28.4 summarizes the conditions utilized with the various pulping processes, and Fig. 28.17 shows the relative production of pulp by the major processes [9].

These various mechanical and chemical pulps vary considerably in the papermaking property and have different applications. In terms of the tensile-tear property [11] (Fig. 28.18), the SBK (semibleached kraft) is the strongest follower by UBS (unbleached sulfite), chemithermomechanical pulp (CTMP), thermomechanical pulp (TMP), refiner mechanical pulping (RMP), and GWD mechanical pulps. In contrast, the trend is reversed with respect to the tensile and light-scattering coefficient [11] (Fig. 28.19). Thus, kraft pulp makes a stronger paper while the GWD is very suitable for newsprint application.

Nonfibrous raw materials include the chemicals used for the preparation of pulping liquors and bleaching solutions and the various additions to the fiber during the papermaking process. For pulping and bleaching, these raw materials include sulfur, lime, limestone, caustic soda, salt cake, soda ash, hydrogen peroxide, chlorine, sodium chlorate, and magnesium hydroxide. For papermaking, they include

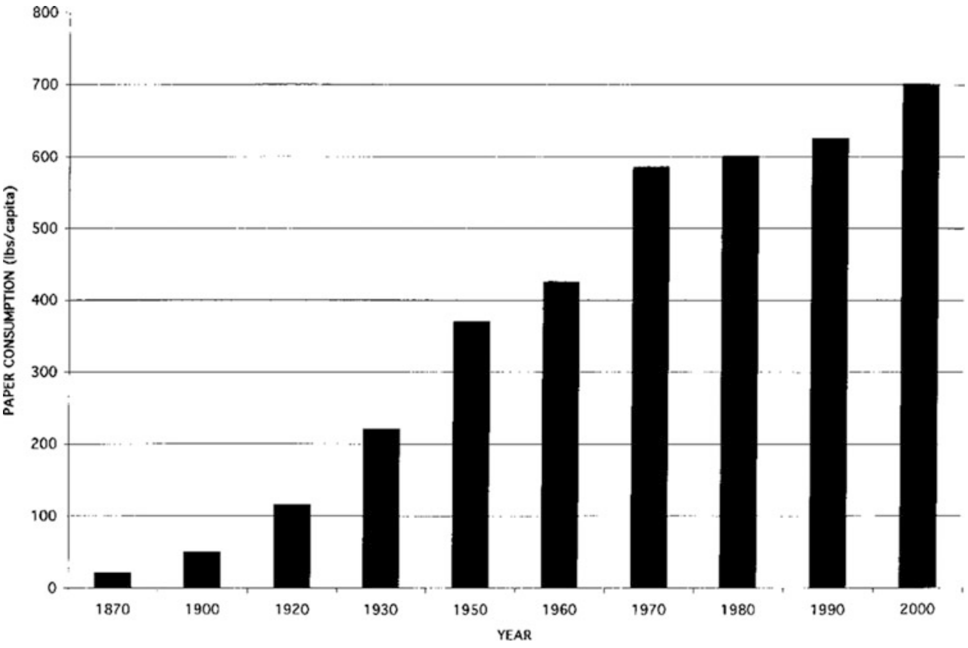


Fig. 28.14 Consumption of paper in the United States since 1860

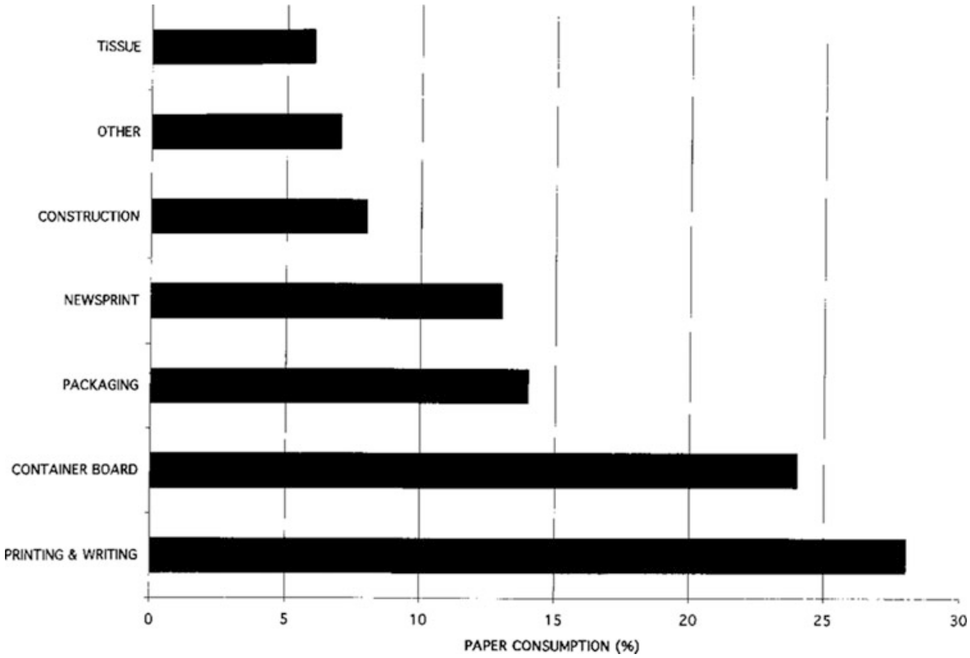


Fig. 28.15 Relative consumption of paper by grade

rosin, starch, alum, kaolin clay, titanium dioxide, dyestuffs, and numerous other specialty chemicals.

Wood Preparation

The bark of trees contains relatively little fiber and much strongly colored nonfibrous material; what fiber there is, is of poor quality. The nonfibrous material usually will appear as dark-colored dirt specks in the finished paper. Therefore, for all but low-grade pulps, bark should be removed as thoroughly as possible, and this removal must be very thorough in the case of groundwood and sulfite pulps if the finished paper is to appear clean.

Debarking usually is done in a drum barker, where the bark is removed by the rubbing action of logs against each other in a large rotating drum. Hydraulic barkers using high-pressure water jets are excellent for large logs and are common on the West Coast. Mechanical knife barkers are becoming more common and are used extensively in smaller

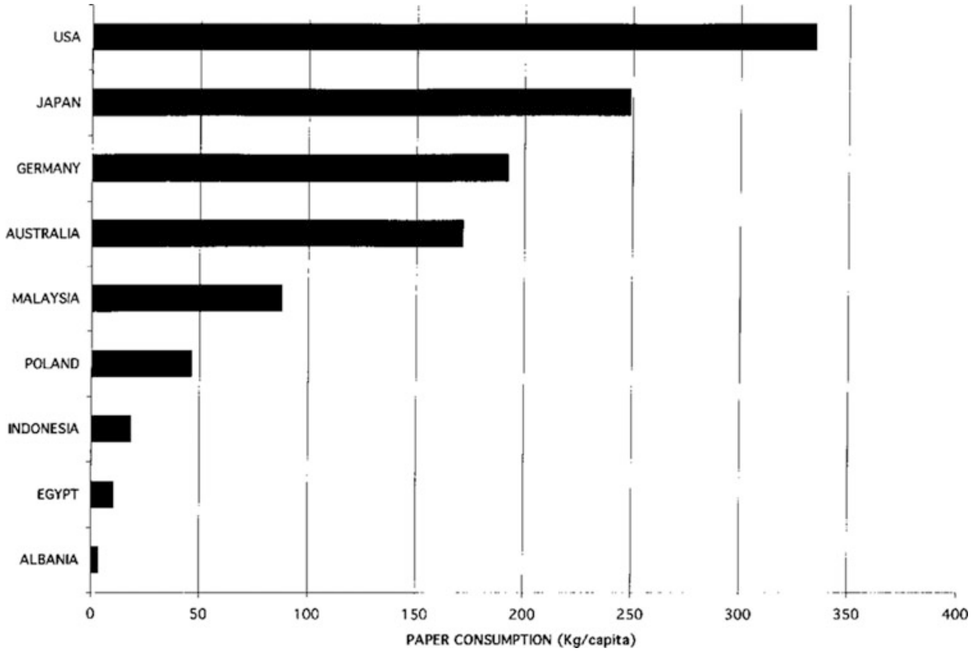


Fig. 28.16 Comparison of paper consumption in developed and developing countries

Table 28.4 Wood pulping by process and yield

Process	Acronym	Treatment Chemical	Mechanical	Pulp yield (%)
Mechanical processes				
Stone groundwood	SGW	None	Ginder	93–95
Pressure groundwood	PGW	None	Ginder	93–95
Refiner mechanical	RMP	None	Disk refiner (pressure)	93–95
Thermomechanical	TMP	Steam	Disk refiner (pressure)	80–90
Chemithermomechanical	CTMP	Na_2SO_3 or NaOH[a]	Disk refiner (pressure)	80–90
Chemimechanical[b]	CMP	Na_2SO_3 or NaOH	Disk refiner	80–90
Semichemical processes				
Neutral sulfite	NSSC	$Na_2SO_3 + Na_2CO_3$	Disk refiner	70–85
Green liquor	GLSC	$NaOH + Na_2CO_3$	Disk refiner	70–85
Nonsulfur	–	$Na_2CO_3 + NaOH$	Disk refiner	70–85
Chemical processes				
Kraft	–	$NaOH + Na_2S$	None	45–55
Acid bisulfite	–	$Ca(HSO_3)_2$ in H_2SO_3[c,d]	None	40–50
Magnetite	–	$Mg(HSO_3)_2$ in H_2SO_3[e]	None	45–55
Soda	–	NaOH	None	40–50
Soda-oxygen	–	$NaOH + O_2$	None	45–55
Soda-anthraquinone	SAQ	NaOH + AQ	None	45–55
Dissolving pulp				
Prehydrolysis kraft	–	Steaming and kraft (two-step process)	None	35
Acid sulfite	–	Acid sulfite (Ca, Na)	None	35

[a] Na_2SO_3 or NaOH, 2–7 % of wood

[b] Also chemical treatment after fiberizing

[c] Also sodium, magnesium, ammonia

[d] pH 2

[e] pH 5

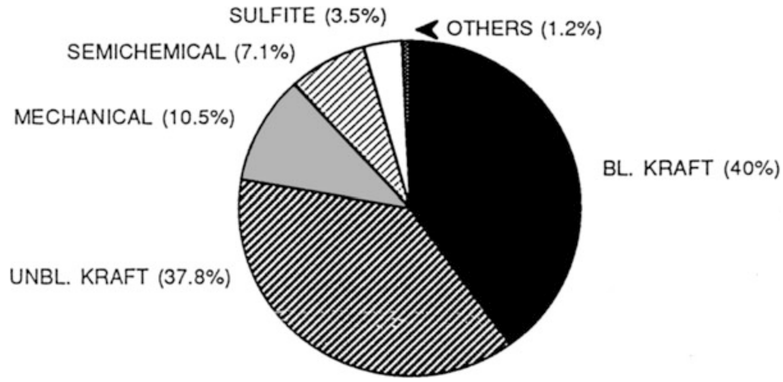


Fig. 28.17 Relative wood pulp production by grade

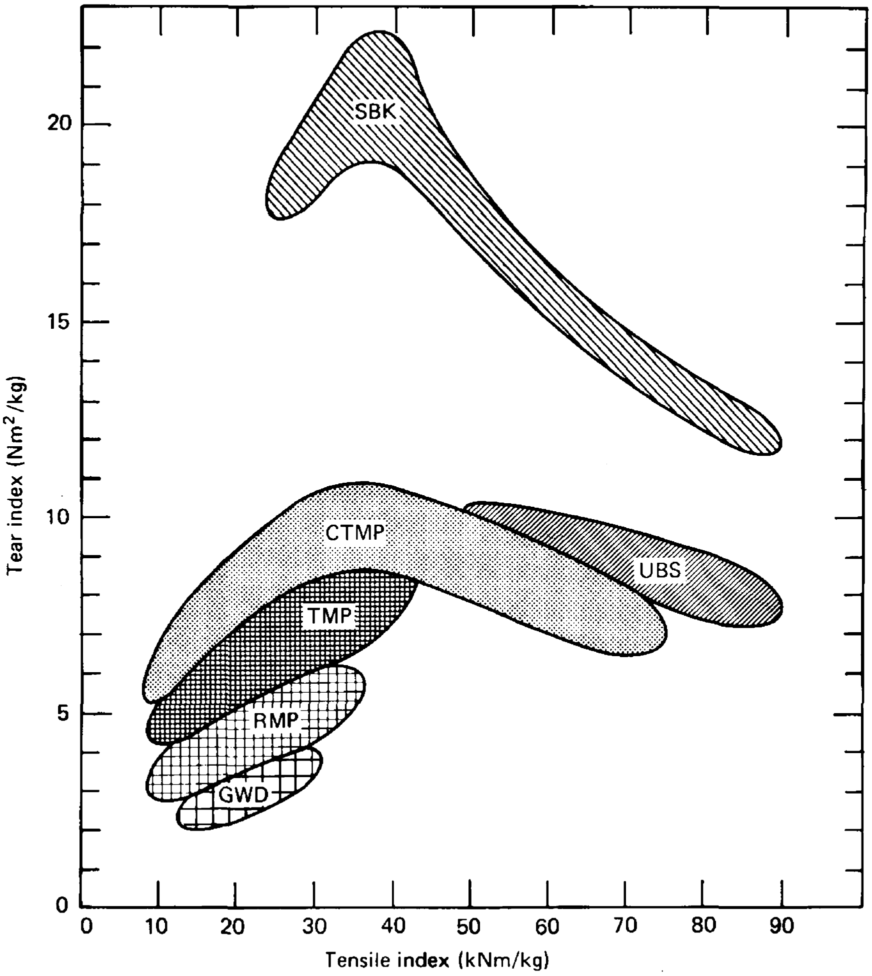


Fig. 28.18 Comparison of the tensile and tear property of stone groundwood (GWD), refiner mechanical pulp (RMP), thermomechanical pulp (TMP), chemithermomechanical pulp (CTMP), unbleached sulfite pulp (UBS), and semibleached pulp (SBK) (from McGovern [11], with permission of the copyright owner)

operations because of their relatively low capital cost. Also they have found widespread use in sawmills to debark logs prior to sawing so that the wood wastes can be used to produce pulp.

Wood cut in the spring of the year during the active growing season is very easy to peel. Much of the spruce and fir cut in the North is still hand-peeled during this season and usually represents the optimum in bark removal.

The standard log length used in the Northeast is 48 in., whereas 63 in. is common in the South. Wood generally is measured by log volume, a standard cord being considered to contain 128 ft^3. Large timber on the West Coast generally is measured in board feet of solid volume. Measurement and purchase of wood on a weight basis are practiced, and are desirable because weight is directly related to fiber content [2].

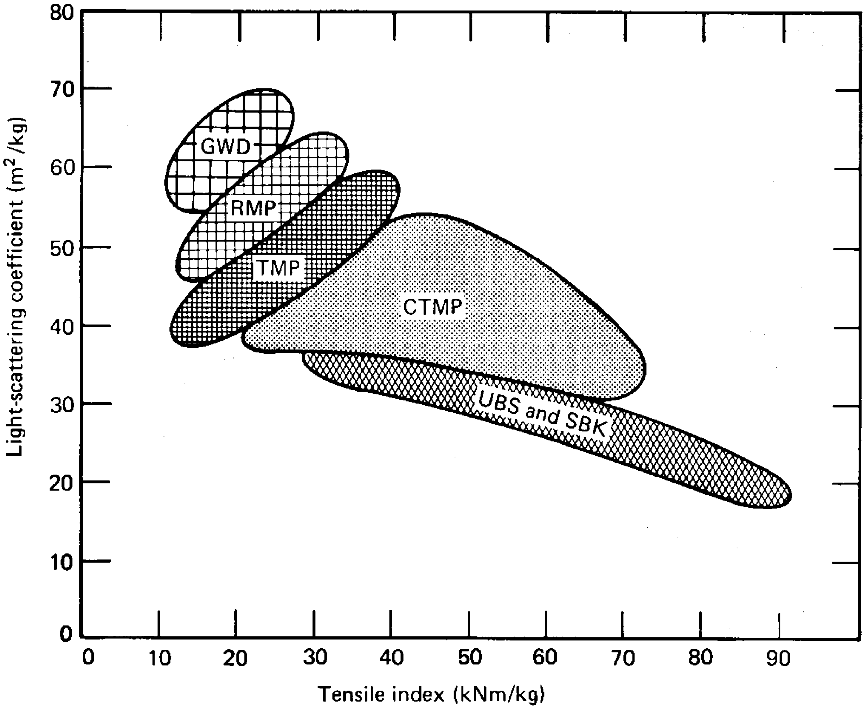


Fig. 28.19 Comparison of the tensile and light-scattering coefficient for stone groundwood (GWD), refiner mechanical pulp (RMP), TMP, CTMP, unbleached sulfite pulp (UBS), and semibleached pulp (SBK) (from McGovern [11], with permission of the copyright owner)

The growth in the use of wood residues has been substantial. By barking the sawlogs, the slabs, edgings, and other trimmings that formerly were burned now can be used to make pulp. Almost 28% of the wood used by the pulp industry is classified as waste wood. Several mills have been built that use no logs whatsoever but depend on residuals from satellite sawmill operations. Special sawmilling equipment has been developed to produce sawdust of a proper size so that it too can be used.

Wood used in producing groundwood pulp requires no further preparation after debarking, but that used in the other refiner mechanical or chemical processes must first be chipped into small pieces averaging 0.5–1 in. in length and about 1/8–1/4 in. in thickness.

Chipping is accomplished with a machine consisting of a rotating disc with knives mounted radially in slots in the face of the disc. Modern chippers have up to 12 knives; the ends of the logs are fed against the disc at about a 45 angle.

Mechanical Pulping

Mechanical pulping, as the name implies, involves mainly a grinding or abrading action to produce wood fibers. The stone groundwood process (SGW) is the conventional mechanical pulping, in which the log bolts are pressed against an abrasive grindstone. Developments in stone grinding include application of pressure to the grinder (PGW) and control of temperature. At one time, natural sandstone was used for the grindstone, but modern stones are either silicon carbide or aluminum oxide grits in a vitrified clay binder. Groundwood pulp is the primary component of newsprint. The characteristics of the stone can be varied to produce pulps "tailor-made" to fit their desired end use. For roofing or flooring felts, fibers must be extremely coarse and free-draining, so they require a stone with large grits, whereas newsprint pulps are very fine and require the use of small grits.

Pulp characteristics also can be varied by changing the stone surface pattern, the stone speed, the pressure of the logs against the stone, and the temperature of the groundpulp slurry. Generally a coarser and more freely draining pulp is obtained with a coarse surface pattern and high speed, pressure, and temperature. The type and the condition of the wood also are factors, but groundwood pulps usually are made from the coniferous or long fibered species because the deciduous or short-fibered species give very weak pulps.

Since 1975, SGW process has been gradually replaced by the RMP of chips. Basically, chips are pressed through the single- or double-rotating plates of a vertical-disc attrition mill. This RMP process has several advantages. In contrast to the SGW being limited to the use of log bolts, RMP fibers can be produced from chips or sawdust with improved strength properties (Fig. 28.18).

Groundwood or RMP pulp is low in strength compared with the chemical pulps. It is composed of a mixture of individual fibers, broken fibers, fines, and bundles of fibers. Papers made from these pulps also lose strength and turn yellow with time. Thus, these pulps are used only in relatively impermanent papers such as for newsprint, catalogs, magazines, and paperboard. Groundwood papers have excellent printing qualities because of high bulk, smoothness, resiliency, and good ink absorption. Newsprint contains about 80% groundwood, and the other publication paper grades contain about 30–70%, with the remainder a chemical pulp to provide greater strength. Groundwood is the cheapest pulp made and also utilizes the entire wood, giving close to a 100% yield. Now, the RMP is dominated in the production of mechanical pulps and their fiber characteristics can also be readily improved by modification of refining conditions leading to the development of many new products.

Major improvements in mechanical pulping technology include the pretreating chips with chemicals, steam, or both. When chips are steamed at about 130°C and 40 psi for 2–4 min followed by pressure refining, it produces a TMP, which has a significantly higher strength than the RMP or SGW fibers (Fig. 28.18). Similarly, chips can be pretreated with chemicals followed by disc refining. Various combinations of chemical and mechanical treatments have been used to produce pulps with specific properties. The chemimechanical pulping (CMP) refers to a mild chemical treatment while the semichemical pulping includes a mild cooking stage for a partial removal of the lignin.

Chemimechanic pulping (*CMP*). CMP pulps usually are produced by mild treatment of chips with caustic soda, sulfite, or peroxide followed by refining under atmospheric pressure. Chemical consumption is very low, and yields are usually 85–95%. When chemical pretreatment and pressure steaming are combined, it is called CTMP. The main purpose of chemical treatments is not to remove but to modify the lignin component. On this aspect, a pretreatment with sulfite is much more effective than alkalis, as it introduces the sulfonate group into the lignin (sulfonation) and softens the cell wall matrix. Thus, the sulfite treatment facilitates the mechanical fiber separation and thus improves the sheet strength property. The strength of CTMP is very comparable to that of the conventional acid-sulfite chemical pulps (Fig. 28.18). In fact, newsprint is produced commercially from the TMP or CTMP fibers alone without the use of any chemical pulps.

Semichemical Pulping

This is a two-stage process involving a mild cooking of chips for a partial removal of the lignin components followed by refining to separate the fibers. This semichemical process differs from the CMP in having a moderate degradation of the lignin component. The chemical stage can apply any of the chemical pulping process with reduced chemical charge, temperature, or cooking time. For hardwood, it is quite common to use a neutral sulfite liquor (Na_2SO_3 + Na_2CO_3) or green liquor (Na_2CO_3 + Na_2SH). The major uses of semichemical pulps are in making corrugating medium and linerboard.

The NSSC is one in which wood chips, usually from hardwoods, are cooked with Na_2SO_3 liquor buffered with either $NaHCO_3$, Na_2CO_3, or NaOH to maintain a slightly alkaline pH during the cook. Unbleached pulp from hardwoods cooked to a yield of about 75% is widely used for the corrugating medium. Although bleachable pulps can be produced by this process, they require large quantities of bleaching chemicals, and the waste liquors are difficult to recover. Currently many NSSC mills are located adjacent to kraft mills, and the liquors can be treated in the same furnace. Thus, the waste liquor from the NSSC mill becomes the makeup chemical for the kraft mill, solving the waste problem. NSSC hardwood pulp is the premier pulp for the corrugating medium and cannot be matched by any other process.

Chemical Pulping

The main purpose of chemical pulping is to remove the lignin portion of wood so that wood fibers can be liberated freely with very little mechanical treatment. The degree of lignin removal (or delignification) depends on the end uses of the final products. For making shopping bags, it is not necessary to remove all the lignin. However, a complete delignification, essential for making a permanent white paper, has to be achieved by separate bleaching operations.

Many variations are used throughout the world; the simplest classification of pulping methods depends on whether they are acidic or alkaline. Each has its own specific advantages and disadvantages, but the kraft (alkaline) process accounts for over 90% of all chemical pulp produced in the United States. The important acidic process is the sulfite, which however can be applied to the whole pH range now. All present commercial processes use aqueous systems under pressure at elevated temperatures, although several organosolv systems have reached the demonstration-plant stage.

Kraft Process

The kraft (sulfate) process uses a mixture of sodium hydroxide (NaOH) and sodium sulfide (Na_2S) as the active chemical. Although the designation of sulfate process could be misleading, it sometimes is used because sodium sulfate is the makeup chemical in the process. The word “kraft” now is mainly used to describe this process, derived from the Swedish or German word for strength, as this process produces the

Table 28.5 Typical kraft pulping conditions

Alkali charge[a]	15–20%
Temperature	170–175°C
Pressure	100–110 psig
Duration	2–3 h
Liquor to wood ratio	4:1 by weight

[a] Based on wood and calculated as Na_2O; consisting of approximately $5NaOH + 2Na_2S$

strongest pulp. Historically, sodium hydroxide alone (soda process) first was used as the alkaline pulping agent, but all these mills have converted to the kraft process because the soda pulp is weak and inferior to kraft pulp. The addition of Na_2S greatly enhances the lignin degradation and removal, which is the key mechanism in alkaline pulping.

The pulping (cooking) process traditionally was performed on a batch basis in a large pressure vessel called a digester. Conditions vary depending upon the type of wood being pulped and the quality of end product desired. Typical conditions for kraft cooking are listed in Table 28.5.

Digesters are cylindrical in shape with a dome at the top and a cone at the bottom. Ranging in size up to 40 ft high and 20 ft in diameter; the largest will hold about 7,000 ft^3 of wood chips (about 35 t) for each charge. The chips are admitted through a large valve at the top, and at the end of the cook they are blown from the bottom through a valve to a large blow tank. During the cook, the liquor is heated by circulation through a steam heat exchanger, which also avoids the dilution of the cooling liquor that would occur from heating by direct injection of steam.

The development of continuous digester significantly facilitated the production of kraft pulps. Figures 28.20 and 28.21 illustrate a two vessel vapor/liquor phase digester system [12] for a modified kraft cooking to achieve a maximum removal of the lignin while maintaining the pulp quality. Key features include the impregnation of wood chips with white liquor, injection of white liquor at three locations to maintain the alkali concentration nearly uniform throughout the cook, and a countercurrent mode to facilitate the lignin dissolution.

Kraft delignification. A key delignification reaction in kraft pulping is attributed to the cleavages of major lignin units (β-aryl ether) and to generate the phenolic hydroxyl group (PhOH). This fragmentation process requires a combination of both sodium hydroxide and sodium sulfide at elevated temperatures (170°C). Once the lignin fragments contain enough of these hydrophilic PhOH groups, they become soluble in alkalis. Figure 28.22 illustrates the overall process of lignin removal in a kraft process [13]. The extent of cooking is expressed as H-factor, which represents a combination of the cooking duration and temperatures.

The lignin-removal process is consisted of three distinct phases. The initial phase occurring during the heating period (up to 170°C) readily removes about 20% of wood lignin. Although the lignin dissolved was often referred to as "extracted lignin," it does involve the degradation of more reactive lignin structure. The bulk phase removes additional 70% of lignin, whereas the remaining residual lignin (5–10%) was very resistant to degradation. In commercial operation, the cooking has to be terminated at the inter-phase of the bulk and residual delignification. The residual lignin of kraft pulps has to be removed by the subsequent bleaching process.

Although kraft process is effective in the sense that it can pulp virtually any kind of wood, it is not very efficient, as about one third of the polysaccharides mainly hemicelluloses are degraded and removed during the process. As illustrated in Fig. 28.23, the delignification selectivity (a ratio of lignin removal to carbohydrate degradation), also displays three distinct stages [14, 15]. Both the initial and residual phases are very nonselective. Again, the curve suggests that the kraft pulping has to be terminated at the end of the bulk phase to avoid an excessive degradation of the polysaccharides to preserve the fiber quality.

The brown stock after cooking is washed in a countercurrent rotary vacuum washer system using three or four stages to remove the residual chemicals. The resulting pulp is ready for bleaching or for making papers such as grocery bags where the brown color is not objectionable. The separated liquor being very dark in color is known as "black liquor." The black liquor generally has a solid content of 15–20% containing mainly the dissolved lignin and hydroxyl acids from carbohydrate degradation with some extractives. After partial evaporation, the "tall oil" skimming containing mainly the wood resins is recovered as by-products (see section on "Chemicals from Wood Extractives").

Chemical recovery. The black liquor is concentrated further by multiple-effect evaporators to 60–65% solids. At this concentration, the quantity of dissolved organic components from the wood (lignin and carbohydrate degradation products) is sufficient to allow the liquor to be burned in the recovery furnace (Fig. 28.24).

By controlling the amount of excess air admitted to the furnace and the temperatures, the organics in the liquor can be burned. The chemical reactions occurring in the furnace consist of three different zones (drying, oxidizing, and reduction). The sodium carbonate and sodium sulfate formed on the top oxidizing zone will be dropped to the bottom of furnace. At the bottom reducing zone, the sodium sulfate formed or added to the liquor as make up will be reduced by carbon to sodium sulfide. After dissolving the molten smelt of Na_2CO_3 and Na_2S in water, the mixture (called "green liquor") is reacted further with slaked lime:

$$Na_2OH + Ca(OH)_2 \rightarrow 2NaOH + CaCO_3$$

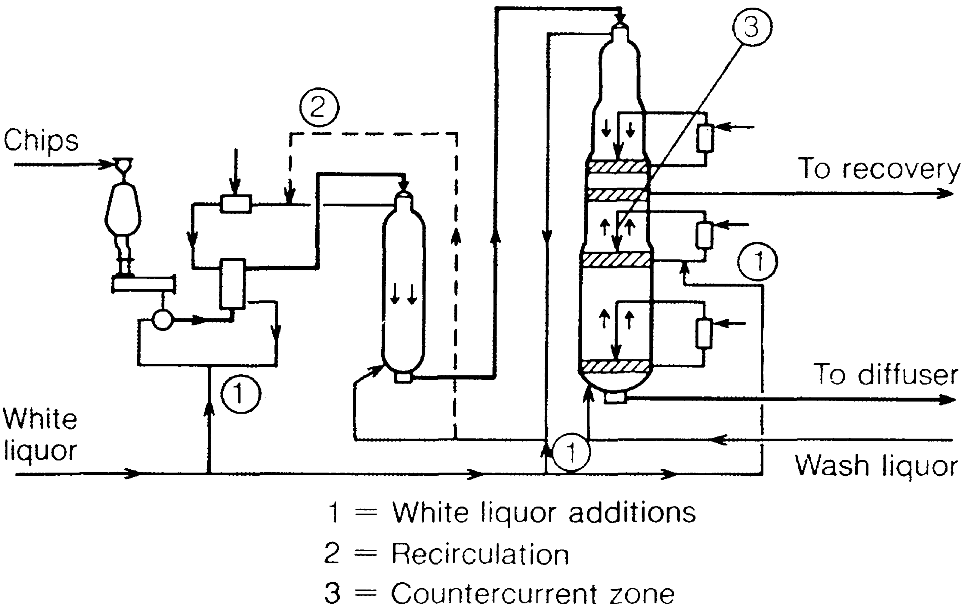


Fig. 28.20 Two-vessel vapor-liquid phase digester adapted for extended delignification (from Johansson et al. [12]; Smook GA (1992) Handbook for pulp and paper technologies, 2nd edn. Angus Wilde Publications, with permission)

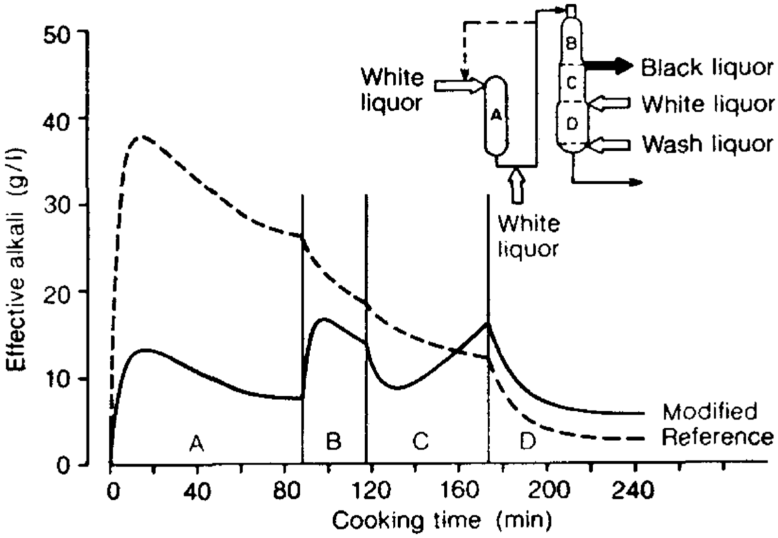


Fig. 28.21 Concentration profile of effective alkali inside chips in a two-vessel vapor/liquid phase digester system—conventional system compared with modified system as calculated by the mathematic model (from Johansson et al. [12]; Smook GA (1992) Handbook for pulp and paper technologies, 2nd edn. Angus Wilde Publications, with permission)

The Na_2S does not react with the lime, so the resultant mixture of NaOH and Na_2S (called "white liquor") can be reused to pulp more wood. The $CaCO_3$ sludge is filtered off, burned in a lime kiln, and reused. Thus, the chemical system is a closed one, as shown in Fig. 28.25, and this minimizes costs and pollution.

Kraft mill odor. A major drawback in kraft pulping is the odor produced associated with the sodium sulfide (Na_2S) used. This compound promotes the lignin degradation and dissolution and induces the formation of hydrogen sulfide (H_2S), methyl mercaptan (CH_3SH), dimethyl sulfide ($CH_3S–CH_3$), dimethyl disulfide ($CH_3S–SCH_3$). These four compounds constitute the so-called kraft mill odor. However, the paper company has done a great job in controlling the emission of these compounds through a combination of various techniques including a black liquor oxidation, improved evaporators and furnaces. Now, one can hardly recognize the existence of a kraft mill just based on the odor alone.

Future prospects. Virtually any kind of wood can be pulped by the kraft process; because it produces the strongest pulps with good economies, it has grown to be the dominant process in the world. Although the odor problem is a major concern, it does not appear that this process will be supplanted in the near future; instead, it will be improved and modified further.

When pulping resin-rich woods such as southern pine, the kraft process yields turpentine and tall oil as valuable by-products. The steam generated in the recovery furnace

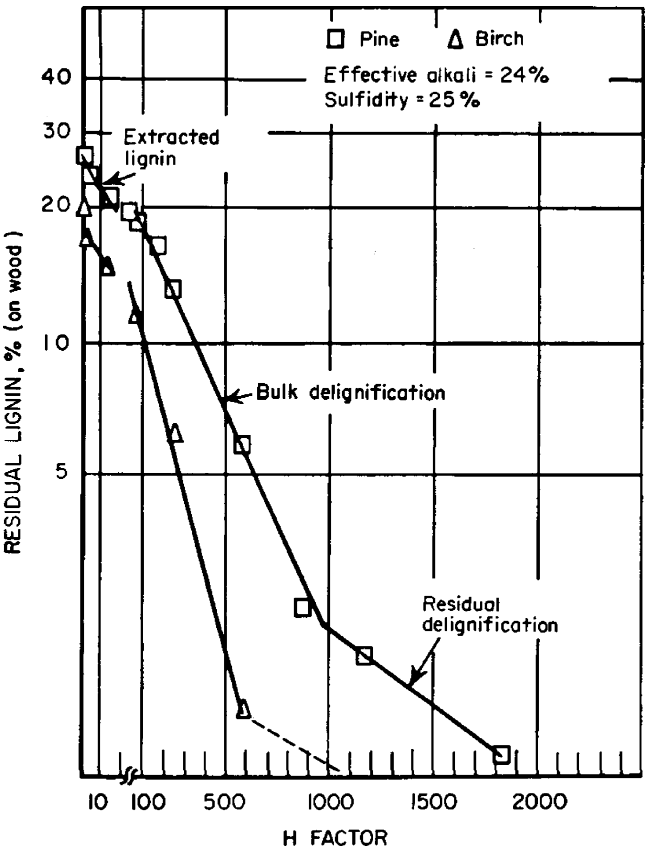


Fig. 28.22 Kraft delignification process as a function of H-factors (from Kleppe and Tappi [13]; Smook GA (1992) Handbook for pulp and paper technologies, 2nd edn. Angus Wilde Publications, with permission)

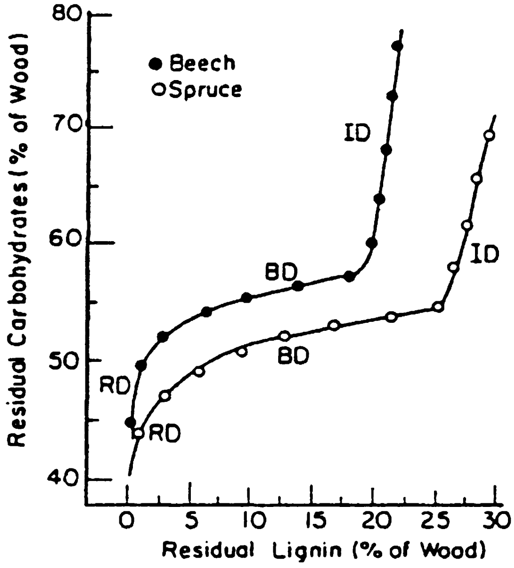


Fig. 28.23 A typical delignification selectivity in kraft pulping of softwood and hardwood [14, 15]

is almost enough to make the pulp mill energy self-sufficient. The discovery that small amounts of anthraquinone (0.05–0.25% on wood) added to alkali liquors can enhance delignification and produce the kraft quality pulp has been of great interest. Although the price of anthraquinone (AQ) is high and the chemical is not recovered, only small quantities are required, and the benefits outweigh the costs sufficiently that a number of mills around the world now routinely add AQ to the kraft pulpwood digester.

Sulfite Process

The conventional acid sulfite pulping is often called sulfite process and uses a cooking liquor of sulfurous acid (H_2SO_3) and calcium salt of the acid at pH around 1.5. This is the only acid process used in commercial pulping operation. For many years, this was the preferred process because it produced pulps of light color that could easily be bleached, it used cheap chemicals in fairly limited amounts so that no recovery was necessary, and it was a relatively simple process to operate.

Although the production of sulfite pulp has remained relatively constant for the last 30 years, the rapid growth of kraft pulping has reduced sulfite's share to less than 10% of the chemical pulp produced. There are several reasons for this, but the primary ones are the inability to cook resinous woods such as pine, problems in producing strong pulps from hardwoods, and of the greatest importance, the lack of a cheap and simple recovery system to reduce water pollution problems. However, this process produces dissolving pulps (pure cellulose) with special qualities such as high alpha-cellulose grades for rayon, so it will continue to be used.

Acidic delignification. The mechanism of sulfite delignification is distinctly different from the kraft pulping. The key reaction in acid sulfite cooking is the sulfonation (Fig. 28.26) and the lignosulfonates are soluble in water. In contrast, kraft delignification relies on an extensive degradation of lignin unit to lower molecular weight fragments. Although these lignin fragments are soluble in alkalis, they are not soluble in an acidic medium. Therefore, the lignosulfonate and kraft lignin have different molecular property and applications as a polymer.

Use of soluble bases. Initially, calcium was the preferred base because it was cheap and convenient to use. However, no recovery system is available for this base, so most calcium-base mills either have ceased operation or have converted to the more soluble bases. With magnesium, the cooking pH can be extended to the bisulfite range while the whole range of pH is applicable by using sodium or ammonium as summarized in Fig. 28.27. As the pH

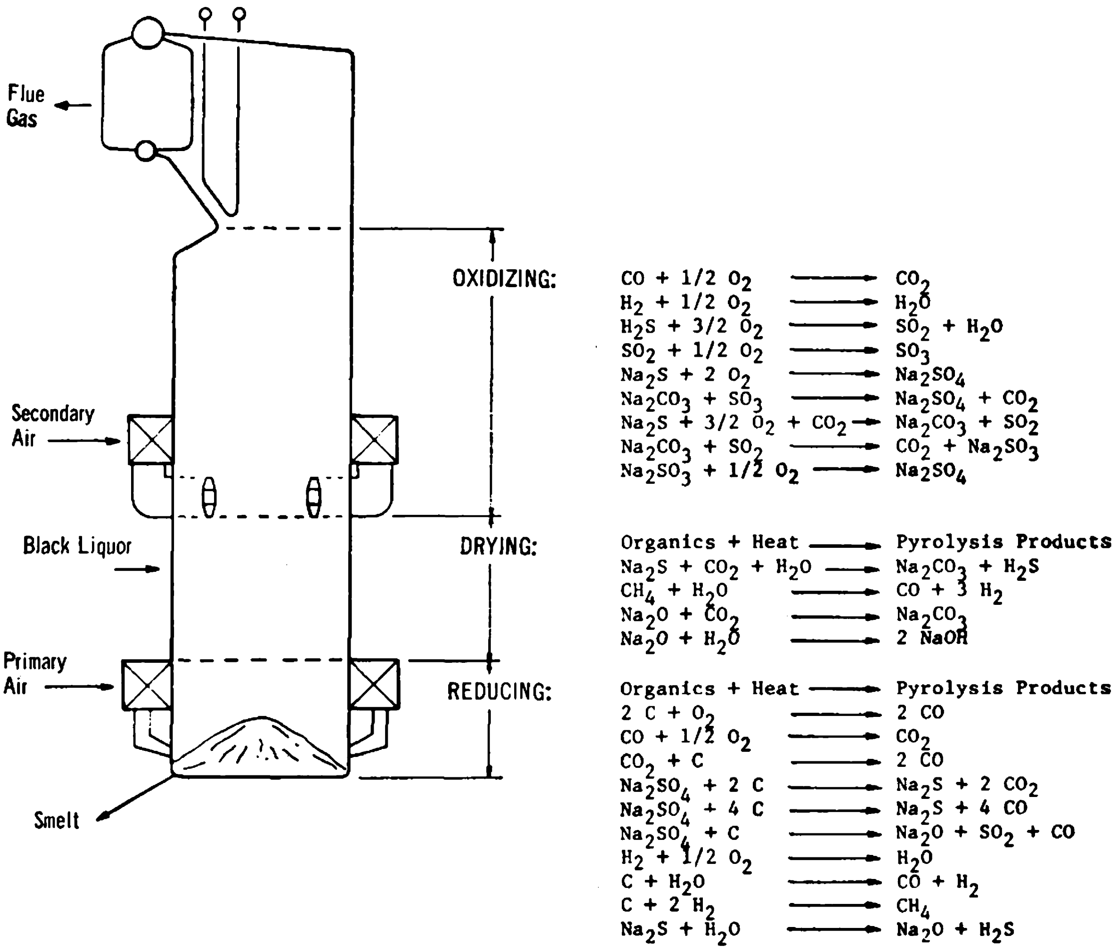


Fig. 28.24 Chemical recovery furnace (from Smook GA (1992) Handbook for pulp and paper technologies, 2nd edn. Angus Wilde Publications, with permission)

increases, the cooking can be conducted at higher temperatures and shorter.

In sulfite process, it is essential to have a well impregnation of chips prior reaching the maximum temperature. Thus, an extremely long cooking duration (7–10 h) is necessary with acid sulfite, whereas 4–5 h is sufficient with bisulfite at a higher temperature. Also, the higher pH cooking produces stronger pulps.

Different bases used were shown having noticeable effects on the pulping operations. Under identical cooking conditions (temperature, pH, and sulfur dioxide concentration), the ammonium base has the highest cooking rate and followed by magnesium and sodium. However, the ammonium bisulfite pulp has a lower brightness and is more difficult to bleach.

Chemical recovery. Sodium base is the easiest one to prepare (Na_2CO_3 or NaOH usually is used as the makeup chemical) and gives the highest quality pulp. However, the recovery processes, though available, is complicated and expensive. The magnesium-based bisulfite (Magnefite) process has been used in several mills and has a cutting edge in the chemical recovery. After cooking, the magnesium sulfite spent liquor (Mg-SSL) is processed and burned just like the kraft black liquor in a recovery furnace with the exception of no melts formation. All the inorganic components would be converted into SO_2 and MgO and they are be recovered by adsorption with scrubbers and by dust collectors, respectively.

In ammonium-based (from NH_4OH) process, the ammonia cannot be recovered, but the liquor can be evaporated and burned without leaving any solid residue, thus reducing water pollution. As long as aqueous ammonia remains low in price, this process will be attractive because the SO_2 can be recovered from the waste gases by passing them through a wet scrubber flooded with fresh ammonium hydroxide.

Modified sulfite process. Although the conventional acid sulfite has limited uses in the production of chemical pulps, the sulfite treatments, which can be conducted in the whole range of pH, has been extensively used especially in

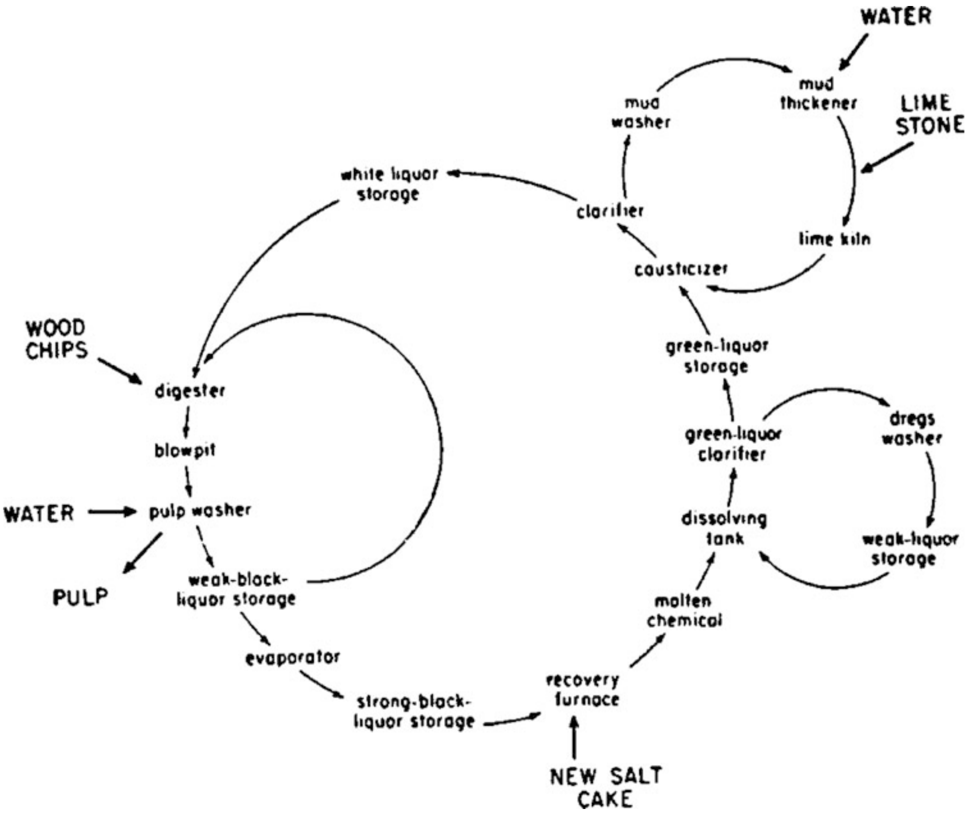


Fig. 28.25 Diagram showing cyclic nature of kraft recovery process (from Casey JP (ed) Pulp and paper, 3rd edn. Wiley-Interscience, New York. Copyright © 1980 John Wiley & Sons, New York; with permission of Copyright owner)

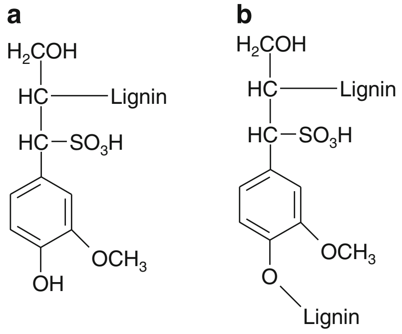


Fig. 28.26 Lignin sulfonation

conjunction with the production of chemimechanical or semichemical pulps. Interestingly, the NS (neutral sulfite)-AQ process can be used to produce chemical pulps from pine and birch wood with yields significantly higher than that from birch wood [16]. However, the NS-AQ delignification is a very slow process and produces a weaker pulp than kraft. More research is needed to improve the NS-AQ process.

Organosolv Pulping

The concept of using solvents to dissolve lignin is not new. In the 1930s, Kleinert and Tayenthal already developed and patented the first solvent (organosolv) pulping process based on aqueous ethanol (25–80% alcohol) [17]. Until recently there was little incentive to pursue alternatives to the kraft process because energy and chemicals were reasonably cheap, there were few environmental regulations, and there was little interest in by-products. However, new, strict environmental regulations, rising energy and chemical costs, and the very large capital investments needed to construct new kraft mills have stimulated research and investment in alternate solvent pulping processes [18–20].

The organosolv pulping includes the use of alcohols (methanol or ethanol), organic acids (acetic acid or formic acid), and others. In general, wood chips are contacted with an aqueous solvent mixture and cooked (digested) for a period of time at elevated temperatures. Digestion pressures range from atmospheric to 500 psig. Cooking initially releases acetic and formic acids from ester functionalities in the wood, and these acids further promote hydrolysis of lignin and hemicelluloses to low molecular weight species. Catalysts such as mineral acids (HCl), organic acids (acetic, formic), sulfonic acids, and Lewis acids and bases ($AlCl_3$, $Fe_2(SiO_4)_3$, $Mg(SO_4)$, $CaCl_2$) can be employed to promote delignification. Cellulose pulp is filtered from the liquor, and the solvent is recovered, usually by distillation. Lignin becomes insoluble at this point and precipitates from the liquor, leaving an aqueous solution of primarily hemicellulosic sugars.

Several processes have reached the demonstration-plant scale and none are currently used in commercial production.

Fig. 28.27 Summary of sulfite pulping processes

	Acid Sulphite	Bisulfite	Neutral Sulfite Semichemical
pH	1.5	4 - 5	8 - 9
Chemicals	H_2SO_3, $NaHSO_3$	$NaHSO_3$	Na_2SO_3, Na_2CO_3
Base	Ca, Na, Mg, NH_4	Na, Mg, NH_4	Na, NH_4
Rate of Cooking (Given temp.)	——— Decreasing ———→		
Temperature	140°C	170°C	170 - 180°C
Brightness	— Slight increase →		— Decrease →
Strength Characteristics (Given yield)	——— Increasing ———→		
Species	Nonresinous	Hardwood & Softwood	Hardwood
Equipment Corrosion	——— Decreasing ———→		

Processes [18–20] that had received much attention are briefly outlined in the following.

ALCELL process. This process based on the original concept patented by Kleinert and Tayenthal [17] is the simplest organosolv pulping system for hardwood using aqueous ethanol. It is an auto-catalyzed process (conducted at 190–200°C for 1–2 h), as the acetic acid generated from the acetyl group of hemicelluloses served as catalyst. A demonstration plant, producing 15 metric tons of organosolv pulp a day, was constructed in New Brunswick, Ontario. The flow sheet for the Alcell pulping process and by-product recovery is illustrated in Fig. 28.28 [20].

Both the Alcell and kraft processes were reported being quite comparable in pulping hardwood in terms of pulp properties (strength and bleachability) [20]. The Alcell process may have an advantage of being a more environmentally friendly process, if all the ethanol used could be quantitatively recovered. Also, the dissolved carbohydrate and lignin can be readily recovered for conversion to biofuel and bioproducts. The Alcell lignin being the only sulfur-free lignin product available in the market would be very attractive for industrial applications. However, the Alcell process being largely confined to pulp hardwood and nonwood species is not applicable for softwood.

Organocell process. This process [18, 20], developed in Germany, was a methanol-based solvent pulping intended for softwood. Wood chips were first impregnated with 50% aqueous methanol. In the original concept, the impregnated chips were subjected to the first cooking stage (operated under 40-bar pressure at 195°C) to remove approximately 20% of the lignin and a major part of the hemicelluloses. The partially cooked chips are then subjected to a second-stage cooking with an alkali charge of 15–20% based on wood in 30% aqueous methanol at 170°C. The combined residence time in both reactors is only 45 min. This two-stage concept was later modified to a totally alkaline process by cooking with a mixture of methanol, alkalis, and AQ (anthraquinone). The strength properties of the pulps produced were only slightly lower than those of corresponding kraft pulps.

ASAM process. This process stands for the Alkaline Sulfite Anthraquinone Methanol process and was derived essentially from the addition of methanol and AQ to an alkaline sulfite liquor [21]. The ASAM process was substantially more effective than the NS-AQ process [16] in delignification. It can be applied to both hardwood and softwood and produce pulps with strength properties being slightly lower than the kraft pulp.

Milox process. This is a three-stage process based on the use of formic acid to achieve both the pulping and bleaching operations [22]. Both the first and third stages use a mixture of formic acid and hydrogen peroxide conducted at 80°C, so peroxyformic acid is a major reactive species. Whereas

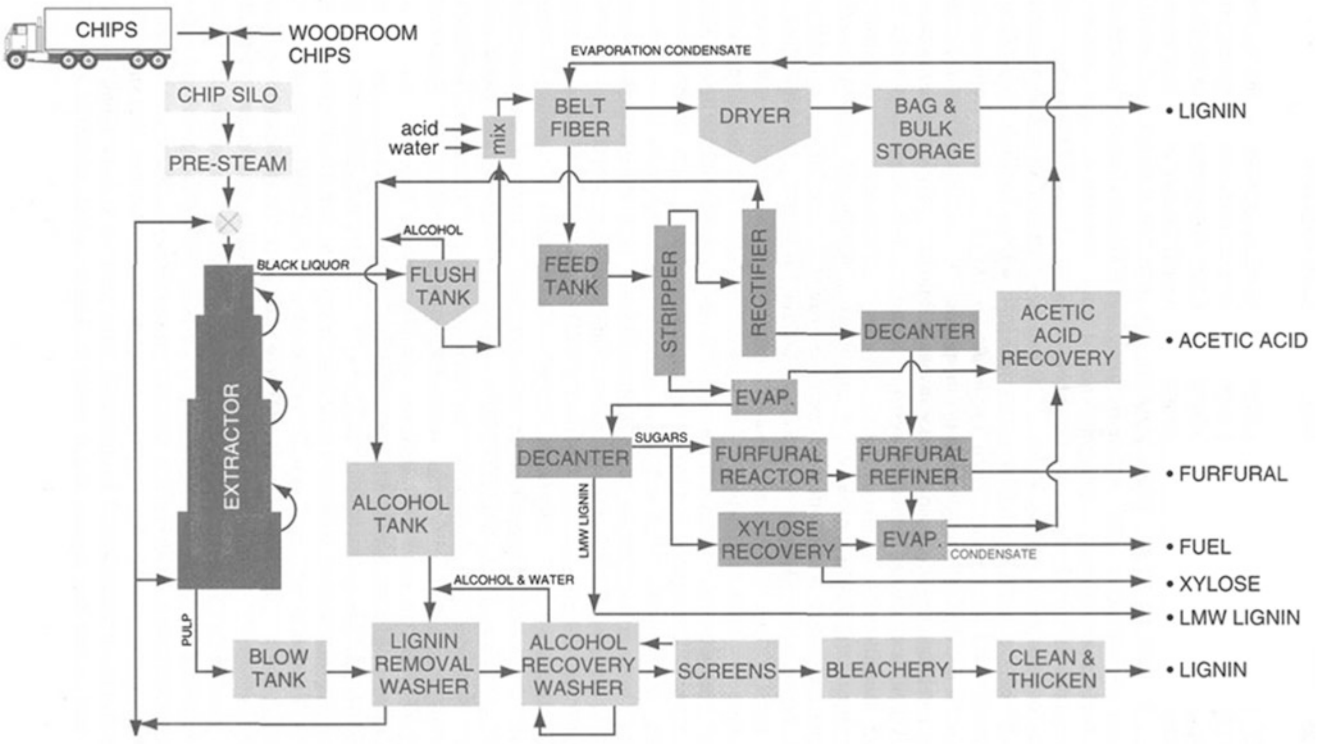


Fig. 28.28 Flow sheet for the ALCELL pulping and by-product recovery (from Hergert [20], with permission)

formic acid is employed in the second stage conducted in 100–140°C range. The Milox process has been successfully applied to all the woody substrates including hardwood, softwood, and nonwood species.

Acetic acid-based systems. A variety of organosolv processes were developed based on the use of acetic acid [20]. For example, the Acetosolv process was conducted in a 95% acetic acid solution containing about 0.15% HCl at 110°C for 2–5 h. This process because of the corrosion problem was modified to an Acetocell process by using a lower concentration of acetic acid (80–90%) conducted at higher temperatures (170–190°C).

For the Formacell process, it utilized a mixture of acetic acid, formic acid, and water in various proportions and was conducted in a series of extractors in the 130–160°C range. Although the organic acid-based systems can be operated at much lower temperatures, they may be labile to equipment corrosion and more difficult in the chemical recovery process.

Screening and Cleaning of Wood Pulp

The desired pulp fibers for papermaking applications are usually between 1 and 3 mm in length with a diameter about one-hundredth as large. Any bundles of fibers or other impurities would show up as defects in the finished paper and must be screened out. Wood knots usually are difficult to pulp and must be removed.

Screening is usually a two-stage process with the coarse material being removed by screens with relatively large perforations $\left(\frac{1}{4}\text{–}\frac{3}{8}\text{ in.}\right)$. Additional fine screening is done with screens using very small (0.008–0.014 in.) slots to ensure the removal of oversized impurities. Screen size openings will depend on the species of wood being processed and the desired quality of the end product. Because of the tendency of the fibers to agglomerate when suspended in water, it is customary to screen at very low solid contents (consistencies) of about 0.5% fiber and 99.5% water.

To meet the ever-increasing demands for cleaner pulps, centrifugal cyclone cleaners have come into almost universal use. The screened pulp is pumped through these units at low consistencies and high velocities. The fiber slurry enters the cone tangentially at the top, and a free vortex is formed with the velocity of the flow greatly increased as the diameter of the conical section is reduced. Heavier particles of sand, scale, or other dirt are forced to the outside of the cleaner and are discharged from the bottom tip through a small orifice. Because of the velocity gradients existing in the cone, the longer fibers (75–95%) are carried into the ascending center column and are discharged through the larger accept nozzle at the top. In a properly designed and operated unit, a shape separation also is made so that round particles, even though of the same specific gravity as the good fibers, will be discharged as rejects through the bottom orifice. In this way, small pieces of bark also are removed. To reduce the quantity of rejects to an acceptable level, they in turn are processed through a second, a third, or even a fourth stage of cleaners, thus holding the final loss of pulp to about 0.25–1% of the feed, depending upon quality demands and the dirt level of the incoming pulp.

Bleaching of Wood Pulp

The color of unbleached pulp varies considerably with the pulping process. It ranges from cream or tan (50–65% brightness) for the acid sulfite and groundwood pulps to dark brown (15–30% brightness) for kraft pulps. These pulps in general are too dark in color for many paper-grade products, although they could be used in the packaging grade or newsprint. The light-absorbing groups (or called chromophores) in unbleached pulps are associated mainly with the lignin component. Thus, a total removal of the residual lignin in unbleached pulps is essential to achieve a high and stable brightness (>90%). This is the purpose in bleaching of chemical pulps.

In case of the high-yield groundwood or chemimechanical pulps, they still contain practically all the wood lignin. The bleaching of these pulps is not by removal of the lignin to retain the advantage of high yield and is achieved by a modification or removal of the chromophores groups. Thus, the bleaching of high yield pulps is generally referred to a brightening process and differs distinctly from the bleaching of chemical pulps.

Bleaching chemicals. Table 28.6 lists the chemicals that have been used in the bleaching or brightening of wood pulps. The commercial bleaching of chemical pulps has gone significant developments in recent years. Among these, the chlorination (C) and hypochlorite (H) are not longer used in pulp bleaching due to the environmental concern. Current bleaching of chemical pulps relies heavily on the oxygen delignification (O), chlorine dioxide (D), and alkali extraction (E).

Although it is possible to improve the brightness (whiteness) of pulps in one stage, the economic achievement of high brightness requires the use of several stages. A bleaching sequence without the use of a chlorination (C) stage is termed as elemental chlorine free (ECF) bleaching in a sequence like ODEDE. Regulatory agencies in Europe, and particularly in Scandinavia, has imposed even greater restriction on emissions from pulp mill bleach plants and another new approach has been developed, namely totally chlorine free (TCF) bleaching of pulps. For TCF, more radical changes are necessary with the substitution of both C and D stages with oxygen (O), ozone (Z), peroxide (P), and enzyme (X) stages in a sequence such as OXZP [19].

Table 28.6 Chemicals used in pulp bleaching

Bleaching stage	Chemical	Symbol
O_2 bleaching	$NaOH + O_2$	O
Alkali extraction	NaOH	E
Alkali extraction reinforced with O_2	$NaOH + O_2$	E_o
Alkali extraction reinforced with H_2O_2	$NaOH + H_2O_2$	E_p
Chlorine dioxide	ClO_2	D
Hypochlorite	NaOCl	H
Hydrogen peroxide	H_2O_2	P
Chelating agent treatment	EDTA or DTPA	Q
Enzyme treatment	Xylanases	X
Ozone	O_3	Z
Dithionite	$Na_2S_2O_4$	Y

Bleaching of chemical pulps. After the kraft or other chemical pulping, the unbleached pulp is then subjected to an O_2 bleaching for further delignification. This oxygen stage is almost a standard practice in the industry now and is commonly referred as a prebleaching stage. A typical O_2-delignification removes about 50% of the residual lignin in kraft pulps in order to avoid an excessive degradation of the cellulose component.

For a typical ECF bleaching, a common bleaching sequence to follow is D_1ED_2 with typical conditions listed in Table 28.7. The amounts of chlorine dioxide charged depended on the content of residual lignin in the pulp and distributed properly between the D_1 and D_2 stages. For some mills, the O_2-delignification was also conducted in two separate stages aimed at further reducing the residual lignin content of pulp prior to entering the D stage.

In addition, the alkali extraction stage (E) commonly used to remove the degraded lignin can be enhanced by the addition of O_2 (Eo), peroxide (Ep), or both (Eop). For a TCF bleaching, the O_2-delignification is generally followed by an ozone (Z) and a peroxide (P) stage in a sequence such as OXZP [19].

Bleaching (brightening) of mechanical pulps. The brightening of high yield pulps is commonly achieved by treatments with a reducing agent, sodium hydrosulfite ($Na_2S_2O_4$) or with an oxidizing agent (peroxide, H_2O_2). Unlike the bleaching of chemical pulps, the brightness from a single treatment of mechanical pulps (Table 28.8) is generally sufficient for making the newsprint-grade papers. Also, a high brightness of pulps (over 80) can be achieved by a two-stage peroxide process. However, the brightness of these pulps is not stable on exposure to light and air. This yellowing problem limits many applications for the high-yield pulps.

Table 28.7 Typical conditions for bleaching chemical pulps

Stage	% Chemical on pulp	Temperature (°C)	Time (min)
O_2 Bleaching (O)	1.5–3.5 (NaOH) O_2 pressure (100 psig)	100	60
Chlorine dioxide (D)	1–2	70	180
Alkali extraction (E)	2–3 (NaOH)	70	60

Table 28.8 Typical conditions for brightening of high-yield pulps

Conditions	Sodium hydrosulfite	Hydrogen peroxide
Chemical charge, % of pulps	1	1–2
Temperature (°C)	40–100	35–60
Retention time (min)	30–60	1–3
pH	4.5–6.5	10.5–11.2
Consistency (%)	3–30	
Silicate (%)	–	5
$MgSO_4$	–	0.05

Recycling

Worldwide, approximately 38% of wastepaper is recycled; however, the rate of recycling varies between countries with the United States at a 45% recycle rate. The United States is a major exporter of waste paper, with the majority going to "fiber-poor" countries such as Korea and China that have much less virgin fibers.

A variety of problems are associated with paper recycling, such as collection, distribution, and wild cyclic swings in the market. However, with landfill sites at a premium and paper representing 30–40% of the municipal solid waste, it makes good sense in the long run to promote paper recycling, which reduces landfill needs and the consumption of virgin timber. The paper industry in the United States has now set a goal of 50% paper recycling rate.

There are a wide range of different grades of wastepaper available, depending on the source and the extent of separation. As the name implies, wastepaper designated as "direct pulp substitutes" is utilized with little treatment before reslushing in a hydrapulper. The direct pulp substitutes are the highest grade of wastepaper.

The majority of recycled paper (about 75%) is used with no attempt to remove inks, dyes, or pigments from the paper. The resultant pulp is of rather poor color and quality, and is used primarily as filler stock in paperboard. The bulk grades are the largest-volume wastepaper source.

Deinked grades of wastepaper require special techniques and equipment to remove inks, coatings, adhesives, and so on. The deinking process is complicated and time-consuming; depending on the quality of deinked pulp required, the process may involve a number of combinations of washing, flotation, dispersion, screening, and the use of cyclone cleaners. The newer noncontact inks present special removal problems, as do wet-strength agents, adhesives, and the so-called stickies.

Stickies are derived from hot-melt adhesives (vinyl acetate polymers and copolymers), pressure-sensitive adhesives (styrene-butadiene), and lattices (natural and synthetic rubber). Tackifiers and waxes also are usually included in these adhesive formulations. A variety of additives are used to help remove stickies and other contaminants from wastepaper, including solvents, nonionic and cationic surfactants, zirconium compounds (to reduce tackiness), and talc. Of course, all of the additional steps and additives add to the expense of recycling. There is also the problem of waste disposal from the deinking process, which must be properly handled.

Stock Preparation

Stock preparation in a paper mill includes all intermediate operations between preparation of the pulp and the final papermaking process [23]. It can be subdivided into (1) preparation of the "furnish" and (2) "beating" or "refining." Furnish is the name for the water slurry of fibers and other chemicals that goes to the paper machine. Beating or refining refers to the mechanical treatment given to the furnish to develop the strength properties of the pulp and impart the proper characteristics to the finished paper.

Cellulosic fibers are unique in that, when suspended in water, they will bond to each other very strongly as the water is removed by filtration and drying, with no need for an additional adhesive. This is so because of the large number of hydrogen bondings that form between the surfaces of fibers that are in close contact as the water is removed. This bonding is reversible and accounts for the well-known fact that paper loses most of its strength when wet. If paper is suspended in water and agitated, it will separate into the individual fibers, a behavior that allows the easy reuse of wastepaper or the processing waste from the paper mill.

Refining. In order to enhance the bonding capability of the fibers, it is necessary to mechanically beat or refine them in equipment such as beaters, jordans, or disc refiners. This treatment of the pulp slurry at about 3–6% consistency is done by passing the pulp between the two rotating surfaces of the refiner. These surfaces contain metal bars and operate at very close clearances. As the fibers pass between the bars, they are made more flexible, and a larger surface for bonding is developed by the mechanical action.

This refining brings about fundamental changes in the pulp fibers and increases the degree of interfiber bonding in the final sheet of paper. Thus, the final properties of the paper can be significantly changed by varying the degree and type of refining. As additional refining is performed, properties such as tensile strength, fold, and density are increased while tear resistance, opacity, thickness, and dimensional stability are decreased. Thus, the proper refining conditions must be selected to bring out the desired properties without detracting too much from other properties.

Furnish. The furnish for a paper machine varies widely, depending on the grade of paper being made. Newsprint usually consists of about 80% groundwood and 20% chemical fiber (sulfite or semibleached kraft). Bag papers and linerboard are usually 100% unbleached softwood kraft. Printing papers are made from bleached pulps and contain both hardwoods and softwoods. By selecting the proper pulps and refining conditions, a wide variety of paper qualities can be achieved.

Additives. The paper industry is a large user of chemicals because relatively small quantities of additives can materially change the properties of paper. The use of 1–2% rosin size and 2–3% alum [$Al_2(SO_4)_3$] will greatly increase the resistance of paper to penetration by water or ink. Pigments such as kaolin clay, calcium carbonate, and titanium dioxide are added in amounts up to 15% to increase opacity and give a better printing surface. Organic dyes and colored pigments are added to produce the highly colored papers used for business and printing papers.

Other additives such as wet-strength resins, retention aids, and starch can be used to give particular properties that are needed. Thus, in order to produce the wide variety of grades of paper now available, the papermaker selects the proper pulps, refining conditions, and additives, and then combines the pulp and additives before sending them on to the paper machine for the final step in the process.

Papermaking Process

Some paper mills are not integrated with pulp mills, and it is necessary for these mills to use dried, baled pulp manufactured at a separate location. Many mills making limited quantities of highly specialized papers fall into this category, as it allows maximum flexibility in selecting the optimum pulps for a particular paper grade. However, the papermaking process is the same regardless of the source of pulp.

After the furnish has been prepared with the proper refining treatment and additives, it is stored in the machine chest and then fed continuously into the paper machine system. A refiner or jordan is placed in this line to give the paper machine operators the opportunity to make small adjustments in the quality of the furnish as needed to give the desired paper properties. Screens and centrifugal cleaners also are included to ensure a high-quality paper.

The papermaking process is essentially a system whereby the pulp is diluted to a very low consistency (about 0.5%) and continuously formed into a sheet of paper at high speeds, and then the water is removed by filtration, pressing, and drying. The basic units of the fourdrinier paper machine are diagrammed in Fig. 28.29 and a picture is shown in Fig. 28.30.

Wet end. The section of the paper machine, where the sheet is formed, is referred to as the “wet end.” The fourdrinier machine is characterized by a headbox that allows the diluted stock to flow through an orifice (slice) onto the flat moving wire. This is actually an endless wire belt that returns on the underside of the machine, thus allowing the process to proceed continuously. Because a low headbox consistency (about 0.5%) is necessary for good formation, the volumes of water handled are very large (about 10,000 gal/min for a machine producing 300 t/day). Much of the water is removed through the wire by the action of the table rolls and foils that support the wires in the forming area immediately following the headbox. At this point the stock consistency has been raised to about 2%, and the stock no longer drains freely. By passing it over suction boxes operating at fairly high vacuum (6–8 in. Hg), the consistency is raised to about 15–20%.

A dandy roll (covered with woven wire) rotating on top of the wet paper is used to improve the formation and can impact a watermark if it contains the proper pattern. The suction roll after the suction boxes has several functions: it removes additional water; it serves as the driving roll for the wire; and it serves as the point at which the paper separates from the wire and passes into the press section while the wire returns to pick up additional pulp at the headbox.

Modern fourdrinier machines are available up to a width of 360 in. and can operate at speeds up to 3,000 ft/min. Newsprint machines are usually the widest and fastest, whereas hose making heavier grades such as linerboard run somewhat more slowly. In the United States, there are several machines capable of producing over 1,000 t/day of linerboard. More specialized grades such as bond and printing papers are usually produced at a lower speed on narrower machines, and 150–250 t/day would be considered a high output.

Many specialized grades such as filter paper and tracing paper are produced on very small, slow machines producing

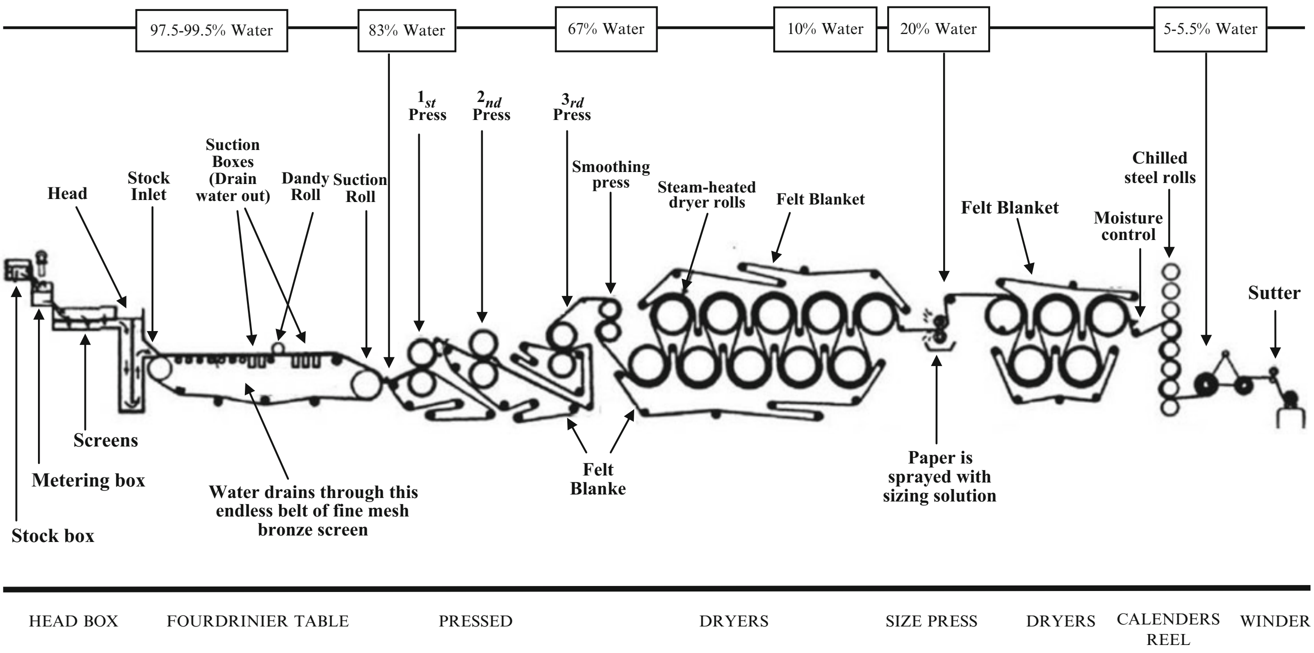


Fig. 28.29 Schematic of a fourdrinier paper machine

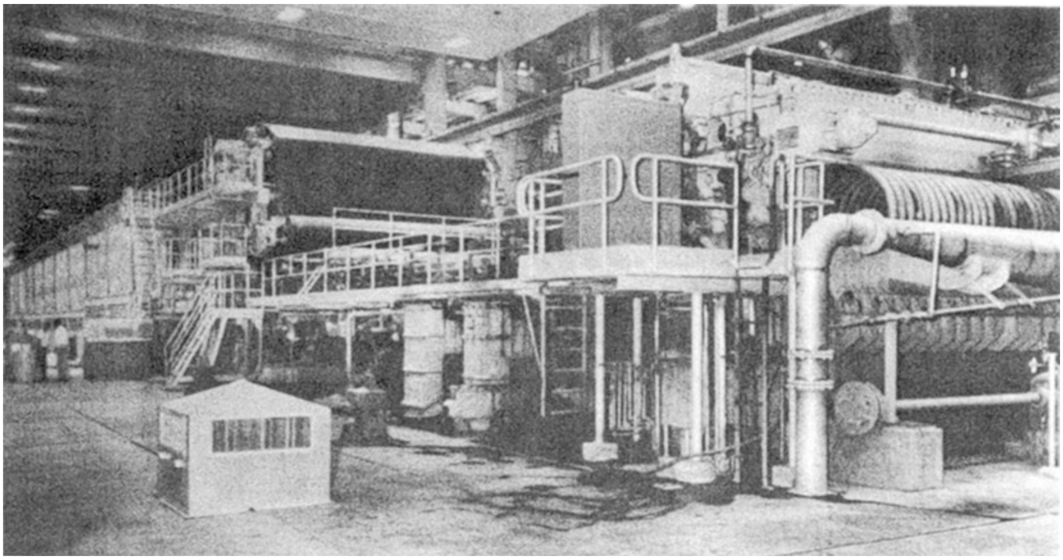


Fig. 28.30 Modern high-speed paper machine (Courtesy Beloit Corp.)

only a few tons per day. Machines making tissue paper for sanitary purposes use modifications of the standard fourdrinier to produce tissue at speeds of up to 5,000 ft/min. Because of the light weight of this paper, it is necessary to make many modifications in the equipment shown in Fig. 28.29.

Other sheet forming methods. Many new methods of forming a sheet of paper are being developed. The most common concept is to introduce the pulp slurry between two converging wires and remove the water from both sides. This is termed “twin-wire” forming, and many variations are now in operation. This technique has many advantages and may eventually replace the fourdrinier.

A second method of forming paper uses the cylinder machine. Actually these machines are used primarily to form the multiply board used in packaging such as cereal boxes. The cylinder wet end consists of one or more (up to eight) cylinder vats, each of which forms a separate wet web of fibers. Each vat contains a woven wire-covered cylinder

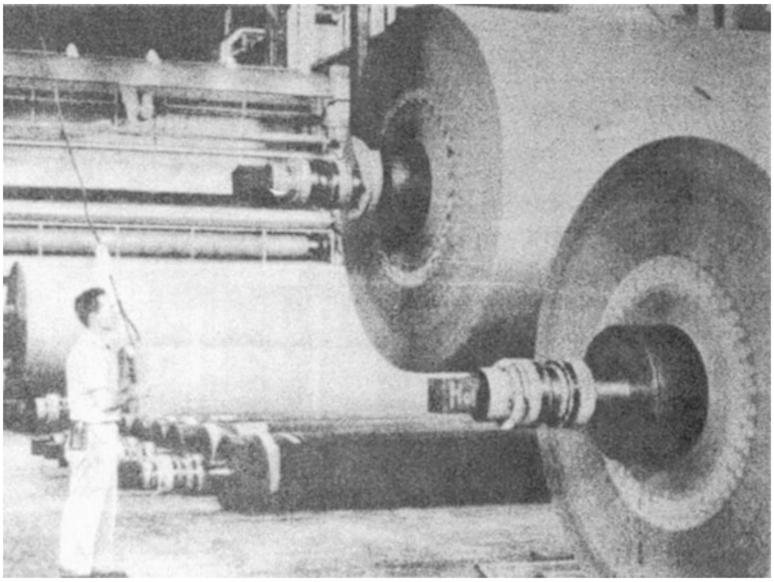


Fig. 28.31 Dry end of a large Fourdriner paper machine (Courtesy of Regis Paper Co.)

rotating in the diluted pulp slurry. The liquid head on the outside of the cylinder is greater than that on the inside, resulting in a flow of water through the wire and out of the vat. The pulp mat that is formed on the face of the cylinder is removed by an endless woolen felt that moves in contact with the cylinder by means of a rubber roll riding on top of the felt.

With machines having more than one cylinder, the same felt moves from one cylinder to the next, and the wet sheets from each cylinder are laminated to each other on the bottom side of the felt. Thus, very heavy papers or boards can be fabricated by multivat machines. Because each vat can be supplied with a different type of pulp, it is possible to make boards with a white surface of high-quality pulp and a center of low-cost pulp from waste newspaper or other cheap grades. Because of hydrodynamic problems, cylinder-machine speeds are limited to about 250 ft/min and widths of about 150 in. However, because of the heavyweight board produced and the widespread use of cheap waste paper for most of the furnish, the cylinder machine is widely used. New forming units looking like miniature fourdrinier units (Ultraformer, Inverformer) have been developed and are rapidly replacing the old-fashioned cylinder vat because their speed is not so limited.

Press section. From the wet end of the machine, the wet sheet is conveyed by woolen felts through a series of roll-type presses for further water removal, increasing the consistency to about 35%. The sheet then is threaded through the dryer section, consisting of a long series of steam-heated cast iron cylinders that reduce the moisture content to approximately 5%, which is about the equilibrium moisture content for cellulosic fibers at 40–50% relative humidity. Tissue machines use one large dryer (called a Yankee dryer), ranging from 8 to 18 ft in diameter. Because of the light weight of the tissue paper, it can be dried at high speeds on a single dryer.

After drying, the paper is compacted and smoothed by passing through a calender stack consisting of a vertical row of highly polished cast iron rolls. Then the paper is wound into rolls on the reel, as shown in Fig. 28.31.

Furnishing and Converting

The objectives in the final stage of the total papermaking system are to improve the paper surface, to reduce rolls and sheets in size, to modify paper for special properties, such as coat or emboss, to convert to finished products, such as bags and corrugated boxes, and to package for shipping. Corrugated boxboard is made by gluing sheets of linerboard to each side of a fluted sheet of the corrugating medium. Papers for packing may be laminated to polyethylene film or aluminum foil, or coated with waxes and hot-melt resins. The printing and bag- and box-making industries depend on the production of many mills that produce the several hundred grades of paper used in the United States, and each user may require special paper characteristics to match its process.

Surface treatments. The quality of many papers is improved by a surface treatment. A size press about two-thirds of the way along the dryer section can apply a solution of starch to improve surface bonding. Also, paper generally is coated to improve printing properties. A surface coating of a pigment (usually kaolin or china clay, calcium carbonate, or titanium dioxide) and an adhesive (such as starch, casein, and others) is applied to the partially dried web by brush, blade, spray, or

Table 28.9 Summary of pollution abatements in pulp and paper mill operations

Type of pollutant	Mill operation	Treatment
Effluents		
Suspended solids (SS): fiber fragments, inerts, clay	Papermaking	Primary—clarification
Pulping residuals	Pulping	Secondary—biological treatment and clarification
Air emissions		
Total reduced sulfur (TRS)	Kraft liquor recovery	Oxidation, precipitation, scrubbing, incineration
Sulfur dioxide	Sulfite pulping	
Particulates	Steam generation	Precipitation, scrubbing
Solid wastes	Effluent treatment	Landfill, utilization, incineration
Toxic wastes: chlorinated compounds	Bleaching	Lime pretreatment, oxidation, biological

other method, and dried during the papermaking (on-machine) operation or in a separate operation. The paper surface is brought to a high finish by passage through the calenders referred to above or through supercalenders. Supercalenders are stacks of alternate steel and densified fiber rolls that create a rubbing action on the sheet, imparting an extra-high gloss to the sheet surface.

Environmental Protection

The manufacture of pulp and paper is a chemical process industry and produces air emissions, effluents, and solid and toxic wastes that are potential hazards. The paper industry uses large volumes of water as a fiber carrier and as a dispersing medium. An increasing volume of water is being recycled, but makeup water still is required to cover losses. A bleached-pulp and paper mill may use 100 m^3 (26,400 gal) of fresh water per metric ton of product and 50,000 m^3 (13.2 million gal) daily for a plant producing 500 metric tons of products.

In addition to the aqueous effluent that the mill must clean up, it must contend with polluted air and solid and toxic wastes. The nature of these emanations, their sources, and their treatments are summarized in Table 28.9. In this connection the paper industry generally has been in good compliance with government environmental regulations although at considerable nonreimbursed capital expense, which amounts to about 10% of the cost of the mill.

Using a revolving cylinder or other equipment (save-alls) in the 1930s, mills recovered for reuse fibers and clay from the paper water (white water) system. In the 1970s, procedures for removal of the fibers and clay from the paper mill effluent were incorporated through settling or clarification or primary effluent treatment. About the same time, secondary effluent treatment (biochemical treatment) of the pulp mill effluent was necessary to remove pulping residuals. The purpose of this treatment is to reduce the biological oxygen demand (BOD) of the effluent, which, if untreated, reduces the oxygen content of the stream to a level incapable of supporting aquatic life. The most common method uses microorganisms that react with the wood sugars and other oxygen-consuming compounds in the spent liquors; this is called the activated-sludge method. The products of primary and secondary treatments are sludges, the handling of which is discussed later under solid wastes.

Two objectionable air emissions have characterized pulp mills for years: the sulfur dioxide of the sulfite pulping mill and the malodorous reduced sulfur compounds (TRS) (mercaptans and hydrogen sulfide) of the kraft mill. Still another less noxious air emission is the particulate matter from steam boilers. Coal-burning boilers also emit sulfur dioxide, as is well known. The treatments for the particulate emissions are given in Table 28.9.

Solid wastes represent the ultimate in mill residues and include the accumulated refuse of the mill and the sludges from primary and secondary effluent treatment. There is difficulty in removing water from the secondary sludge; the primary and secondary sludges often are mixed to aid in water removal, which is important if the sludge is to be incinerated for disposal. The sludges from pulp and paper mills are handled mostly as landfill, and sometimes, if not toxic, they are spread for agricultural purposes.

Since chlorine is no longer used in pulp bleaching, the toxicity of most mill solids is greatly reduced. This toxicity can be further reduced with lime pretreatment and biological treatment. It was a main concern of governmental regulating bodies in the past.

Chemical Commodities Associated with Paper Industry

Although paper is the dominant use of pulp fibers, a variety of commodities are produced from the pulping operations. Besides papermaking fibers, the paper industry produces a specialty pulp called dissolving pulps (pure cellulose) for making a variety of cellulose products. To prepare this cellulose pulp, a prehydrolysis stage to remove the hemicellulose components is needed prior to the kraft delignification [23].

Also after pulping, the black liquor contains, in addition to residual chemicals, significant amounts of dissolved wood components that are useful for many applications. The chemical nature of these organic materials differs considerably between the acid sulfite and kraft process and the wood species.

Cellulose Products

Major uses of dissolving pulps are in the making of the so-called regenerated cellulose and cellulose derivatives [24]. Traditionally, the regenerated cellulose is prepared from cellulose by a series of mercerization (treating with high concentration of alkali), dissolution (by reacting with carbon disulfide in alkalis), and precipitation in acid bath in the form of film (cellophane) or fiber (rayon).

Recently, a new alternative to prepare the regenerated cellulose is by a direct dissolution of cellulose in N-methyl-morpholine-*N*-Oxide (NMMO) followed by a subsequent precipitation of the cellulose filaments in a NMMO–water mixture [24]. These regenerated fibers are called Lyocell fibers, which like rayon fibers are used mainly in the textiles applications.

A variety of cellulose derivatives has been produced including cellulose acetates (for film), cellulose nitrate (explosives), and carboxymethylcellulose (CMC). The latter CMC being a water-soluble polymer is widely used as dispersing or thickening agent in the food, pharmaceutical, and cosmetic industry.

Products from Sulfite Pulping Liquor

Lignosulfonate. The current world-wise production of linosulfonate is about 1,000,000 t/year and the Borregaard Lignotech is the major producer with facilities in six different countries [24]. Lignosulfonates are used mainly in areas of industrial binders and dispersing agents. Also, lignosulfonates have been used to produce vanillin by alkali treatment at elevated temperatures under oxidative conditions. The production of vanillin is around 3,500 t/year [24]. It is used as flavoring agent for food and beverages. Also, vanillin can be obtained from natural sources such as vanilla beans.

Dissolved carbohydrates. In acid sulfite pulping, a portion of wood polysaccharides was hydrolyzed to hexoses and pentoses. These sugars have been used to produce ethanol. The Domsjö mill in Sweden produces the ethanol about 10,000 t/year [24].

Products from Kraft Pulping Liquor

Many useful products can be derived from the kraft pulping liquor of softwood. These include turpentine and tall oil from the wood extractives as well as dimethyl sulfoxide (DMSO) from the kraft lignin. The production of kraft turpentine and tall oil will be discussed in more detail in a later section on "Naval Stores."

Kraft turpentine. It contains mainly the volatile monoterpenes and was isolated in a yield of approximately 10 kg/t of pulp. The major uses of turpentine are in the area of resins, fragrances, and flavors especially for the pine oil production.

Tall oil. It is derived from the nonvolatile extractives, mainly soaps of resin acids and fatty acids. They are separated from the black liquor by acidification. The yield of tall oil might reach ~50 kg/t of pulp. Main components of tall oil are rosin and fatty acids, whereas the remainder forms tall oil pitch. Much of the rosin products are used in the sizing of paper as adhesives and in printing inks whereas a major use of the fatty acids is in alkyd resins.

Kraft lignin. It can be isolated from the black liquor by precipitation with flue gas or acid. Only a few mills produce the kraft lignin due to its limited applications. At present [24], the total production is in the order of 100,000 t/year from Westvaco (USA) and Lignotech (Norway). The major use of kraft lignins is as a dispersant in aqueous solution. Quite often, kraft lignin is further modified by sulfonation to yield products similar to the lignosulfonate derived from the acid sulfite process.

Dimethyl sulfoxide. Dimethyl sulfoxide (DMSO) is an important solvent for many applications including a medicinal of interest in the treatment of arthritis. It is produced from the oxidation of dimethyl sulfide (DMS). DMS is an odorant for natural gas and a solvent recovered from the black liquor. The content of DMS in the black liquor can be increased by reaction with elemental sulfur.

Board and Structural Materials

Lumber, boards, sheets, panels, and other structural materials are manufactured from solid wood and wood veneers, fibers, and particles. Traditional industries producing these products are not generally classified as chemical industries. Recent trends have been on the development of using wood as an engineering material. Among many wood-based composites, the production of wood–polymer composites is one of fast growing segments of the forest

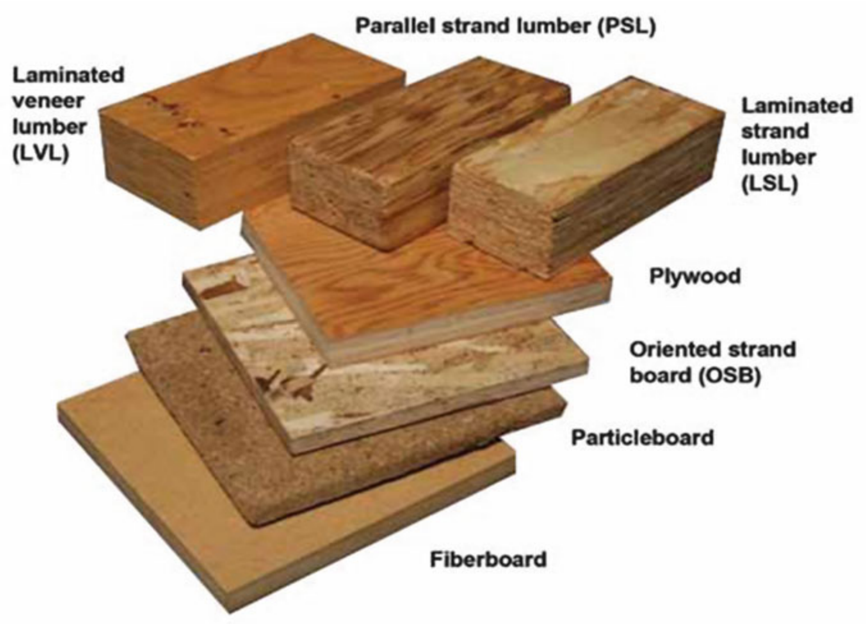


Fig. 28.32 Examples of various composite products. From *top left*, clockwise: LVL, PSL, LSL, plywood, OSB, particleboard, and fiberboard (from Stark et al. [25])

product industry and is a close association between the wood and chemical industry. The composite products represent a more efficient utilization of renewable resources, as all types of wood materials can be applied. Thus, wood adhesives play an increasing role in the forest products industry. Figure 28.32 illustrates the examples of some wood-based composites [25].

Lumber

A log, when sawn, yields round timber, ties, or lumber of varying quality. There are two different types of mills (hardwood and softwood) for manufacturing lumber. Hardwood lumbers are used commonly for flooring, paneling, or manufacture into furniture, cabinetwork, or pallets, whereas softwood lumbers are suitable for both construction and manufacture applications.

A hardwood sawmill typically produces 1–15 million board feet annually. Most of the lumber produced is 4/4 thick (4/4, pronounced “four quarter,” is the number of quarter inches of thickness; 8/4 would be 2 in. thick). Lengths run from 4 to 16 ft; width is random. The primary quality factor is the amount of large clear areas. The lumber is dried and then sawn into smaller pieces for furniture, millwork, and cabinets.

A softwood sawmill typically produces in excess of 200 million board feet annually. Most of the lumber is 8/4 nominal thickness (would be sold as 2 in. by a specified width), although the actual thickness is 1.5 in. Lengths are usually 8–20 ft; widths would typically be 4, 6, 8, or 10 in. nominal size or 3–1/2, 5–1/2, 7–1/2, and 9–1/4 in. actual size. Most softwood lumber will be dried, planed, and graded, and then used “as is” without additional manufacturing, except cutting to length at times.

A modern softwood sawmill producing lumber for construction is highly automated, with high manufacturing efficiency and high production, thereby assuring good profitability and good stewardship of the natural resource. In a typical mill, logs are debarked, scanned for metal, scanned for size to determine the best sawing pattern, and sawn into lumber and cants at the headrig. Then cants are resawn into lumber. Any pieces with wane (wane is the absence of wood) are edged. Lumber is then trimmed and sorted by size. Most of these processing procedures are computer controlled with manual override possible. Volume conversion efficiencies can exceed 65%; the 35% of the log not converted into lumber is used for pulp and for fuel. After the green lumber is produced, the lumber is kiln dried, planed, and graded according to the rules of the American Lumber Standard.

Wood Adhesives

Adhesives play a decisive role in making composite materials and their performance. Thermosetting (heat-curing) resins are commonly used as adhesives in the manufacture of wood-based composites [25]. These include the phenol-formaldehyde, urea-formaldehyde, melamine-formaldehyde, and isocyanate resins.

Phenol-formaldehyde (PF) resin. The PF resin is typically used in the manufacture of construction plywood and oriented strandboard (OSB). The slow-curing process of this phenolic resin generally requires a higher temperature and longer press time than other resins. Cured phenolic resins remain chemically stable at elevated temperatures. The darker color of phenoic resin as compared with other resins may be less attractive for product applications such as interior paneling and furniture.

Urea-formaldehyde (UF) resin. The UF resin is commonly used in the manufacture of products for interior applications. It is lighter in color and easier to cure but more sensitive to moisture and heat [25]. Excessive heat exposure could breakdown the resin and release the formaldehyde, which is a growing health concern.

Melamine-formaldehyde (MF) resin. This resin is more expensive than the PF resin and is used primarily for decorative laminates and paper coating. Also, it is often used as an ingredient to improve the performance of UF resin.

Isocyanate-based resin. This isocyanate adhesive is a polymeric methylene diisocyanate (pMDI). It is more costly than the PF resins but has an advantage of being easier to cure. Although the cured pMDI resin poses no recognized health concerns, special precautionary protective measures are needed for the facility handling the pMDI, as the uncured resin can result in chemical sensitization for persons exposed to it [25].

Future perspectives. Current emphases on using renewable (green) products have stimulated a renewed interest in using bio-based adhesives [25]. The protein-based glues were widely used in the construction plywood prior to the early 1970s. They were replaced in mid-1970s by the PF resin being a better bonding resin. However, the performance of soy protein-based resin systems has been improved considerably in recent years and could be quite comparable to phenolic resin in performance now.

Also, useful adhesives have been prepared from the phenolic lignin (a major wood component) or tannin (being rich in wood bark). Extensive efforts are continued to improve the performance of these natural adhesives. The development of lignin-based adhesives is of special significance, as plenty of technical lignin is readily available as by-products from the pulping operations.

Laminated Timber (Glulam)

Glulam is one of the oldest glued engineered wood products [25]. It consists of two or more layers of lumber that are glued together with the grain of all layers parallel to the length. The maximum lamination thickness is about 2 in. A major advantage of this process is that the larger size of timber can be produced from small-diameter trees. By combining the lumber in glulam, the production of large structural elements like 100 ft is very common. This represents an improved utilization of forest resources.

Structural Composite Lumber

The structural composite lumbers [25] are produced from gluing smaller pieces of wood together into sizes common for the solid-sawn lumber. This process like Glulam greatly enhances the utilization of small trees. Different composite lumbers have been produced based on the nature of wood elements used.

Laminated veneer lumber (*LVL*). This LVL lumber is produced by laminating veneer with all plies parallel to the length. Specially graded veneer are selected for the product to achieve the desired engineering properties.

Parallel strand lumber (*PSL*). PSL is a composite of wood strand elements with fibers oriented primarily along the length. The least dimension of the strands must not exceed 0.25 in. The product is often used for large beams and columns, typically as a replacement of solid-sawn lumber or glulam.

Laminated strand lumber (*LSL*) *and oriented strand lumber* (*OSL*). Both LSL and OSL represent an extension of the technology used to produce OSB. The main difference is that the aspect ratio of strands used in LSL is higher than for OSL. For example, one type of LSL uses strands about 12 in. in length, which is somewhat longer than that commonly used for OSB. The LSL and OSL products are used for band joists in floor construction and as substitutes for studs and rafters in wall and roof construction.

Conventionally Wood-Based Composite Panels

Wood-based composites panels contain several products that differ primarily on the size of wood elements used. These variations lead to the production of plywood, oriented strand board, particle board, and hardboard (Fig. 28.33) [25].

Plywood. Plywood panels consist of several large thin wood veneer-type layers joined with adhesive. Some or all of the layers are sheets of veneer. Other layers, particularly in the core, may be particleboard, hardboard, lumber strips, or special materials. The fiber direction of each layer is at right angles to that of the adjoining layer. This cross-banding makes plywood more uniform and less anisotropic than

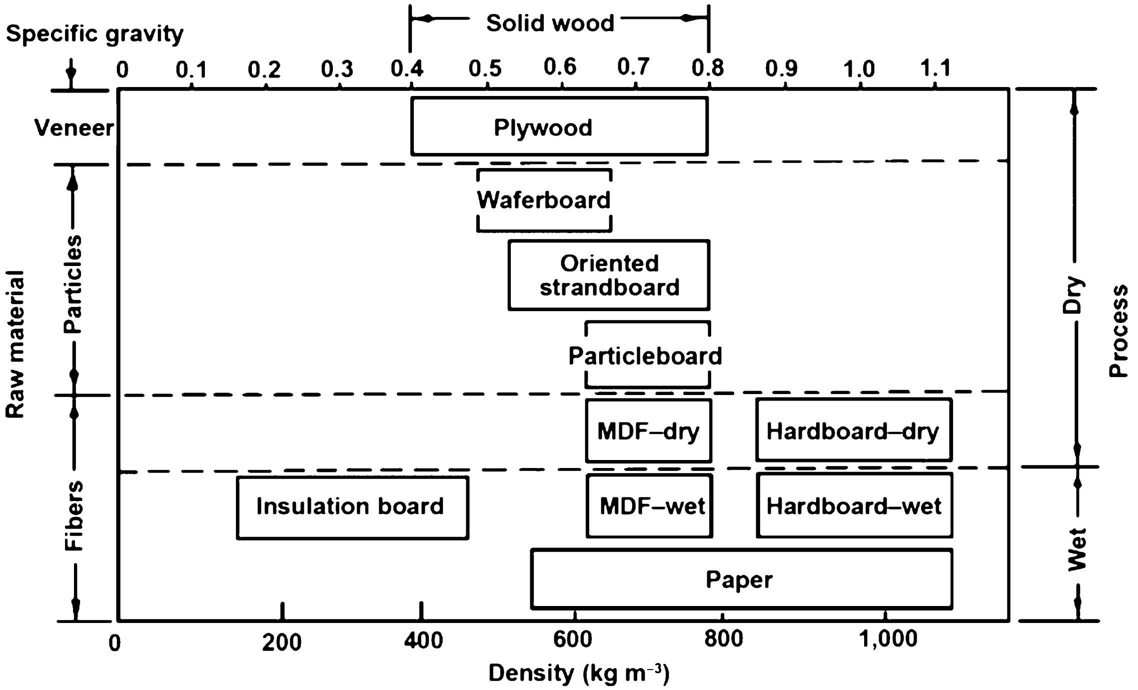


Fig. 28.33 Classification of wood composite panels by particle size, density, and process [76]. Note that insulation board is now known as cellulosic fiberboard (from Stark et al. [25])

lumber; its properties in the direction of panel length resemble those in the direction of panel width. Cross-banding affects strength in a logical way in both directions of the plane.

The panel is as strong as the combined layers. Because in any one of the two directions transverse layers contribute practically no strength, plywood is roughly one-half as strong as lumber is lengthwise. But by the same principle plywood is stronger than lumber in the direction of width and can therefore be thinner. Moreover, it does not split like solid-wood products. Plywood and lumber properties naturally are the same in the thickness direction, provided that the layers are adequately bonded together. Cross-banding also imparts dimensional stability to the plywood.

Many plywood panels are 6 mm (1/4 in.) thick; the thinnest measure about 1 mm, and the thickest several centimeters. In addition to house sheathing and siding, much plywood is used in cabinets, billboards, furniture, bookshelves, concrete forms, skins of flush doors, paneling, boxes, in mobile homes and for trailers. Plywood is also used as a component in other engineered wood products like the LVL mentioned earlier.

Oriented strandboard (*OSB*). OSB is an engineered product from thin wood strands bonded together with a water-resistant resin. The wood strands typically have an aspect ratio (strand length divided by width) of at least 3. Since small trees and crooked stems can be effectively utilized, the raw material cost, compared to product cost, is about 38%. Plywood requires more than double the raw material cost. Lumber requires even more. As a result, OSB has rapidly gained acceptance in the construction market. OSB plant capacities exceed 20 million cubic meters annually; more new plants are being constructed each year.

The typical material flow in an OSB plant provides very high conversion efficiencies. Any “waste” is used for fuel for the manufacturing process. Logs are debarked, heated in hot water, cut into short lengths, flaked into wafers or strands using knives. The strands are dried in large drum driers, coated with adhesive, and assembled into a thick mat. The strands in the mat are oriented to provide parallel strands in the various layers in the mat. The mat is compressed and heated, and cut to size. By controlling the panel density, the amount of adhesive, and the strand orientation, a panel can be designed and produced with properties that meet the engineering needs of the building designer and engineer [2].

If lumber were used to do to the same job as OSB, an excess of eight billion more board feet of lumber would be needed annually, and this lumber would have to come from larger, higher quality, more expensive trees. It is easy to understand why OSB is considered as the wood product of future in the construction business.

Similarly, the LVL, PSL, and LSL, and OSL are also rapidly moving into the market as engineered wood products [26, 27].

Particleboard. Particleboards are composed of discrete particles of wood bonded together by a synthetic resin adhesive, most commonly urea-formaldehyde or phenol-formaldehyde. The material is consolidated and the resin cured under heat and pressure. The strength of the product depends mainly upon the adhesive, although the size and shape of the particles influence strength properties. They may be fine slivers, coarse slivers, planar shavings, shreds, or flakes, and they are divided into two main groups: (1) hammer-mill-produced particles (slivers and splinters from solid wood residues, feather-like wisps to block-shaped pieces from planar shavings) and (2) cutter-type particles, sometimes termed "engineered" particles (flakes and shreds).

Hammer-milled particles usually vary appreciably in size. Dry raw material produces greater amounts of fines than green wood. Cutting machines (either cylinder-type or rotating-disc-type) give more uniform particles, with the length dimension in the direction of the grain of the wood. The thickness, size, and shape of particles influence the strength of the board. Boards made from sawdust have the lowest strength properties, hammer-milled particles give boards of intermediate strength, and solid wood, cut to flakes, gives boards of highest strength. The latter often are referred to as wafer boards.

Particleboards may be made in a wide range of densities. Low-density or insulating types are a comparatively recent development in Central Europe, whereas the high-density-hardboard types are a U.S. development. Most particleboard production is in the middle-density range.

Particleboards most commonly are used as core stock for veneer in furniture and in doors, as interior panels for walls and ceilings, as subflooring, as sheathing and siding, and as components in interior millwork. The dense types are used in the same way as fiberboard hardboard, described earlier. Both dense particleboards and hardboards, after receiving a surface coating, may be printed with decorative designs.

Particleboard production has increased rapidly, both in the United States and worldwide. Until recently the most significant limitation on the market for particleboard was the availability of inexpensive plywood. Now plywood is no longer inexpensive relative to particleboard, and the cost factor would seem to favor continued growth of the particleboard industry.

Fiberboard. Fiberboard is the name for rigid or semirigid sheet materials of widely varying densities and thicknesses manufactured from wood or other vegetable fibers. The board is formed by the felting of the fibers from a water slurry or an air suspension to produce a mat. Bonding agents may be incorporated to increase the strength, and other materials may be added to give special properties, such as resistance to moisture, fire, or decay.

Fiberboards are manufactured primarily for panels, insulation, and cover materials in buildings and other structures where flat sheets of moderate strength and/or insulating capacity are required. They also are used as components in doors, cupboards, cabinets, furniture, and mill work. Fiberboard is normally classified by density and can be made by either dry or wet processes.

Dry processes are applicable to boards with high density (hardboard) and medium density (MDF). Wet processes are applicable to both high-density hardboard and low-density cellulosic fiberboard. These products are manufactured for many specific uses, and they vary accordingly.

The dry-process fiberboard is made in a similar fashion to the formation of particleboard. In general, the resin and other additives are applied to the fibers by spraying in a blender or refiner. The adhesive-coated fibers are then air-laid into a mat followed by pressing.

The wet-process hardboard differs from the dry-process fiberboard in several ways. First, water is used as the fiber distribution medium for the mat formation. The technology is really an extension of papermaking process. Second, some wet-process boards are made without additional binders, as the hemicelluloses and lignin present may serve as the binder. Also, under heat and pressure, the lignin will further serve as a thermosetting adhesive. Thus, refining plays a significant role in developing the strength of the wet-process hardboards.

For making the cellulosic fiberboard, the process is similar to the papermaking technology. It does not require a binder as in the wet-process hard board. After drying, the board may be treated for specific applications. Cellulosic fiberboard products include sound-deadening board, insulation boards, structural and nonstructural sheathings, backer board, and roof decking in various thicknesses.

Wood-Polymer Composites

The wood-polymer composites (WPC) represent a class of composites containing a wood and a polymer component [25, 28, 29]. The wood element is usually a flour or fiber. Among others, the plastic polymers are of special significance and consist of two major types (thermoset and thermoplastic). The thermoset polymers are the plastics and once cured cannot be remelted by heating. These include the phenolic resin used as adhesive in making the wood-based composites discussed earlier. In contrast, the thermoplastics can be melted repeatedly and are used in making many diverse commercial products ranging from milk jugs, grocery bags to the siding for homes. The production of these wood thermoplastic composites is one of the fast growing segments of the wood industry.

Wood elements. Wood flour is the most commonly used in the WPC application and contains mainly very short fiber bundles. It can be produced from any woody materials including scrap wood from various processing plants (planer shavings, chips, and sawdust). In case of using wood fibers, they can be produced by a variety of pulping methods (Table 28.4) depending on the desired quality of the WPC. Fibers can be used instead of flour and serve as reinforcement of the matrix. They enhance the mechanical property but fibers are more difficult to handle due to the low bulk density. More intensive mixing and the use of special feeding equipment may be necessary to handle longer fibers.

Thermoplastic polymers. Due to the low thermal stability of wood flour, the plastics that can be processed at temperatures below 200°C are commonly used in the WPC. These polymers include the polyethylene (PE), polypropylene (PP), polystyrene (PS), and polyvinyl chloride (PVC) and the PE is wildly used in North America.

Wood-thermoplastic composites (*WTC*). WTC has two types based on the relative proportion of wood elements and polymers [25]. The first type has a high thermoplastic content (>60%) and the wood element serves as reinforcement of the plastic matrix or filler. In another type, the thermoplastic serves as a binder for the wood elements much like the conventional wood-based composites.

The bulk of commercial WTC products have a high thermoplastic content. A variety of WTC products are manufactured by varying the forming processes. The most common types of product-forming involve forcing molten material through a die (sheet or profile extrusion) or into a cold mold (injection molding); or pressing in calendars (calendering) or between mold halves (thermoforming and compression molding). A profile extrusion yields products such as decking, siding, railings, and window profiles. An injection-molded process produces the consumer household goods and furniture parts, which are gaining acceptance.

Also, a major application of the molded products would be for interior uses in automobiles, such as head liners, door panels, and dashboards. Although this is a low-cost and low-performance application, it represents a very large-volume market. Indeed, wood is already utilized in applications of this type, but as a finely ground flour that serves as filler (up to 40%) in extrusion-molded polyolefin products. The use of recycled fiber offers the potential of even greater cost reductions and could also alleviate the solid waste disposable problems.

Modified Wood

Wood has a number of disadvantageous properties, as described earlier in this chapter. It is biodegradable, flammable, dimensionally unstable to water, and degradable by ultraviolet light, acids, and bases. There has been a considerable amount of research expended over the years to correct these defective properties for both solid wood and reconstituted wood products.

Heat treatments. A wood product of improved dimensional stability can be produced by heat treatment of wood at temperatures varying from 150 to 320°C [30]. For good temperature control and avoiding the strength loss, a heat treatment is performed under the surface of molten metal or a fused salt for exclusion of oxygen. The modified wood product is referred as Staybwood in United States or Thermonwood in Finland. The heat treatment causes the loss of water followed by modification and degradation of the wood components, especially the hemicelluloses. The wood could become brittle from an extensive heat treatment.

Polyethylene glycol treatment (*PEG*). Treatment of wood with polyethylene glycol (PEG) is the most commonly used method to impart dimensional stability to wood. PEG in a water solution can be used to permeate the lumens and penetrate the cell wall. A 30–50% solution of PEG in water generally is used on green wood; no pressure is required. Treatment times are adjusted so that a uniform uptake of 25–30% of PEG is achieved, based on the dry weight of the wood; temperatures range from 21 to 60°C. The wood may require weeks of soaking, depending on species, density, and thickness. After treatment, the wood is dried in a well-ventilated space, preferably heated.

The PEG is not "fixed" or cured in the wood, and it leaches out if the wood contacts water. For this reason, PEG-treated wood usually is finished with a penetrating oil or polyurethane varnish. PEG is used primarily to treat cross-sections of trees for plaques, clock faces, and tabletops. It also is used to stabilize bowls and other turnings, carvings, and rifle stocks, and to treat archaeological water-logged wood.

Resin impregnation. Resin impregnation can be used to add bulking agents to the wood cell wall. These resins have a low degree of polymerization, or the monomers are polymerized after impregnation. The best fiber-penetrating thermosetting resins are the highly water-soluble, phenol formaldehyde resin-forming systems with low initial molecular weights. Green or dry veneers usually are treated by soaking or pressure treatment with a water-soluble resin in a 30% aqueous solution. Following impregnation, the wood is dried slowly and then heated at about 150°C for 20 min to set the resin.

Similarly, laminates of treated wood are constructed by gluing the individual sheets together. This product, called Impreg, contains 25–35% resin and has a density about 20% higher than the original wood and an antishrink efficiency

(ASE) of about 65%. As with PEG, the water-soluble resin penetrates the cell wall and keeps the wood in a swollen state. Unlike PEG, the phenol formaldehyde resin is cured by heating to form a high-molecular-weight, water-insoluble resin in the cell wall. Thus, Impreg can be used in contact with water. Its compressive strength is higher than that of untreated wood, but it has much less toughness. It shows considerable resistance to decay and termite and marine-borer attack.

Treating of wood with phenolic resins imparts high acid resistance, greatly increases the temperature to which wood can be heated without appreciable change in properties, and increases electrical resistance. The largest industrial application of Impreg is in die molds for automobile body parts and other uses. Despite its high cost, the dimensional stability and plasticity of Impreg make it commercially viable.

Compreg. Compreg is a highly dense product obtained by applying pressure to dry, resin-treated veneers during heat cure. It offers most of the advantages of Impreg. Its mechanical properties are appreciably better than those of the original uncompressed wood because of a 2–3-fold increase in density (1,350 kg/m^3). The strength properties of Compreg are increased in proportion to the degree of compression. Only its toughness is lower than that of untreated wood, although greater than that of Impreg. Because of the plasticizing action of the resin-forming materials, treated wood can be compressed under considerably lower pressures than dry, untreated wood. For example, when subjected to a pressure of only 1.7 MPa (244 psi) at 149°C, treated spruce, cottonwood, and aspen wood, dried to a moisture content of about 6% but not cured, are compressed to about half the original thickness and a specific gravity of about 1.0.

In a 24-h water-soaking test, Compreg has an ASE value of 95%. The rate of water pickup is so slow that complete swelling equilibrium of a 13-mm specimen is not reached after 1 year at room temperature. The product is brown and can be buffed to a high polish. Compreg is produced commercially in small quantities and is used for knife handles, gears, certain musical instruments, and decorative articles.

Vinyl polymer composites. Composites of wood with vinyl polymers have been developed. Because the vinyl polymers are clear, colorless, thermoplastic materials, they do not significantly discolor the wood; thus its natural beauty is retained, whereas the phenolic resins darken the wood. However, the hygroscopic characteristics of the wood substance are unaffected because little, if any, resin penetrates the cell walls; the ASE is only about 10–15%. The high resin content (70–100%) greatly reduces the normally high void volume of wood, however. Thus, the elimination of this important pathway for vapor diffusion slows down the response of the wood substance to changes in relative humidity, and its moisture resistance is greatly improved compared with the original wood. The hardness of the wood–plastic composite also is greatly improved.

A variety of vinyl monomers, such as methyl methacrylate and styrene, may be used. Complete filling of the cell lumens and other voids (in the "full-cell process") is easily accomplished by first subjecting the wood to a partial vacuum (about 0.3 in. of Hg) and then covering it with the monomer and soaking it for 2–6 h, depending upon the species of wood and its dimensions. Some penetration of the monomer into the cell walls also may be obtained by using a diffusion process, such as a solvent-exchange method.

Polymerization of the vinyl monomer in the wood may be done with either radiation or free radical catalysts. The polymerization of the vinyl monomers in both processes depends upon the same mechanism, that is, initiation by free radicals. In the radiation process, the gamma rays passing through the monomer and the woody tissue create a large number of excited and ionized molecules, many of which break into fragments, namely organic free radicals (R^-). These act as the initiator for the polymerization of an unsaturated monomer. Alternatively, the free radicals may be formed by thermal decomposition of compounds involving a weak bond. Commercially, the catalyst 2,2′-azobisisobutyronitrile now is most widely used, as it forms free radicals at a lower temperature than benzoyl peroxide [31].

If the end use of the wood–polymer composite requires an abrasive (sanding) or cutting process that brings about high temperatures, the thermoplastic polymer will melt, causing machining difficulties. To prevent such melting, a cross-linking substance such as diethylene glycol dimethacrylate is added to the monomer before impregnation into the wood (about 5% of the volume of the monomer). Wood–plastic materials are used in parquet flooring, certain sporting equipment, musical instruments, and decorative articles.

Chemical modifications. Wood also can be modified by chemical reactions directly with the hydroxyl groups present in wood polymer. This treatment confers bulk to the cell wall with the help of a permanently bonded chemical. Many reactive chemicals have been used to modify wood. For best results, the chemicals should be capable of reacting with the wood hydroxyl groups under neutral or mildly alkaline conditions at temperatures below 120°C. The chemical system should be simple and must be capable of swelling the wood structure to facilitate penetration. The complete molecule must react quickly with wood components, yielding stable chemical bonds; the treated wood still must possess the desirable properties of untreated wood.

The treatments of wood [32] with anhydrides, epoxides, and isocyanates appear to be most attractive conferring ASE

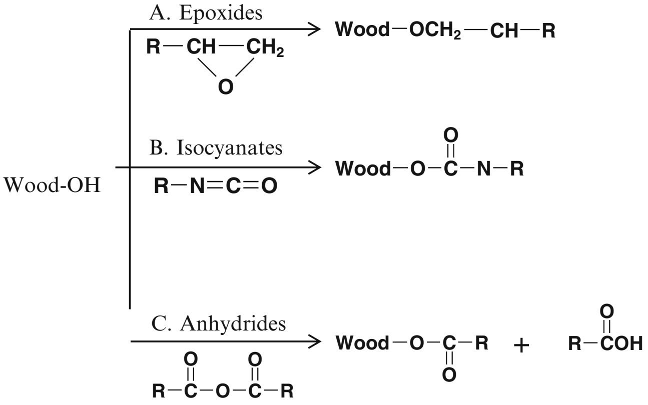


Fig. 28.34 Chemical modification of wood based on the reactions of hydroxyl groups

values of 60–75% at chemical weight gains of 20–30%. These chemical reactions as illustrated in Fig. 28.34 induced the formation of stable chemical linkages with the hydroxyl groups of wood polymers.

Of several anhydrides studied, acetic anhydride reacted the most readily. Reactions were carried out by refluxing the wood in a xylene/acetic anhydride solution or with acetic anhydride vapors alone at 120°C. With this system, for each mole of acetate bonded onto the wood a mole of acetic acid is generated as a by-product. Although this by-product generation is a disadvantage of the process, the chemical system does penetrate and react quickly with wood, without a catalyst. It is not so sensitive to moisture as are the epoxide and isocyanate systems.

A very facile dip acetylation procedure was developed for wood chips and fibers. The wood is dried at 105°C and then dipped in acetic anhydride for 1 min and drained. Then the sample is placed in a preheated (120°C) reactor for different lengths of time, depending on the degree of substitution of acetyl groups desired in the sample [33]. The swelling of aspen flakeboard made from such chemically modified flakes is dramatically reduced. For example, while unmodified flake board swelled by over 60% after immersion in water for 5 days, the modified material swelled by only about 8–22%, depending on the extent of chemical modification.

Acetylated wood [30] is now a commercial product being sold by Titanwood in Amhem, The Netherlands. It is under the name of Accoya and available in several species of wood in Europe, North America, and Asia.

Preservative Treatment of Wood

Wood, as a natural plant tissue, is subject to attack by fungi, insects, and marine borers. Some species of wood are more resistant to decay than others (e.g., the heartwood of cedars, cypress, and redwood) because of the presence of natural toxic substances among the extractable components. Most woods, however, are rapidly attacked when used in contact with soil or water, or when exposed to high relative humidities without adequate air circulation. Wood for such service conditions requires chemical treatment with toxic chemicals, collectively termed wood preservatives. The service life of wood may be increased 5–15-fold, depending upon the conditions of preservative treatment and the nature of the service.

The preservative treatment of wood is the second largest chemical wood-processing industry; pulp and paper manufacture is the most important. The more important types of wood products treated are shown in Fig. 28.35.

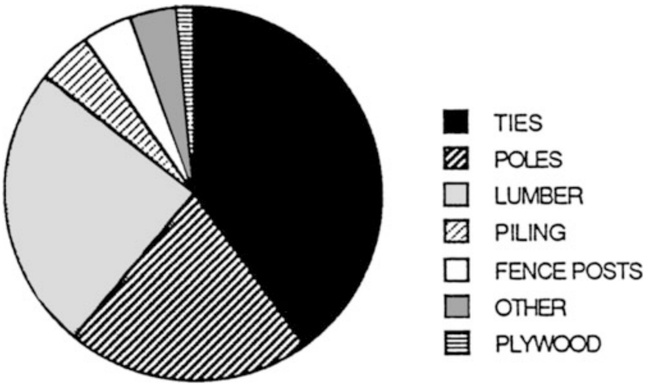


Fig. 28.35 Relative distribution of products treated with preservative

Preservative Chemicals

Owing to their toxic nature, all the commercial wood preservatives presently used in the United States are effective in preventing attack by microorganisms. However, because of concern that these chemicals will have a deleterious effect on the environment, alternative methods based on nontoxic procedures also are being investigated. These nonconventional approaches are based on chemical modification, and are described briefly at the end of this section. Wood preservatives are often classified as the oil-borne and water-borne preservatives [34].

Oil-borne type. Major preservatives of this type are creosote and pentachlorophenol. Creosote from coal tar is the most widely used wood preservative for several reasons: (1) it is highly toxic to wood-destroying organisms; (2) it has a high degree of permanence due to its relative insolubility in water and its low volatility; (3) it is easily applied, with deep penetration easily obtained; and (4) it is relatively cheap and widely available. For general outdoor service in structural timbers, poles, posts, piling, and mine props, and for marine uses, coal-tar creosote is the best and most important preservative. Because of its odor, dark color, and the fact that creosote-treated wood usually cannot be painted, creosote is unsuitable for finished lumber and for interior use.

Coal-tar creosote is a mixture of aromatic hydrocarbons containing appreciable amounts of tar acids and bases (up to about 5% of each), and has a boiling range between 200 and 355°C. The important hydrocarbons present include fluorene, anthracene, phenanthrene, and some naphthalene. The tar acids are mainly phenols, creosols, xylenols, and naphthols; the tar bases consist of pyridines, quinolines, and acridines.

Often coal tar or petroleum oil is mixed with the coal-tar creosote, in amounts up to 50%, as a means of lowering preservative costs. Because coal tar and petroleum have a low toxicity, their mixtures with creosote are less toxic than is creosote alone.

A number of phenols, especially chlorinated phenols and certain metal–organic compounds, such as copper naphthenate and phenylmercury oleate, are effective preservatives. Pentachlorophenol and copper naphthenate are most commonly used, and are carried into the wood in 1–5% solutions in petroleum oil. Pentachlorophenol is colorless, and can be applied in clear volatile mineral oils to mill-work and window sash requiring a clean, nonswelling, and paintable treatment.

Water-borne type. Copper is the primary biocide used in many wood preservative formulations. Chromated copper arsenate (CCA) and ammoniacal copper zinc arsenate (ACZA) had been the two major waterborne preservatives used in the United States, especially when cleanliness and paintability of the treated wood are required. Wood treated with CCA (commonly called green treated) dominated the treated-wood market from the late 1970s until 2004. Due to concerns over the exposure and disposal issues of CCA, now, CCA-treated products are generally not available at retail lumber yards. It can no longer be used for treatment of lumber to be used in residential decks or playground equipment. Substitutes for CCA include a copper boronazole compound and ammoniacal copper quat (ACQ). Those preservatives contain amine- or ammonia-complexed alkaline copper (II) and an organic co-biocide to control copper-tolerant fungi.

Preservation Process

Two methods are generally used for applying preservatives to wood: nonpressure and pressure processes. A variety of nonpressure processes have been used ranging from brushing, soaking to a thermal or vacuum process. The pressure process being more effective is commonly used in industrial applications. It consists of a full-cell process (Bethell) and an empty-cell process (Rueping and Lowry). They give deeper penetrations and more positive results than any of the nonpressure methods. The wood, on steel cars, is run into a long horizontal cylinder, which is closed and filled with preservative. Pressure is applied, forcing the preservative into the wood.

Nonpressure process. Brush and spray treatments usually give only limited protection because the penetration or depth of capillary absorption is slight. Dip treatments give slightly better protection. Organic chemicals dissolved in clear petroleum solvents often are applied to window sash and similar products by a dip treatment of 1–3 min.

The cold soaking of seasoned wood in low-viscosity preservative oil for several hours or days and the steeping of green or seasoned wood in waterborne preservatives for several days are methods sometimes employed for posts, lumber, and timbers on a limited basis. The diffusion process employs the water-borne preservatives that will diffuse out of the treating solution into the water in green or wet wood.

The most effective of the nonpressure processes is the thermal method of applying coal-tar creosote or other oil-soluble preservatives, such as pentachlorophenol solution. The wood is heated in the preservative liquid in an open tank for several hours, after which it is quickly submerged in cold preservative in which it is allowed to remain for several hours. This is accomplished either by transferring the wood at the proper time from the hot tank to the cold tank or by draining the hot preservative and quickly refilling the tank with cooler preservative. During the hot treatment, the air in the wood expands, and some is expelled. Heating also lowers the viscosity of the preservative so that there is better penetration. When the cooling takes place, the remaining air in the wood contracts, creating a partial vacuum that draws the preservative into the wood. For coal-tar creosote, the hot bath is at 210–235°F, and the cold bath at about 100°F. This temperature is required to keep the preservative fluid.

The hot- and cold-bath process is widely used for treating poles and, to a lesser extent, for fence posts, lumber, and timbers. This process gives the most effective results of the common nonpressure processes, which most nearly approach those obtained by the pressure processes.

The vacuum processes involve putting the wood under a vacuum to draw out part of the air. The wood may be subjected to a vacuum alone or to steaming and a vacuum before being submerged in a cold preservative. These methods are used to a limited extent in the treatment of lumber, timber, and millwork.

Pressure process. In commercial practice, wood is most often treated by immersing it in a preservative in a high-pressure apparatus and applying pressure to drive the preservative into the wood. Pressure processes differ in details, but the general principle is the same.

The empty-cell process seeks deep penetration with a relatively low net retention of preservative by forcing out the bulk liquid in the wood cells, leaving the internal capillary structure coated with preservative. The preservative liquid is forced under pressure into the wood, containing either its normal air content (Lowry process) or an excess of air, by first subjecting the wood to air pressure before applying the preservative under pressure (Rueping process). In the former case, the preservative is put in the cylinder containing the wood at atmospheric pressure, and, in the latter case, under air pressure of 25–100 psi. After the wood has been subjected to the hot preservative (about 190–200°F) under pressure (100–200 psi in the Lowry process and 150–200 psi in the Rueping process) and the pressure has been released, the back pressure of the compressed air in the wood forces out the free liquid from the wood. As much as 20–60% of the injected preservative may be recovered, yet good depth of penetration of the preservative is achieved.

The full-cell process seeks to fill the cell lumens of the wood with the preservative liquid, giving retention of a maximum quantity of preservative. The wood in the cylinder first is subjected to a vacuum of not less than 22 in. Hg for 15–60 min, to remove as much air as possible from the wood. The cylinder then is filled with hot treating liquid without admitting air. The maximum temperature for creosote and its solutions is 210°F, and for water-borne preservatives it is 120–150°F, depending upon the preservative. Then, the liquid is placed under a pressure of 125–200 psi, and the temperature and pressure are maintained for the desired length of time, usually several hours. After the liquid is drawn from the cylinder, a short vacuum is applied to free the charge of surface-dripping preservative.

Preservative Retention

Retention of preservative is generally specified in terms of the weight of preservative per cubic foot of wood, based on the total weight of preservative retained and the total volume of wood treated in a charge. Penetration and retention vary widely between different species of wood, as well as with woods of the same species grown in different areas. In most species, heartwood is much more difficult to penetrate than sapwood. Also, within each annual growth ring there is variability in penetration, the earlywood generally being more easily treated than the latewood.

The American Wood-Preservers' Association Standards specify methods of analysis to determine penetration and retention. They also specify minimum retention amounts for different preservatives according to the commodity, the species, the pretreatment of the wood, such as kiln drying, and the end use of the commodity. Heavier retention is required for products in contact with the ground (poles, timbers, etc.) or with marine waters (piles, timbers, etc.). Unprotected wood in contact with the ground is subject to severe attack by fungi and insects, and, in contact with seawater, it is quickly destroyed by marine borers. For wood products to be used in contact with the ground or marine waters, creosote is the major preservative employed because it can be readily impregnated to give high retention and good protection, and it is not leached out by water.

Increasing environmental regulations are causing greater restrictions on the use of traditional wood preservatives and alternate, nonpolluting methods are urgently needed. The use of CCA to treat wood for residential structures was phased out at the end of 2003. Arsenic, a carcinogen, imparts resistance to rot and termite damage, but there are concerns that it leaches from wood and exposes children. This forces the industry to develop more environmentally friendly preservatives [35].

Nonconventional Wood Preservation

The chemical modification of wood as a possible preservative treatment for wood is based on the theory that enzymes (like cellulases) must contact directly with the substrate (cellulose), and this substrate must have a specific configuration. If the chemical nature of cellulosic substrate is changed, this highly selective enzymatic reaction cannot take place. A major modification process is to change the hydrophilic nature of wood, which can be achieved by a thermal or chemical process as indicated earlier. In most cases, water a necessity for decay organisms is excluded from biological sites.

The chemicals used for modification need not be toxic to the organism because their action renders the substrate unrecognizable as a food source to support microbial growth. For wood preservation, this means that it is possible to treat wood in such a manner that attack by wood-destroying fungi will be prevented, and the material will be safe for humans to handle. For applications of wood in which human contact is essential and nontoxic preservatives may well be specified or required in the future. An added benefit to most of the chemically modified woods is that the resulting bulking action gives the treated wood very good dimensional stability.

A key chemical modification of wood involves the reaction of their hydroxyl groups with chemical reagents to form a covalent bond between the two (Fig. 28.34) as discussed earlier related to the dimensional stability treatments. A review on this subject shows that treatments with acetic anhydride, dimethyl sulfate, acrylonitrile, butylene oxide, phenyl isocyanate, and propiolactone all give good rot resistance at a 17–25 weight percent gain (WPG) [36]. The decay resistance of acetylated wood is proportional to the WPG, and the degree of dimensional stability also is proportional to the WPG; so the exclusion of cell wall or biological water may be a very important factor in the decay resistance mechanism [37].

In preliminary tests, alkylene oxide-treated southern pine was found to be resistant to termite attack and attack from the marine borers *Teredo* (shipworm) and *Limnoria*.

In conclusion, the modification of wood by heat treatments or by chemical reactions using acetic anhydride results in treated wood that is dimensionally stable and resistant to attack by fungi without the use of toxic chemicals [30]. The high chemical treatment level required for effectiveness may result in a rather expensive treatment. However, for those products where both the Rot resistance and dimensional stability are important, the present state of the technology is close to a viable industrial process as reflected in the commercial production of acetylated wood [30].

Fire-Retardant Treatment of Wood

The Romans first treated wood for fire retardancy in the first century AD. They used solutions of alum and vinegar to protect their boats against fire. In 1820, Gay-Lussac advocated the use of ammonium phosphates and borax for treating cellulosic material. Many of the promising inorganic chemicals used to day were identified between 1800 and 1870. Since then, the development of fire retardants for wood has accelerated. Commercially treated wood became available after the U.S. Navy (1895) specified its use in ship construction, and New York City (1899) required its use in buildings over 12 stories tall. Production reached over 65 million board feet in 1943, but by 1964 only 32 million board feet was treated annually [38].

However, starting in 1979 changes were made in some building codes that allowed structural changes if fire-retardant (FR) treated wood was used in place of untreated plywood for roof sheathing. The structural changes resulted in sufficient savings such that FR-treated wood was utilized more extensively. Also the use of FR-treated plywood was mandated at other specific sites such as prisons.

Fire-retardant treatments for wood can be grouped into two general classes: (1) those impregnated into the wood or incorporated into wood composite products and (2) those applied as paint or surface coatings [39]. Chemical impregnation has the greater use, primarily for new materials, whereas coatings have been limited primarily to materials in existing constructions. There are advantages and disadvantages to each class. Coatings are applied easily, and they are economical. Chemical impregnation usually involves full-cell pressure treatment and can be costly.

A coating is subject to abrasion or wear that can destroy the effectiveness of the fire retardant. Chemical impregnations deposit the fire retardant within the wood, so that if the surface is abraded, chemicals are still present. On-site application of surface coatings requires strict control of the amount applied to ensure correct loading levels for a particular flame spread rating.

Both coating and impregnation systems are based on the same chemical compounds although the formulations for each vary. Among the most commonly used chemicals for impregnation treatments are diammonium phosphate, ammonium sulfate, borax, boric acid, zinc chloride, and most recently the leach-resistant amino-resin systems.

The fire-retardant chemicals have different characteristics with respect to fire resistance. Ammonium phosphate, for example, is effective in checking both flaming and glowing; borax is good in checking flaming but is not a satisfactory glow retardant. Boric acid is excellent in stopping glow but not so effective in retarding flaming. Because of these different characteristics, mixtures of chemicals usually are employed in treating formulations.

Fire-Retardant Formulations

Many chemicals have been evaluated for their effectiveness as FRs. Today most FRs for wood are based on phosphorus, nitrogen, boron, aluminum trihydrate, and a few other compounds. Phosphorus and nitrogen frequently are used together because they behave synergistically; amino-resins are an example of such a combination.

Most FR formulations are not resistant to leaching by water. Therefore, there have been increased efforts to develop leach-resistant chemicals that can be impregnated into wood products for use in exterior or high-humidity applications. Some of the proposed leach-resistant systems include chemical combinations that form insoluble complexes, amino-resin systems, and monomers that polymerize in the wood. A common amino-resin system for exterior use is dicyandiamidephosphoric acid formaldehyde. Guanylurea phosphate-boric acid also is a commonly used organic phosphate salt for modern commercial FR wood.

The American Wood-Preservers' Association Standards specified the four types of FR formulations given in Table 28.10. Many newer formulations have been developed by commercial enterprises and are proprietary.

The impregnation methods are similar to those employed for the preservative treatment of wood by water-borne salts using pressure processes. The maximum temperature of the solution must not exceed 140°F for formulation Types A, B, and D, and must not be above 160°F for Type C. After

Table 28.10 Formulations of fire retardants

	Type			
Chemicals (%)	*A*	*B*	*C*	*D*
Ammonium phosphate	–	–	10	–
Ammonium sulfate	–	10	60	35
Boric acid	–	10	–	25
Sodium dichromate	20	40	–	5
Sodium tetraborate	–	–	10	
Zinc chloride	80	40	–	35

treatment, the wood must be dried to a moisture content of 19% or less. For most uses, the wood is kiln-dried to a moisture content of below 10%.

For effective fire protection larger amounts of chemicals must be deposited in the wood than the amounts necessary when water-borne chemicals are used for decay prevention. Whereas for the water-soluble toxic salts retentions from 0.22 to 1.00 lb/ft^3 of wood are specified according to commodity standards in order to give good protection against decay and insects, as much as 5–6 lb of some fire retardants may be required for a high degree of effectiveness against fire. Usually, however smaller amounts will give a good degree of protection. For example, formulation Type B when impregnated in amounts of 1.5–3 lb/ft^3 of wood provides a combined protection against fire, decay, and insects. Because the amino-resin systems are leach-resistant, high application levels are not necessary in practice.

Fire retardancy of wood involves a complex series of simultaneous chemical reactions, the products of which take part in subsequent reactions. Most FRs used for wood increase the dehydration reactions that occur during thermal degradation so that more char and fewer combustible volatiles are produced. The mechanism by which this happens depends on the particular FR and the thermal–physical environment. The effectiveness of a FR treatment depends upon the performance rating of the treated material when tested in accordance with ASTM E84 (no greater flame spread than 25).

Conversion of Wood to Energy, Fuels, and Chemicals

Wood has been a source of energy and chemicals for hundreds of years and continues to be an important raw material for specific chemicals. The use of wood as a primary source of industrial chemicals decreased dramatically in the 1940s when oil became the preferred raw material. Current interest in wood-derived fuels and chemicals has been rekindled by an increasing demand of energy and the soaring prices of petroleum products. This current biorefinery development is sustainable and also would help to reduce our dependence on fossil fuels and the greenhouse gas emission to the environment. The term “silvi-chemicals” sometimes is used to refer to wood-derived chemicals analogous to petrochemicals.

One of the recent developments in wood-based biorefinery is on the preextraction of hemicelluloses from wood as this would greatly enhance the overall utilization of wood [4]. The conversion of wood to energy or chemicals in general involves a chemical, biological, thermal, or a combined process. The specific process used varies with the nature of intended product. In this section, the improved wood utilization by preextraction of hemicelluloses, direct combustion, saccharification, fermentation, and a variety of thermal decomposition process (carbonization, pyrolysis, gasification, and liquefaction) is discussed to reflect their recent developments [40, 41].

Improved Wood Utilization by Pre-extraction of Hemicelluloses

The paper industry as noted earlier represents the largest scale in the isolation of wood polysaccharides as papermaking fibers. In the production of pulp fibers for making stable white papers, the lignin component has to be totally removed by a series of pulping and bleaching operations. Although the dominant kraft pulping is an effective process in the sense of being able to pulp all kind of plant materials, it is not an efficient operation, as approximately one third of the polysaccharides are degraded and lost resulting from the alkaline degradation reactions (Fig. 28.23).

Naturally, many wood-based biorefinery projects have emphasized on the extraction and isolation of those hemicelluloses [4, 42–45], which otherwise would be degraded in a typical kraft cooking process. The extraction methods examined ranged from a hot-water (autohydrolysis) extraction to a mild alkali treatment. The extent of the hemicellulose extraction would vary with the intent to utilize those extracted chips for the production of papermaking fibers or solid wood products like wood pellets.

A technology based on the hot-water extraction of sugar maple in the 160–170°C range has been extensively studied with respect to the characteristics and potential application of the water extracts and the extracted chips [4, 42, 45]. For example, the sugar components of water extracts after hydrolysis were shown being very responsive to fermentation for the production of biofuels (ethanol and butanol) or biopolymers like polyhydroxy-alkanoates (PHA). The extracted chips can be used for the production of papermaking fibers or reconstituted wood products with improved properties. Interestingly, the water extraction step significantly enhanced the subsequent alkaline (Kraft) delignification reaction [45].

Direct Combustion

The concept of using wood as a source of energy through direct combustion dates back to the very beginning of human existence. As soon as early people learned to use fire, wood became the major source of energy. It is important to note that fuel is still the major use of wood on a worldwide basis, amounting to an estimated 1.2×10^9 m^3/year compared with 0.8×10^9 m^3/year for lumber and other solid wood products, 0.34×10^9 m^3/year for pulp, and 0.2×10^9 m^3/year for other uses. Thus, on a global basis, nearly 50% of the wood harvest is used directly for fuel. In many nonindustrialized countries, fuel amounts to 80–90% of total wood consumption, compared with an average of 7% for industrialized regions.

The average stored energy content of wood is 8,600 BTU/lb. Variations in heat values for wood as a function of species occur within a relatively narrow range of 8,000–10,000 BTU/lb. Bark values are slightly higher than wood values, about 10,000 BTU/lb.

The greater the oxygen content of a carbon compound or carbonaceous substance, the lower the heat of combustion. Lignin, with an elemental analysis of $C_{10}H_{11}O_2$ compared with $C_6H_{10}O_5$ for cellulose, has a lower degree of oxidation and a higher heat of combustion than cellulose. The fossil fuels that have very low or no oxygen content have much higher energy contents, for example 12,000–13,500 BTU/lb for coal, 18,000–19,900 BTU/lb for oil, and 18,550 BTU/lb for natural gas.

Wood and other biomass materials are hygroscopic and retain substantial amounts of moisture (15–60%). Hence, some of the thermal energy of the wood is lost in vaporizing this water to steam during combustion. In addition, wood has a lower density than coal, so that a greater volume of material must be gathered and transported to provide the same heating value as that of coal. Wood has a negligible sulfur content, less than 0.1% compared with much larger amounts in coal (avg. 2%) and petroleum. Sulfur from coal and petroleum creates serious environmental problems by causing acid rain. Also, ash from coal presents a disposal problem and creates unhealthy fly ash, whereas the ash content of wood is less than 0.5%. The only environmental hazard produced from burning wood as a fuel is the production of particulates.

Most wood and wood-derived materials (spent pulping liquors) that are used for energy are consumed by the forest products industry itself. The source of this fuel is almost entirely in the wood-processing and manufacturing operations, termed “manufacturing residuals” or “mill residues.” Forest residues from logging operations and the noncommercial trees in the forest also are utilized for fuels.

Wood wastes supply the forest products industry with roughly 40–50% of the energy requirements in the form of spent liquors and wood and bark residues. The spent pulping liquor contains 40–50% of the wood as dissolved lignin and a large variety of other organic compounds. The spent liquor is used almost totally, after evaporation to about 65% solids, to produce steam and to recover the inorganic chemicals used in the pulping operation. The Tomlinson furnace is used almost universally and is the most efficient system for combined energy and chemical recovery yet devised.

The mixed wood and bark residues burned directly are collectively termed “hog fuel.” Over the past decade, there has been a trend to channel more and more of the cleaner, drier, and larger-sized mill residues into raw materials for products. The coarse wood residues (slabs, trim, edgings, etc.) are chipped and sent to pulp mills. Also, considerable sawdust now is being pulped instead of being burned for fuel. Some of the residues, including shavings, also are used for particleboard. This means that hog fuel contains the less desirable, dirtier, and wetter forms of waste.

Combustion technology. The hog fuel boiler represents the conventional technology of using wood for energy to produce process steam. Modern wood-fuel boilers have traveling grates and use mechanical draft fans. Many are of gigantic size, handling as much as 500,000 lb of wood per hour. Spreader stokes distribute the fuel evenly over the large grate areas to ensure efficient combustion. The ash can be removed continuously with traveling grates. Modern instrumentation meters air flow, controls fuel-to-air ratios, and meters combustion efficiency and stack emissions. With these developments, the energy recovery and environmental acceptability of wood-burning systems have been improved in the forest products industry.

The deterioration of hog fuel quality referred to above makes it more difficult to obtain efficient combustion, high reliability, and low stack emissions simultaneously. To overcome this difficulty, the fuel can be processed before combustion to remove dirt and moisture and produce clean-burning, efficient fuel. This can be done by drying, screening, grinding, washing, and pelletizing or densification of the hog fuel in order to remove noncombustible dirt and moisture, reduce the size of large and slow-burning material, and agglomerate small and fine material.

Size uniformity is important for ease in transportation, storage, and fuel feeding. Dirt and moisture reduction reduces transportation costs, promotes efficient combustion, and minimizes air emissions. Another approach for using low-quality hog fuel is the development of improved combustors. Examples are: (1) the fluidized-bed combustors, which obtain excellent wood combustion at relatively high efficiencies with low-quality, nonuniform fuels; and (2) the pyrolytic burner, which has very low stack emissions with relatively high thermal efficiencies and wood combustion rates.

Present-day wood-fired boiler systems are complicated and cost considerably more than a comparable petroleum-fired installation. As much as 25% of the capital cost is in the fuel-handling equipment, and another 20% is in the air pollution control system. Because of the high capital costs and the lower thermal efficiencies of burning wood compared with oil (68% vs. 82%), the success of the wood-fired systems depends on the low cost of the wood fuel supplies.

The North American forest products industry has been successful in developing useful technologies for the production of energy from wood residues as process heat or steam, which are cost-effective. New technologies promise even greater economic benefits.

Upgrading of solid fuels. To overcome the drawbacks of using the hog fuels as noted earlier, the recent production of densified wood fuels like pellet has been increased quite substantially. The raw materials include shavings and sawdust from sawmills as well as tree branches or low-quality wood. These solid fuel products are uniform in size and quality, and also easy to handle and transport. Wood pellets have a cylindrical form being 6–8 mm in diameter with varying length and are increasingly used in residential heating and industrial operations for generating heat and power. The energy value for 1 t of softwood pellets was estimated to be about 5 MWh, which is equal to 0.5 m^3 of oil [40]. Thus, wood pellets appear to be attractive as a good substitute for fossil fuels in many applications.

Cogeneration technology. Cogeneration is the concurrent generation of electricity and the use of exhaust heat, usually in the form of process steam, for manufacturing operations. This is done by burning fuel (in this case, wood) to make high-pressure steam, 600–1,200 lb/in.^{2}, passing this steam through a back-pressure or extraction turbine to drive a generator, and then using the steam exhausted from the turbine at lower pressures, 50–300 lb/in.^{2}, for process heat. This technology gets full use of the energy contained in the fuel. Wood at 55% moisture will generate power at about 60% efficiency.

The forest products industry is a major user of cogeneration technology, as it requires large quantities of process steam as well as electricity. It produces about 50% of its electricity needs in this way. The pulp and paper industry alone is the largest producer of energy by cogeneration of any U.S. industry. Electricity self-sufficiency is likely to increase to 80–90% for forest industries in the future through cogeneration. However, there is a minimum plant size for economical power generation. Steam usage should be more than 70,000–120,000 lb/h, equivalent to 3–5 MW of back pressure, for a plant economically to employ cogeneration systems.

Saccharification

Saccharification represents a key step to convert wood polysaccharides to liquid fuels and other chemicals. It involves the hydrolysis of cellulose and hemicelluloses to a mixture of hexoses (glucose, galactose, and mannose) and pentoses (xylose and arabinose). These monomeric sugars can then be fermented to yield ethyl alcohol (ethanol or grain alcohol, C_2H_5OH) by proper microorganism in the same way that ethanol is produced from grains or fruits. Obviously the concept of ethanol production from wood is not a new one and the polysaccharide character of wood has been known for over 100 years. A major challenge on the use of wood for the ethanol production as compared to other substrates like sugarcane or corn has been on the difficulties in separating and hydrolyzing the crystalline cellulose component in wood.

The ethanol production from the nonwoody materials is rather straightforward. To cope with the high oil prices, Brazil took the dramatic step of shifting to a much greater use of fuel alcohol produced from the sugars in sugarcane. One wood hydrolysis plant was constructed, but it was uneconomical to operate and was shut down. Since 2002, they have used the gasoline–bioethanol blend (gasohol) for motor fuels. Brazil is now second to the U.S. in the production of bioethanol.

In the United States, the interest in producing alcohols from renewable resources was revitalized by the dramatic increase in the price of petroleum in the 1970s and the push to decrease oil imports by substituting gasohol. Now, all the gasoline at gas pumps contains bioethanol (10%), which is produced largely from corn and the commercial production of cellulosic ethanol may be expected to begin in few years. Bioethanols produced from starchy substrates or sugars in food crops are generally referred as the first generation or noncellulosic ethanol, whereas the second generation ethanol refers to the cellulosic ethanol being derived from the inedible lignocellulosic materials like wood or switch grass.

Both acids and enzymes can be used to hydrolyze the cellulose to glucose, but so far only acids have been utilized commercially for wood hydrolysis to sugars notably in Russia. They continued the expansion of wood hydrolysis facilities, and about 40 such plants are presently in operation. All the Russian plants are based on dilute sulfuric acid in percolation towers. Cellulose being a partially crystalline material is closely associated with other cell wall components (hemicelluloses and lignin) in woody materials and relatively inaccessible to the chemical or enzymatic actions. A key technical challenge is in developing an economically viable saccharification process for cellulose. The characteristics of the acid and the enzymatic hydrolysis process are discussed below.

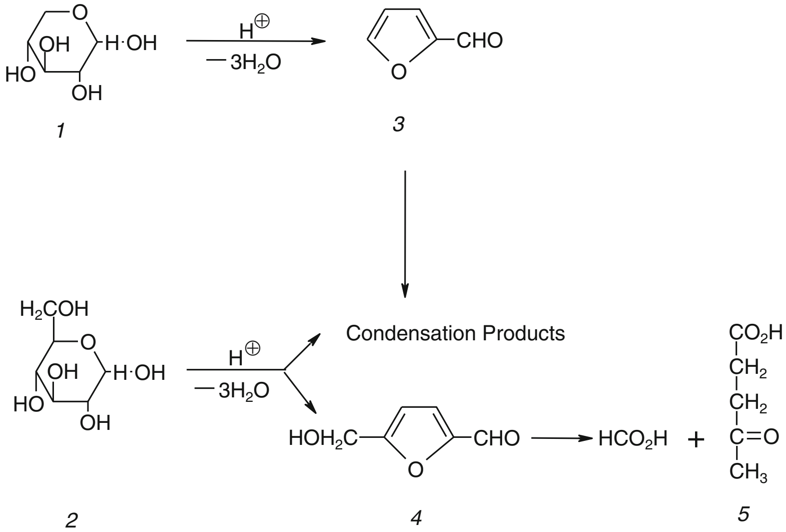


Fig. 28.36 Acid-catalyzed degradation of sugars: *1* (xylose), *2* (glucose), *3* (2-furaldehyde), *4* (5-hydroxymethyl-2-furaldehyde), and *5* (levulinic acid)

Acid hydrolysis. The acid hydrolysis of wood polysaccharides is a classic reaction and has been extensively studied related to the sugar formation [46]. Cellulose, the major component of wood, gives over 90% yield of pure glucose under laboratory hydrolysis conditions. Hemicelluloses hydrolyze much more easily and rapidly than the cellulose and give a mixture of sugars (xylose, arabinose, mannose, galactose, and glucose). Temperatures and acid concentrations that hydrolyze the cellulose to glucose in a matter of a few hours readily convert much of the hemicelluloses into simple sugars in minutes. Under industrial conditions of hydrolysis, the sugars formed undergo decomposition, with the pentoses decomposing more rapidly than the hexoses (Fig. 28.36). Thus, the conditions of hydrolysis cause variations in the ratio and yields of the various sugars due to (1) their different rates of formation by hydrolysis and (2) their different rates of decomposition. The sugar degradation products especially the furan-types compounds are the major inhibitors for the subsequent fermentation process.

The polysaccharide components of wood (holocellulose) may be hydrolyzed by two general methods: by strong acids, such as 70–72% sulfuric acid or 40–45% hydrochloric acid; or by dilute acids, such as 0.5–2.0% sulfuric acid. The hydrolysis by strong acids proceeds somewhat like a homogeneous system following a first-order reaction, as the cellulose component swells quickly and become soluble as illustrated in Reaction Scheme A (Fig. 28.37).

In dilute-acid hydrolysis, the reactions are heterogeneous and more complex, because no swelling and solubilizing of the cellulose occurs. Cleavage of the insoluble cellulose (B_3) gives the low-molecular-weight oligosaccharides (intermediate products), which are then rapidly hydrolyzed to simple sugars as indicated in Scheme B. The Reaction B_3 proceeds slowly and is the limiting factor for the whole process.

A variety of industrial processes on the hydrolysis of lignocellulosic materials for the ethanol production have been developed and evaluated over the years. They can be grouped into the concentrated acid and the dilute acid process as briefly discussed in the following.

The concentrated acid process. This process type is based on the concept that concentrated acid is capable to decrystallize the cellulose molecule and this facilitates the subsequent dilute acid hydrolysis. The yield of sugars is quite high being over 90%. Around the World War II period, the Bergius–Rheinau process was developed in Germany and was based on concentrated cold hydrochloric acids conducted at low temperatures. On the other hand, a concentrated sulfuric acid process was further developed by the U.S. Department of Agriculture (USDA) at the Northern Regional Research Laboratory in Peoria and the process is often termed the “Peoria Process.” A similar process was developed in Japan and they commercialized a concentrated sulfuric acid process in 1948.

For the concentrated acid process, the high yield of sugars is offset by higher capital costs for corrosion-resistant equipment and higher operating costs for acid recovery plus acid losses. In the 1980s, considerable efforts to improve the Peoria Process were conducted in U.S. especially at Tennessee Valley Authority (TVA) and at Purdue University. Among others, how to minimize the use of sulfuric acid and the development of a cost-effective recycling of the acid are key factors affecting the economic feasibility of the process.

Also, there is a continued interest in the concentrated hydrochloric process. In 2004, a pilot plant unit was installed in Domsjö, Sweden to evaluate the relative merit of a concentrated hydrochloric acid process (CHAP) [24, 40]. Interestingly, a recent patent indicates that the HCl can be

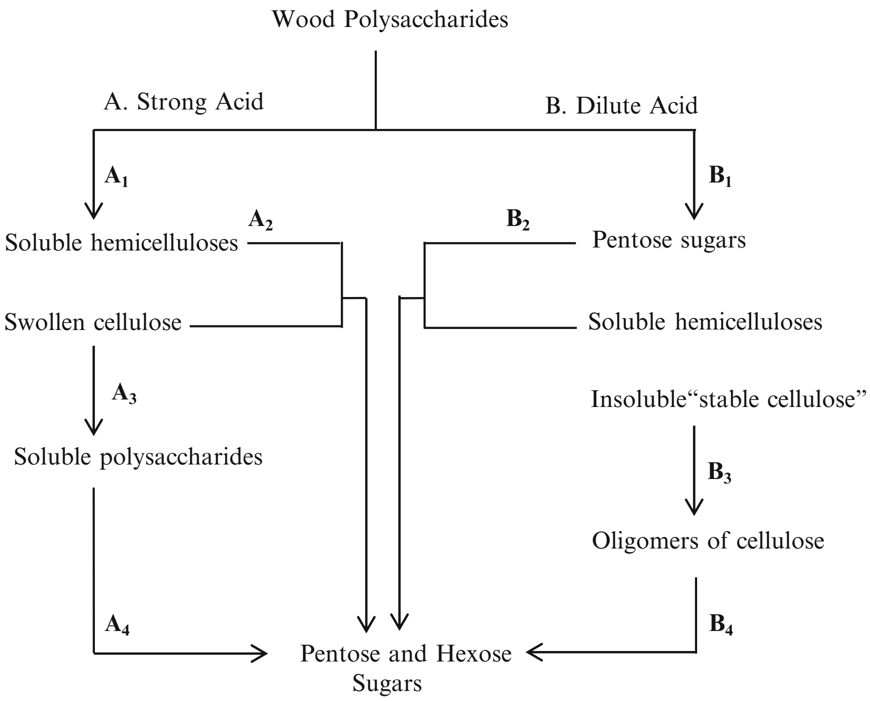


Fig. 28.37 Reaction schemes for the acid hydrolysis of wood polysaccharides

reconverted directly from an aqueous solution by using an immiscible extractant in a hydrocarbon diluent having a boiling above 210°C [47]. This improved process if confirmed would make the Bergius–Rheinau or CHAP process more attractive economically.

The dilute acid process. The Scholler–Tornesch process is a classic example for wood hydrolysis with dilute sulfuric acid conducted in the 170–180°C (338–356°F) range. It was developed originally in Germany and later modified by the research conducted at the U.S. Forest Products Laboratory (FPL) in Madison, Wisconsin. This improved process is also known as the Madison process.

In the Madison process, dilute sulfuric acid with an average concentration of 0.5% is pressed through wood in the form of sawdust and shavings. Regular flow of the acid and of the resulting sugar solution is one of the two principal requirements; the other is a lignin residue that can be discharged from the pressure vessels without manual labor. Both depend upon careful charging of the wood, which should not contain too many very fine particles, and upon maintaining a pressure differential of not more than 5–6 lb/in.^{2} between the top and the bottom of the digester.

The digesters or percolators are pressure hydrolyzing vessels, commonly employing a pressure of 150 lb/in.^{2}, and having a capacity of 2,000 ft^3 each. In the original Scholler plants in Germany, the digesters were lined with lead and acid-proof brick; in the Madison process a lining of "Everdur" metal was found to give sufficient protection.

The wood, about 15 t, is pressed down with steam and then is heated by direct steam, after which the acid is introduced. The practice in Germany was to bring the dilute acid into the digester in several batches, with rest periods of about 30 min, heating the wood to temperatures of 130°C (266°F) at first, then to 180°C (356°F), while keeping the temperature of the entering acid 10–20°C lower. A total of about 14 h was required to exhaust the wood, yielding about 50 lb of carbohydrates for 100 lb of dry wood substance. In the Madison process, continuous flow of the acid, and, correspondingly, of the sugar solution, is provided— in other words, continuous percolation. The cycle is thereby reduced to 6 h, and the yields are increased somewhat. The lignin is blown out of the digester by opening the specially constructed bottom valves while the vessel is still under pressure.

The sugar solutions usually contain about 5% sulfuric acid. The solutions, still under pressure (150 lb/in.^{2}), are flash-evaporated to 35 lb/in.^{2}, neutralized with lime at that pressure, and filtered. Calcium sulfate is much less soluble at the elevated temperature corresponding to the pressure than it is at 100°C. This is a fortunate circumstance, for it must be removed to an extent sufficient to avoid difficulties caused by the formation of incrustations in the subsequent alcohol distillation. The filtered solution is cooled by further flash evaporation and heat exchanged with water to the fermentation temperature. Sugar yields from coniferous woods (softwoods) are about 50% at an average concentration of 5%. When fermented, the average ethyl alcohol yield per ton of drywoods is 50–60 gal and sometimes higher.

The U.S. FPL in cooperation with the TVA had studied a two-stage dilute acid hydrolysis process based in part on studies of Cederquist in Sweden during the 1950s. The first stage (prehydrolysis) selectively removes the hemicellulosic sugars with dilute sulfuric acid at about 170°C prior to hydrolysis of the lignocellulosic residue to glucose in the higher temperature (230°C) second stage. The two-stage dilute acid process has a number of important advantages: (1) the carbohydrates are fractionated into hemicellulosic sugars and glucose, so that the separate utilization of each fraction is facilitated; (2) glucose is isolated in moderately good yield (~50%); (3) the glucose solution from the second stage is moderately concentrated (~10–12%); and (4) the consumption of acid and steam is relatively low.

A detailed model for the dilute acid hydrolysis of cellulose was developed at FPL in connection with studies of the two-stage dilute sulfuric acid hydrolysis process [48]. The model incorporates the effect of the neutralizing capacity of the substrate, the presence of readily hydrolyzable cellulose, and the reversion reactions of glucose in acid solution. Although general in nature, the model was developed specifically for application to the dilute, sulfuric acid hydrolysis of prehydrolyzed wood. A computer program simulating the new model can be used to predict yields of free glucose, reducing sugars, reversion material, remaining cellulose, and glucose loss due to dehydration as a function of acid concentration, temperature, and reaction time.

Also, Sweden has actively engaged in an international development project on the ethanol production from lignocellulosic feedstock [24, 40]. In 2004, a pilot plant was installed in Domsjö to evaluate the relative merits for three different saccharification processes. These include a concentrated hydrochloride acid process (CHAP) conducted at 20–40°C, a dilute acid process termed CASH (Canada-America-Sweden-Hydrolysis), and an enzymatic hydrolysis process. The CASH process includes the use of SO_2 and dilute sulfuric acid in two separate stages and appears to be the most promising among the three processes [24]. The pilot plant has the capacity to process 2 t of woody biomass per day to produce about 400 L of ethanol. Interestingly, both the CASH and the FBL-TVA systems share the same process features, as the sugars derived from the easily hydrolyzed hemicelluloses were removed prior to the subsequent hydrolysis of the cellulose component. This two-stage process naturally would minimize the further degradation of the pentoses derived from the hemicelluloses. The overall benefits of these two-stage processes as noted earlier would increase the total sugar production, reduce the formation of inhibitors like furfural (Fig. 28.36), and thus improve the overall fermentation process.

Major concerns of the acid-catalyzed saccharification as noted earlier are the equipment corrosion and the recovery of the acid used. In case of the dilute acid process, the generation of furfural and 5-hydroxyfurfural (Fig. 28.36) is also a major concern, as they are inhibitory to the subsequent fermentation process. Thus, a detoxification step is needed prior to the fermentation process.

Enzymatic hydrolysis. Saccharification of wood polysaccharides to sugars can also be accomplished by enzymatic techniques. In general, a mixture of cellulases containing endoglucanases, exoglucanases, and β-glucosidases was used to hydrolyze the cellulose [49, 50], whereas hemicellulases are used for the hemicellulose hydrolysis [51]. Important cellulolytic enzymes are produced from the fungus *Trichoderma reesi*. Enzymatic hydrolysis can be conducted as such or combined together with fermentation as a single process. The latter simultaneous saccharification and fermentation (SSF) process would greatly facilitate the ethanol production process [51, 52]. However, the most common microorganism used for the fermentation of hydrolyzates from the lignocellulosic substrates is *Saccharomyces cerevisiae*, which ferments the hexoses but not the pentoses.

The enzymatic hydrolysis as compared to acid hydrolysis has the advantages of being conducted under milder conditions with temperatures up to 60°C in the pH range of 2.5–5.5. It also produces less undesirable products like furfural which inhibits the fermentation reaction. However, the requirements of extensive pretreatment of wood substrates [52–54] and the high cost of enzymes [55] are the major concerns for its industrial applications [56].

Enzymatic methods show the biggest promise for conversion of waste paper from municipal waste into glucose for ethanol production. Because paper is composed primarily of wood cellulose fibers, the enzyme inhibition due to lack of accessibility with whole wood is partially alleviated. As mentioned previously, waste paper can represent up to 50% of typical municipal waste. Currently the separated paper from the waste is just burned for its fuel value.

Pretreatments. The current emphasis in the United States has been on the enzymatic hydrolysis process and how this hydrolysis process could be facilitated by a thermal or chemical pretreatment [52–54]. The main purpose of pretreatments is to enhance the accessibility of cellulose, which can be achieved by a controlled removal of the hemicelluloses or lignin components. Treatments with chemicals include dilute acid, alkali, ozone, organosolv, sulfite (SPORL), or green liquor systems. Thermal treatments include the steam explosion, ammonia fiber explosion, and carbon dioxide explosion process. The specific application of these methods seems to vary with nature of substrates and potential applications of the treated materials. For example, the softwood after a bisulfite-based SPORL treatment was very responsive to enzymatic hydrolysis and the subsequent fermentation process. An ethanol

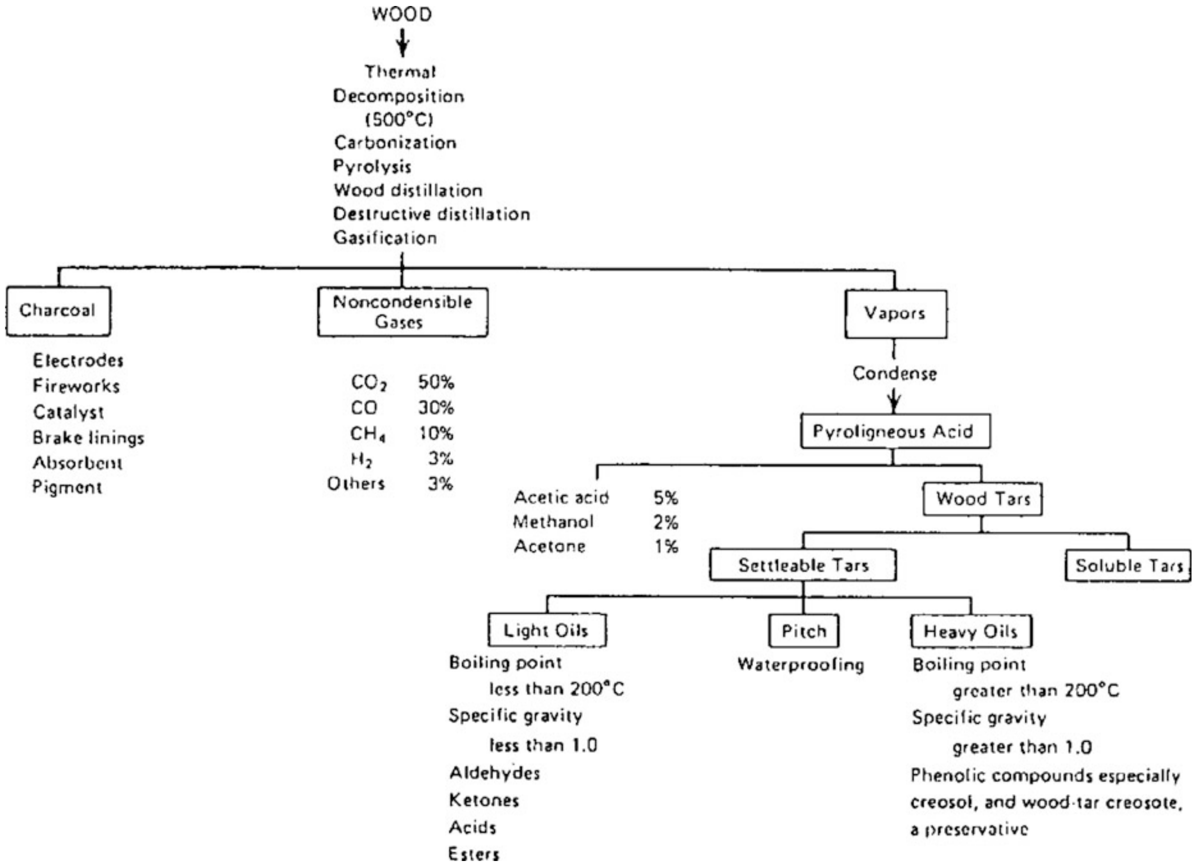


Fig. 28.38 Product obtained from thermal decomposition of wood (from Young and Giese [2], Copyright © 2004, with permission of John Wiley & Sons Publisher)

yield of 276 L/t of wood was reported for softwood and this corresponds to a 72% theoretical yield [53].

Fermentation

All the sugars derived from the hydrolysis of wood polysaccharides can now be fermented to yield ethanol and a variety of other products by a proper selection of the microorganisms.

Alcohols. As noted earlier, the yeast fermentation of hexoses with *S. cerevisiae* is essentially the same way that ethanol is produced from grains or fruits. Whereas the fermentation of five carbon sugars (xylose and arabinose) to ethanol can be achieved by many native and engineered microorganisms including *Escherichia coli* and *S. cerevisiae* [52, 53]. In addition, strains, which are able to ferment both the five- and six-carbons sugars, have been developed. This would greatly facilitate the production of ethanol from a sugar stream derived from the woody biomass [57].

Other products. In addition to ethyl alcohol, the fermentation process can yield many other useful products ranging from simple compounds like xylitol and sorbitol to biodegradable polymers like polyhydoxyalkanoate (PHA) and polylactic acid [58].

Thermal Decomposition

When wood is heated in the absence of air or with only limited amounts of air, thermal degradation takes place. This begins at about 100°C and increases with rising temperature. At about 270°C exothermic reactions set in, causing a rise in temperature (usually held at 400–500°C) bringing about complete carbonization. The resulting products are charcoal, tars (condensable liquids), and noncondensable gases. All these fractions may be used as fuels or chemicals as illustrated in Fig. 28.38.

For example, the wood gas has a fuel value of 300 BTU/ft^3. Methanol obtained from the destructive distillation of wood represented the only commercial source until the 1920s. The yield of methanol from wood by this method is low, only about 1–2% or 20 L/metric ton (6 gal/t) for hardwoods and about one-half that for softwoods. With the introduction of natural gas technology, the industry gradually switched to a synthetic methanol formed from a synthesis gas (syngas) produced from reformed natural gas.

During World War II in Germany, automobiles were fueled by the gases produced from thermal decomposition of wood; and research is ongoing today on the more efficient gasification of wood. Destructive distillation has been used throughout most recorded history to obtain turpentine from pinewood as discussed later in the chapter.

The range of materials or chemicals derivable from the thermal decomposition of wood (Fig. 28.38) varies with the species of woods used, the type of equipment and system (pyrolysis or gasification) employed, and the process variables (temperature, heating rate, and gas residence time). For example, the oil and tar yields can be varied between 1 and 40% and the char between 40 and 10% or less. Low temperatures favor liquids and char, low heating rates favor gas and char, and short gas residence favors liquids. Conversely, high temperatures favor gas, high heating rates favor liquids, and long gas residence times favor gas. Thus, the various product fractions can be preferentially manipulated by proper combinations of these variables. The thermal conversions of wood by the pyrolysis, gasification, and liquefaction processes are discussed further in the following.

Carbonization

The production of charcoal and tars by destructive distillation is the oldest of all chemical wood-processing methods. Charcoal probably was first discovered when the black material left over from a previous fire burned with intense heat and little smoke and flame. For centuries, charcoal has been used in braziers for heating purposes. Destructive distillation of hardwoods has been carried out with charcoal the product sought and volatiles as by-products; with softwoods (pines), volatiles were the principal products (naval stores), with charcoal considered a by-product.

In the United States, charcoal production began in early colonial days. During this period, principal uses of charcoal were as a fuel in blast furnaces for the production of pig iron and as an ingredient of gunpowder. Charcoal needed by the iron industry had to have a high crushing strength, and thus was from dense hardwoods such as maple, birch, oak, and hickory. A softer charcoal was preferred for making gunpowder and was produced from willow and basswood. The first successful blast furnace was built in Saugus, Massachusetts, in 1645. Other furnaces were started in close proximity to iron ore deposits. These early furnaces were small, usually producing only 1–3 metric tons of pig iron per day. In the late 1880s, wood charcoal lost its metallurgical market to coke, which was better suited to demands of larger.

Basic techniques for producing charcoal have not changed over the years although the equipment has. Charcoal is produced when wood is burned under conditions in which the supply of oxygen is severely limited. "Carbonization" is a term that aptly describes the thermal decomposition of wood for this application. Decomposition of carbon compounds takes place as the temperature rises, leading to a solid residue that is richer in carbon than the original material. Wood has a carbon content of about 50%, whereas charcoal of a quality suitable for general market acceptance will be analyzed as follows: fixed carbon 74–81%, volatiles 18–23%, moisture 2–4%, and ash 1–4%. Charcoal with a volatile content over 24% will cause smoking and is undesirable for recreational uses.

Earthen "pit kilns" originally were used to produce charcoal. A circular mound-shaped pile of wood (15–45 cords) was built up with an open core 30–60 cm (1–2 ft) in diameter to serve as a flue. The entire surface of the pile, except for the top flue opening and several small openings around the bottom periphery, then was covered with dirt or sod sufficiently thick to exclude air. The mound was allowed to "coal" for 20–30 days to give the final product.

In the second half of the nineteenth century, brick or masonry "beehive" kilns came into widespread use. The capacity of many of these kilns was from 50 to 90 cords, and operation was essentially the same as for the pit kilns. Many other types of kilns have been used from time to time. Small portable sheet-metal kilns of 1–2 cords have been widely used, and have rectangular masonry block kilns of various designs. These were predominantly used by farmers and small woodlot owners.

Large-scale production of charcoal was done by distilling the wood in steel buggies in long horizontal ovens. The buggies rode in steel rails that carried the cars in line from predryers to the ovens and then to coolers. Charcoal was produced by this method in a matter of 24 h. The latest types of charcoal-producing equipment are designed for continuous operation and make use of residues instead of roundwood. An example is the Herreshoff multiple hearth furnace, in which several hearths or burning chambers are stacked on top of one another, the number depending on capacity. Production in this type of furnace is from 1 to 2.5 metric tons of charcoal per hour.

Pyrolysis

On the pyrolysis of wood in the absence of oxygen, the temperature and heating rate plays a key role in affecting the formation and nature of the products. Conventional pyrolysis is often termed as slow pyrolysis and has been used traditionally for the production of charcoal as described earlier. This essentially is a carbonization process and includes the destructive distillation of wood to produce the charcoal and tars.

A fast pyrolysis [40, 59] refers to a high temperature process (500–700°C) with a high heating rate (300°C/min) for a short time. The major product is the bio-oil. Interestingly, the pyrolysis of pine wood samples releases significant amounts of the wood extractives (turpentine and tall oil). These products constitute the so-called naval stores chemicals and are discussed in the later wood extractives section.

Also, the pyrolysis may be conducted with a very high heating rate and a short reaction time (only a few seconds). This intensive process is often termed as flash pyrolysis and requires very small size of particles (105–250 µm).

Bio-oils. The organic liquid fraction from the wood pyrolysis is often termed as bio-oil. It is a complex mixture of alcohols, aldehyde, ketones, ester, and phenolic compounds from the fragmentation of cellulose, hemicelluloses, lignin, and extractives. The fast pyrolysis is generally employed for bio-oil production. The yields of bio-oil from wood are in the range of 70–80%. Biomass containing a higher content of lignin tends to give a lower bio-oil yield but with a higher energy density. The heating value of bio-oils is about 40–45% of that for hydrocarbon oils.

Bio-oils like the petroleum-based products can be stored, pumped, and transported. They can also be burned in boilers, gas turbines, and slow and medium-speed diesel engines for the heat and power generations.

An increasing interest in the bio-oils applications is to use as substrates like coal for conversion to syngas, as precursors for fuels and chemicals. This biorefinery of bio-oils would be parallel to the conventional oil refining.

Gasification

Gasification [60, 61] is the thermal degradation of wood or other carbonaceous material in the presence of controlled amounts of oxidizing agents, such as air or pure oxygen. It is carried out at higher temperatures than those used in the pyrolysis process, up to around 1,000°C. Hence, the reaction rates are very fast, making equipment design critical. The thermal efficiency for conversion of wood to gases is 60–80%, which compares favorably with that of coal. Wood gasification offers several advantages over coal: (1) much lower oxygen requirements, (2) virtually no steam requirements, (3) lower costs for changing H_2/CO ratios, which are already higher in wood gas, and (4) no or very little desulfurization costs. Coal has an advantage in that larger plants can be built for coal gasification than is normally the case for wood, because of its procurement advantage.

Several gasifiers have been designed to handle municipal refuse, wood wastes, and other biomass materials. Basically, gasifiers fall into two types: (1) those that use air and (2) those using oxygen. An example of the air system is the Moore–Canada gasifier. The wood residues are carried on a moving bed through stages of drying, reduction, and char oxidation, and the ash is discharged in granular form. The maximum temperature in the reaction zone is about 1,222°C. The hydrogen content of the crude gas is increased from 8–10% up to 18–22% by adding steam to the air intake.

The Union Carbide Company Purox gasifier is an example of the oxygen system. This unit also uses a moving-bed reactor. Pure oxygen is the oxidizing agent used to convert the char into CO and CO_2. Molten ash leaves the bottom at about 1,670°C. If air is used, the crude gases contain about 46% nitrogen, which must be removed by cryogenic means; but if oxygen is used, it first must be separated from air, with oxygen and nitrogen separated by a cryogenic system (see Chap. 27). The crude gases also contain an oil and tar fraction, about 2% of the wood (dry basis). Gasifiers designed for wood operate at atmospheric pressure, in contrast to coal gasifiers that operate at pressures up to 400 psi.

Wood gasification products. The composition of the wood gas varies according to the technology used. If the limited oxygen required is supplied by air, a typical gas might contain H_2 (10–18%), CO (22–23%), CO_2 (6–9%), N_2 (45–50%), and hydrocarbons (3–5%), and have a heating value of about 1,700 kcal/m^3 (180 BTU/ft^3). If pure oxygen is supplied to the process, the gas might contain H_2 (24–26%), CO (40%), hydrocarbons (10%), and CO_2 (23–25%), and have a heating value of 2,900 kcal/m^3 (350 BTU/ft^3). The gas may be used directly as a fuel in a variety of ways; for example, supplying boiler energy in an industrial plant at the rate of 2.5×10^8 kcal/h from about 136 kg/day of dry wood.

Figure 28.39 illustrates some examples of products derived from the wood gasification and from the Fischer–Tropsch synthesis of the syngas. Under gasification conditions, the carbon from woody biomass is converted to gas (CO and CO_2). With a combined reactions with oxygen and steam, a mixture of CO and H_2 is formed by a well-known Lurgi process (Eq. 1). The gas mixture is called biosyngas to differentiate it from the syngas derived from coal. A variety of products can be derived from the reforming of syngas by the so-called Fischer–Tropsch (FT) process. The products range from alkane (Eq. 2), methane (Eq. 3), to methanol (Eq. 4), and dimethyl ether (DME) (Eq. 5). The latter DME derived from the dehydration of methanol can be used as fuel in diesel engine.

The technology of gasification [60, 61] is under active development in both equipment and process design to reduce capital costs, and in the chemistry of the process to improve

Fig. 28.39 Examples of products from wood gasification and the Fischer–Tropsch synthesis of syngas [40]

Lurgi process

$$C + H_2O\ (\text{Steam}) \xrightarrow{\text{Heat}} CO + H_2 \tag{1}$$

Fischer-Tropsch synthesis

$$CO + H_2 \xrightarrow[\text{Heat, pressure}]{\text{Catalyst}} \text{Alkane} + H_2O \tag{2}$$

$$CO + 3H_2 \xrightarrow[\text{Heat, pressure}]{\text{Ni}} CH_4 + H_2O \tag{3}$$

$$CO + 2H_2 \xrightarrow{\text{Catalyst}} CH_3OH \tag{4}$$

$$2CH_3OH \xrightarrow[\text{Heat, pressure}]{\text{Catalyst}} CH_3OCH_3 \tag{5}$$

yields and obtain more favorable gas ratios. The latter includes the use of catalysts to enhance the process and to promote the formation of specific products, such as methane or ethylene or increased thermal efficiency or for feed-stocks for chemical synthesis.

Among the various chemicals or products that could be derived the wood gasification, the potential production of methanol and FT diesel has attracted much attention notably related to the paper industry. The gasification of the black cooking liquors after the kraft pulping may provide a means to recover the cooking chemicals (see the Kraft pulping section) and the production of useful chemicals. Recently, the ThermoChem Recovery International, Inc. assisted by the US Department of Energy has installed a pilot plant operation in Durham, North Carolina to produce diesel fuels and chemicals from gasification of woody biomass. This technology should be applicable to the kraft pulping liquor solids as well.

Methanol production. With the introduction of natural gas technology, the industry gradually switched to a synthetic methanol formed from a synthesis gas (syngas) produced from reformed natural gas. Two volumes of H_2 and one volume of CO are reacted in a catalytic converter at pressures of 1,500–4,000 psi to produce methanol. Presently, 99% of the methanol produced in the United States is derived from natural gas or petroleum.

The yield of methanol with present technology is about 390 L/t of oven dry wood. Conversion of wood to methanol is less efficient than producing methanol from natural gas or coal—about 38% compared with about 60%, based on the heat value of methanol as a percent of the total energy input into the plant. The raw materials cost is the most significant operating cost. The production cost for methanol from wood is estimated to be almost twice that produced from natural gas. Improving the efficiency of the gasifier reactor to increase the quantities of CO and H_2 produced from wood would enhance the process.

Recently in Austria [62], a model study was conducted to examine the optimal location of wood gasification for the methanol production with heat. The model covers the competition in all levels of a biofuel production chain like the supply of biomass, biofuel, and heat as well as the demand for bio- and fossil fuels. The results indicated that Austria could be self-sufficient in the production of methanol for blending with gasoline up to M20 (containing 20% methanol). The methanol production for the M5, M10, and M20 blends would require about 2, 4, and 8% of the arable lands, respectively. Austria had approximately 1.4 million ha of arable lands in 2005.

Liquefaction

The thermochemical liquefaction of wood, based on the nature of the product may be discussed in two major types. One is emphasized on the production of liquid fuel (bio-oil) [40, 63]. The other is on the preparation of liquefied wood as chemical feedstock for making adhesive [64] or other modified liquefied wood products [65].

Bio-oil production. The thermochemical liquefaction for the biofuel production generally requires a high-pressure and high-temperature treatment of wood chips in the presence of hydrogen gas or syngas. For example, an oil of a heating value of about 35,200 BTU/kg can be obtained by reaction of wood waste at 750°F under a pressure of 5,000 psi in the presence of syngas and a catalyst for 1 h. The feasibility of the process has been tested in a pilot plant in Albany, Oregon, based on laboratory work conducted at the U.S. Bureau of Mines. So far, it has been determined that a barrel of oil equivalent to No. 6 bunker fuel can be produced from about 405 kg of wood chips [63].

Alkali salts like sodium carbonate can be used as catalyst for the liquefaction process. Since the heavy oil obtained is a viscous tarry lump, some organic solvents (like propanol or butanol) are often added to the reaction system for easy handling. The yield of bio-oil is about 30% for the noncatalytic liquefaction and higher (60%) for the catalytic process [66]. Since the liquefaction process requires a need for more expensive reactors and fuel-feeding system, it is less attractive than the pyrolytic process for the bio-oil production [40].

Liquefied wood. The wood liquefaction can be conducted in a reagent–solvent system and an acid catalyst at temperatures below 180°C, which is significantly lower than that required for the bio-oil production (300–350°C). Phenol is a commonly used reagent solvent while sulfuric acid or oxalic acid can act as catalyst. The resulting liquefied wood can be used for the preparation of phenolic resin by reacting with formaldehyde.

Also, if the ethylene glycol and glycerol are used as a reagent–solvent system, the resulting liquefied wood contains a large number of hydroxyl groups, which can serve as reactive sites for chemical modifications. A variety of liquefied wood polymers have been synthesized including the polyester, polyurethane, and epox resin. These products were reported to be more biodegradable than those derived from the petrochemicals. Since liquefied wood represents a concept for total utilization of the woody biomass, it is an area of continued interest with many potential applications.

Chemical Conversion

Chemical reactions provide a fundamental basis for many of the wood conversion processes especially those related to the pulping and bleaching operations as discussed earlier. In the following, the chemical conversion of wood polymers to chemicals and fuels is further illustrated by the potential production of furan-type products from wood polysaccharides and lignophenols from wood lignin.

Furan-type products. The acid-catalyzed hydrolysis of wood polysaccharides is known to produce the furfural *3* and 5-hydroxymethyl furfural (HMF) *4* (Fig. 28.36). These two compounds are known to be inhibitors to the fermentation process. The furfural *3* is derived from the acidic degradation of hemicelluloses (mainly xylan) while the HMF *4* is from the degradation of hexoses (mainly from cellulose).

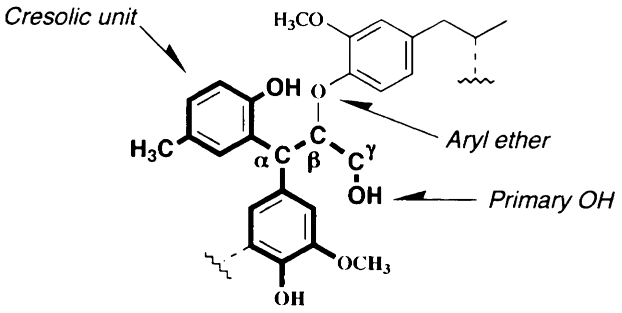


Fig. 28.40 The key structural feature of lignocresol [70]

Improved methods for the isolation of these two furan compounds have been developed by modifying the acidic treatment conditions [67].

Furfural *3* is a well-known industrial chemical and can be reduced to furfuryl alcohol (FA) and methyl furan (MF). FA has been used for polymer applications [68], while MF can be used as fuel additive. Similarly, HMF *4* can be converted to different types of useful acids, alcohols, or 2,5-dimethyl furan (DMF). The latter DMF, like MF, can be used as a fuel and its energy content is similar to that of gasoline.

Levulinic acid. This acid *5* (Fig. 28.36) is derived from the acid degradation of HMF and is a cellulose-derived product. It can be produced from a variety of woody biomass (wood wastes, agricultural residues, or municipal solid waste). Levulinic acid can be converted to a variety of industrial chemicals including tetrahydrofuran, 1,4 butanediol, diphenolic acid, and succinic acid [69].

Lignophenol. Lignophenols [70] represent a unique class of modified lignin polymers derived from lignocellulosic substrates by a phase-separation process. It is prepared by first mixing woodmeal samples with *p*-cresol in a small beaker and stirred for 10 min. After the addition of sulfuric acid (72%), the mixture was stirred vigorously at ambient temperatures for 1 h. The mixture is then separated into a cresol layer containing mainly the wood lignin and an aqueous layer of the carbohydrates.

The separated cresol layer was dropped into diethyl ether with vigorous stirring. The crude lignocresols precipitates were further purified by dissolving in acetone. The acetone-soluble material was then precipitated into the ether to give the purified lignocresols.

The key structure of lignophenol is in the formation of α-condensed 1,1-diarylpropane units (Fig. 28.40). The lignocresol is a highly phenolic polymer resulting from the grafted cresolic units and the cleavages of acid-labile aryl ether units. The total PhOH content of lignocresol per C_9 unit was about 1.3 for softwood and higher from hardwood (1.5) as compared to a value of 0.3 for typical milled wood lignin (MWL) samples.

Lignophenols have been demonstrated being quite reactive and suitable for many industrial applications including the areas of adhesives, cellulose composites, carbon molecular sieving membranes, and adsorbents for protein. Also, the dissolved carbohydrate components can be recovered for conversion to fuels or chemicals.

Thus, the phase-separation process would represent an important new approach to achieve an effective utilization of all the major biomass polymers including both the lignin and polysaccharide components. Several pilot-scale plants have been installed in Japan for further process evaluation.

Chemicals from Extractives

The most common chemicals derived wood extractives [71] are the turpentine and tall oil. They can be recovered by several processes including from the kraft pulping operation, by extraction of wood with suitable solvents, or from the exudates (oleoresin or gum) of living pine trees. The nature of these chemicals along with the tannin and other medicinal is outlined in the following.

Naval Stores

The U.S. naval stores industry began in the very early colonial days, when wooden vessels used tar and pitch from the crude gum or oleoresin collected from the wounds of living pine trees. The demand for tar and pitch from crude gum is now of minor importance. The industry is centered in the southeastern United States and is confined to the longleaf and slash pine areas. There was also a small, but locally important, naval-stores-producing area in the Landes region of southwestern France, based on the maritime pine.

There are three routes by which naval stores are produced. The oldest method is the tapping of living trees to cause a flow of oleoresin. The second method is the removal of naval stores by solvent extraction of wood. The latter extraction process now has replaced steam distillation as a means of recovering turpentine. The third route is from kraft (sulphate) pulping of pine, during which turpentine and tall oil are recovered as by-products of kraft pulp manufacture. This subject was briefly mentioned earlier in the pulping section. A fourth process, no longer used in the United States, is recovery of turpentine and pine oils by the destructive distillation of pine wood [72].

Worldwide, about 60% of naval stores are produced by tapping living trees. In the United States, the most important route is from the kraft pulping operation, whereas gum naval stores account for only 4% of the U.S. production.

Naval stores comprise two major types of products (turpentines and tall oil). Turpentines are a mixture of volatile terpenes while tall oil is consisted of three major types of components (resin acids, fatty acids, and unsaponifiables). Thus, naval stores also include important fatty acids, as well as steroids and other products. These naval stores products have a wide range of uses from ordinary household commodities to complex industrial applications [73].

Gum naval stores. The crude gum or oleoresin is obtained from healthy trees by exposing the sapwood. The lower part of the tree is faced; that is, a section of bark is removed, giving a flat wood surface for the gutters, which are inserted into a slanting cut made by a special ax. The gutters conduct the gum to a container that can hold 1–2 quarts of gum. At the top of the exposed face, a new V-shaped strip of bark is removed about every 2 weeks. The operations of inserting gutters, hanging cups, and cutting the first bark preferably are done in December or January, as early facing stimulates early season gum flow. The gum continues to flow until November, with the height of the season being from March to September.

The collected gum is distilled from a copper still; turpentine and water pass over, and the rosin is left in the still. The remaining molten rosin, plus impurities, is passed through a series of strainers and cotton batting to remove dirt particles. Then the liquid rosin is run into tank cars, drums, or multiwall paper bags for shipment.

Increases in yield of naval stores are brought about by chemical treatment of the exposed wood, especially with paraquat herbicides (dipyridyl compounds). This treatment stimulates extensive oleoresin formation and diffusion into the wood, extending to the pith of the tree and several feet above the treatment level. As much as 40% oleoresin content in the wood has been produced. Such treatment could double naval stores production, for both gum and sulfate processes. It also has the potential of providing a new type of wood naval stores by solvent extraction prior to kraft pulping or a combination of both methods.

Wood naval stores. “Wood” naval stores are produced by solvent extraction of resin-rich wood from old southern pine stumps and roots. The depletion of these stumps from the large trees of virgin forests, combined with high labor costs, has brought about a major decline in the production of wood naval stores by this process. In modern practice, all the resin products are removed from the shredded wood by solvent extraction; the solvent retained by the extracted wood chips is recovered by steaming. Extraction is carried out with naphtha (b.p. 90–115°C fraction). Multiple extractions are carried out in a series of vertical extractors in a countercurrent manner, whereby fresh solvent is used for the final extraction of a charge. The solution from the extractors is vacuum-distilled and the solvent recovered. The remaining

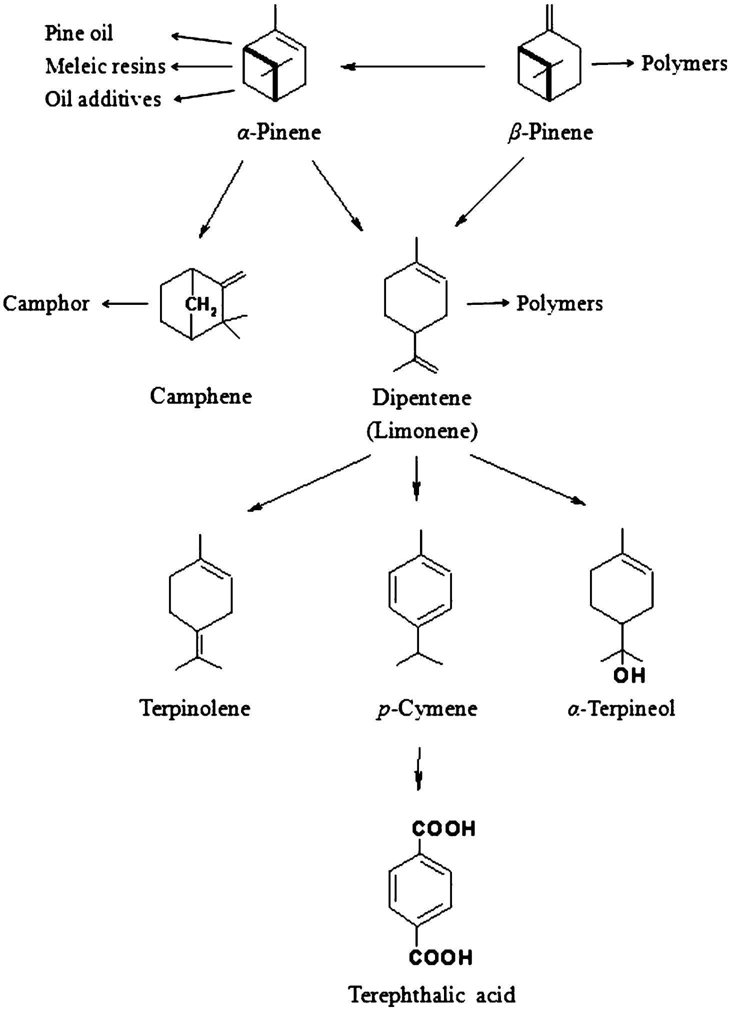


Fig. 28.41 Some reactions of α-pinene and β-pinene (adapted from Goldblatt LA. Yearbook of Agriculture, USDA, 1950–1951)

terpene oils are fractionally redistilled under vacuum and recovered as turpentine, dipentene, and pine oil. The non-volatile rosin is of dark color and is upgraded by clarification methods, such as selective absorption of its solution (bed-filtering).

Sulfate naval stores. Sulfate turpentine is obtained as a by-product during the kraft pulping of pine woods. Vapors periodically released from the top of the digesters are condensed, and the oily turpentine layer is separated and purified by fractional distillation and treatment with chemicals to remove traces of sulfur compounds. The spent black liquor from the kraft pulping of pines contains the less volatile product of the wood resin in the form of sodium salts or soaps. The liquor first is concentrated in multiple-effect evaporators, and then the concentrate is sent to settling tanks. The soaps rise to the surface, are skimmed off, and then are acidified with sulfurous or sulfuric acid. The crude tall oil rises to the top and is mechanically separated.

Turpentines. In the United States, about 80% of the annual production of turpentine is from the kraft pulping of southern pines (approximately 125 million liters). Turpentine is a volatile oil consisting primarily of terpene hydrocarbons, having two basic isoprene units (2-methylbutadiene, C_5H_8). These monoterpenes can have many different arrangements as illustrated in Fig. 28.41. Only six monoterpenes are present in appreciable amounts in commercial turpentines and their boiling points are in the 150–170°C range. These include α-pinene (156°C), β-pinene (164°C), camphene (159°C), Δ^3-carene (170°C), dipentene (limonene) (176°C), and terpinolene (188°C).

The composition of sulfate turpentine is very similar to that of gum turpentine obtained from the oleoresin of the tree (Table 28.11). The major components are α-pinene (60–70%) and β-pinene (25–35%). Sulfate turpentine from western North America woods contains appreciable amounts of Δ^3-carene which is used as a solvent. In contrast, wood turpentines contain some camphene (4–8%) with little of β-pinene.

Table 28.11 Composition of turpentines from different sources

	Types of turpentine		
Component%	Gum	Sulfate	Wood
α-Pinene	60–65	60–70	75–80
β-Pinene	25–35	20–25	0–2
Camphene	–	–	4–8
Other terpenes	5–8	6–12	15–20

Turpentine is used directly as a solvent, thinner, or additive for paints, varnishes, enamels, waxes, polishes, disinfectants, soaps, pharmaceuticals, wood stains, sealing wax, inks, and crayons, and as a general solvent. The chemistry of its monoterpenes offers many possibilities for conversion to other substances, as illustrated in Fig. 28.41. There is increasing use of turpentine to produce fine chemicals for flavors and fragrances. An important use of turpentine is in conversion by mineral acids to synthetic pine oil. It also is a raw material for making terpin hydrate, resins, camphene, insecticides, and other useful commodities. These uses are included in the following summary of its applications.

- Solvents for paints, etc. (11%)
- Synthetic pine oil (48%), used for mineral flotation, textile processing, solvents, odorants, bactericides, and conversion to their chemicals such as terpin hydrate, etc.
- Polyterpene resins (600–1,500 mol. wt.) (16%), used for adhesives, pressure-sensitive sizes (dry cleaning, paper, chewing gum)
- Camphene → toxophene insecticides
- Flavor and fragrance essential oils (9%)

Dipentene. This compound is also known as limonene derived from the higher boiling fractions of wood turpentine. Dipentene is used in paints, varnishes, and as a penetrating and softening agent in rubber reclamation.

Pine oil. Pine oil obtained from wood naval stores has similar uses to those of the synthetic pine oil made from turpentine, given earlier.

Tall oil. In the United States, tall oil like turpentines is largely derived from the kraft pulping operation. For each metric ton of pulp produced, northern pines yield about 50 kg of tall oil and higher (125 kg) from southern pines. The U.S. capacity for fractional distillation of tall oil is nearly one million metric tons per year.

Crude tall oil from southern pines contains resin acids (40–60%) and fatty acids (40–55%), and unsaponifiables or neutral substances (5–10%). These individual components are obtained by fractional distillation under vacuum. One metric ton of crude tall oil yields rosin (350 kg), fatty acids (300 kg), and the head and pitch fractions (300 kg).

Rosin. Rosin is a brittle solid that softens at 80°C. Chemically it is composed of about 90% resin acids and 10% of neutral matter. Rosin is graded and sold on the basis of color, the color grades ranging from pale yellow to dark red (almost black). The color is due almost entirely to iron contamination and oxidation products. Fresh oleoresin, as it exudes from the tree, will yield a rosin that is nearly colorless. Color bodies are removed by selective solvents and selective absorption from a 10 to 15% gasoline solution passed through beds of diatomaceous earth. About 70% of the world's rosin is produced in the United States.

Resin acids are tricyclic mono-carboxylic acids. These diterpene derivatives are often discussed in two groups (abietic or pimaric type) with the abietic type being dominant in nature (Fig. 28.8). A typical range of composition for rosin is indicated below:

1. Abietic type
    - Abietic (42–43%),
    - Neoabietic (4–5%),
    - Palustric (10–12%)
    - Lepopimaric (<0.5%)
    - Dehydroabietic (30%)
    - Other resin acids (8–12)
2. Pimaric type
    - Pimaric (4–5%),
    - Isopimaric (3–5%)
    - Sandaracopimaric acids (1–2%)

Rosin is used mainly in some modified forms. Because all the resin acids contain a carboxyl group and double bonds, they are reactive and can be used to produce salts, soaps, esters, amines, amides, nitriles, and Diels–Alder adducts. Also, they can be isomerized, disproportionated, hydrogenated, dimerized, and polymerized to provide a variety of useful products. When destructively distilled, rosin produces a viscous liquid, termed rosin oil, used in lubricating greases.

The paper industry uses large amounts of the sodium salt of rosin as paper size, which accounts for the greatest single use of rosin. The synthetic rubber industry is the second most important user of rosin. In making styrene-butadiene rubber, disproportionated rosin soaps are used alone or in combination with fatty acid soaps as emulsifiers in the polymerization process. Disproportionation decreases the number of double bonds in the abietic type of rosin, making a more stable material.

The adhesives industry is the third most important market for rosin. Rosin, modified rosins, and rosin derivatives are used in several types of adhesives, including the pressure-sensitive, hot-melt, and elastomer-based latices, and solvent rubber cements. Protective coatings are the fourth major or derivative form. Varnishes and alkyds are the most common types of protective coatings using rosin. Rosin is combined

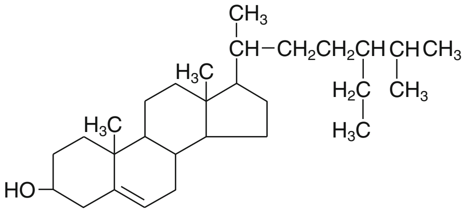


Fig. 28.42 β-Sitosterol

with a heat reactive phenol-formaldehyde resin to produce a widely used varnish. Printing inks also use substantial amounts of rosin. The above rosin uses are summarized as follows.

- Paper size (33%)
- Chemical intermediates and rubber (42%)
- Resins and ester gums (18%)
- Coatings (3.9%)
- Other uses (3.1%)

Fatty acids. The total fatty acids produced annually in the United States, amounting to more than 450,000 t and 35% of these come from tall oil. The solvent extraction of pine wood yields only 1% fatty acids and their esters. The yield is not increased, however, by paraquat (dipyridyl herbicides) treatment. Hence, the kraft-pulping industry will continue to be the major source of fatty acids from wood.

The fatty acids from tall oil have the following components: oleic (50%), linoleic (35%), conjugated oleic (8%), stearic (2%), palmitic (1%), and others (4%). From tall oil heads, a fraction is produced that is composed predominantly of saturated acids, containing 55% of palmitic acid. Oleic acid is distilled in grades that are 99.5% pure.

The approximate uses of fatty acids in various areas are:

- Intermediate chemicals (43%)
- Protective coatings (28%)
- Soaps and detergents (11%)
- Flotation (3%)
- Other uses (15%)

Additionally, the fatty acids of tall oil can be converted to bio-diesels by simple esterification reactions.

Unsaponifiables. The neutral or "unsaponifiable" materials present in tall oil contain mainly phytosterols (25–35%), higher alcohols (5–10%), hydrocarbons (33–60%), and many other minor components. For example, in the neutral fraction of southern pine tall oil soap, 80 compounds have been identified and the main component is sistosterol (25%), which accounts for 2–3% of the crude tall oil. Sistosterol is an alcohol (Fig. 28.42) and has potential applications. For example, this chemical has potential use in the synthesis of cortisones and other steroids and hormones by fermentation processes. Also, food products containing plant sterols are known to reduce the cholesterol levels.

Medicinals

The bark of the cascara tree of the northwestern region of the United States yields cascara, a laxative used in medicine. Several hundred tons of bark are harvested annually.

One of the most promising anticarcinogenic compounds discovered in the last 25 years was founding the bark and wood of the Pacific Yew tree from the Pacific Northwest of the United States. This compound, taxol, was found to be quite effective by the National Cancer Institute for the treatment of breast and ovarian cancers. The compound is now produced by partial synthetic routes and is available commercially under the generic name, paclitaxel.

Also, Gingo extracts or nuts have been used medicinally for thousands of years in China. The red gum tree of the southern United States exudes a yellowish balsamic liquid or gum from wounds, which is known as stora. It is produced by removing a section of bark and incising the wood in much the same manner as that used for the production of naval stores gum described earlier. Storax is used in medicinal and pharmaceutical preparations, such as adhesives and salves, and as an incense, in perfuming powders and soaps, and for flavoring tobacco.

Tannins

The tissues of wood, bark, and the leaves of trees contain a great variety of chemical substances of considerable scientific interest and some of practical value. Turpentine, pine oil, and rosin from the resins of pines are the most important commercial extractives from American woods. Tannin is a commercially important substance that can be extracted from the wood, bark, or leaves of certain trees and other plants. Tannins are complex dark-colored polyhydroxy phenolic compounds, related to gallic or ellagic acids (Fig. 28.8), and vary in composition from species to species. They have the important property of combining with the proteins of animal skins to produce leather. Recent interest in tannins has been on their pharmacological properties for potential applications as antiviral and antitumor agents.

For many years, most of the leather in the United States was tanned with domestic tannins from hemlock and oak bark and from chestnut wood. Today only a small amount of tannin comes from these and other domestic sources. The most important source of vegetable tannin today is the wood of the quebracho tree, which grows mainly in Paraguay and Argentina. The tannin content of this tree and a few other sources of vegetable tannin are given in Table 28.12.

For the tannin production, the wood or bark is reduced to chips and shreds by passing the material through hoggers or hammer mills. Then it is extracted with warm water in diffusion batteries. The dilute solutions are evaporated to

Table 28.12 Tannin content of some plant material

Plant material	Percent tannin
Domestic sources	
Eastern hemlock bark	9–13
Western hemlock bark	10–20
Tanbark oak	15–16
Chestnut oak	10–14
Black oak	8–12
Chestnut wood	4–15
Sumac leaves	25–32
Foreign sources	
Quebracho heartwood	20–30
Mangrove bark	15–42
Wattle (*acacia* bark)	15–50
Myrobalan nuts	30–40
Sicilian suma leaves	25–30

the desired concentration. Loss of solubility of the tannin can be counteracted by treatment of the concentrate with sodium sulfite.

Biotechnology

Biotechnology plays an important role in many aspects of the forest product industry as summarized by Henriksson and Teer [74]. These developments range from the breeding of genetically modified trees for improved wood quality and processing ability to the discovery of new enzyme systems for improved hydrolysis of wood polysaccharides, the fermentation technology, and the pulping and papermaking operations.

Genetic Modification of Trees

Advances in plant biotechnology have demonstrated that genetic engineering can improve the wood growth and quality of hardwood like poplar trees. For example, the growth of transgenic poplar as compared to the controlled tree showed an increased in height by 41% and in stem diameter by 36%. This improved wood yield represents an important trait of forest trees. Also, the lignin content of transgenic poplar was significantly reduced and the lignin was shown to have a higher proportion of the syringyl unit (S) and a lower guaiacyl (G) units. Wood lignin has a higher S/G ratio is generally more reactive and easier to delignify during the pulping operations.

As discussed earlier, softwood lignin unlike hardwood lignin does not contain the S units and is more difficult to degrade. Thus, intensive research has been conducted aiming to incorporate the S unit into the softwood trees. Initial results were quite promising. The positive outcome would facilitate the pulping and bleaching operations and preserve the fiber quality for improved products.

Enzymatic Deinking

Deinking is an essential process to reuse the recycled fibers for making white papers. The deinking process is generally based on detergents and separation of ink particles by floating. It can be substantially improved by the addition of various enzymes to the detergent mixture. Common enzymes used are fungal amylases (starch degrading enzymes), cellulases, and to some extent, lipases and hemicellulases. The enzymatic deinking is one of the most common enzymatic techniques used in the pulp and paper industry.

Biomechanical Pulping

The pulping of wood is at present based on either mechanical or chemical methods or combinations thereof, as previously described in this chapter. The interfiber lignin bond is broken down by the mechanical and chemical treatments to free the cellulose fibers for papermaking. A biomechanical pulping method that has reached the pilot plant stage and will probably be commercialized in the near future. This approach involves the use of white rot fungi to first breakdown the wood structure, and apparently modify the lignin with the result that the wood chips can then be mechanically disintegrated with much less energy input. After screening of hundreds of different species of white rot fungi at the U.S. FPL, it was found that *Ceriporiopsis subvermispora* produced the greatest reduction in energy consumption and also substantially improved the strength properties of the pulp in comparison with conventional refiner mechanical pulp. The approach is also suitable for the production of pulp from agro-based materials as well [19, 75].

Biobleaching

It is well known that the xylanase pretreatment of kraft pulps enhances the subsequent bleaching with gaseous chemicals such as chlorine dioxide or ozone. This is a well-established technology and it can reduce the chemical consumption by 10–20% while improving the pulp quality. In addition to xylanases, other hemicellulases, such as mannanases also have a similar effect.

Among the ligninases, the laccase-mediator-system (LMS) has received the most attention in delignification of pulps. This system requires the use of a mediator such as

1-hydroxybenzotriazole (HBT) or 2,2′-azino-bis-(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS) and is conducted under mild ambient conditions. The extent of LMS delignification of kraft pulps is very comparable to that of the O_2 bleaching. However, the cost of enzymes and mediator has prevented its commercial applications.

Biofuels and Bioproducts Conversion

As discussed earlier, the conversion of biomass materials can be achieved by enzymatic saccharification followed by fermentation to yield fuels (like bioethanol and butanol), chemicals (lactic acid and xylitol), biodegradable polymers (PHA), and many others [5, 58, 73]. Significant progress has been made in identifying the cellulase enzyme complex for hydrolyzing cellulose to glucose. The enzyme complex apparently contains both the decrystallizing and the hydrolysis enzymes that work together to convert cellulose to glucose. Isolation of the specific enzymes and genetic engineering could provide a more efficient complex available in the future.

Acknowledgement The author wishes to acknowledge the use of material from an earlier edition by Dr. R. A. Young.

References

1. Core HA, Cote WA, Day AC (1976) Wood structure and identification. Syracuse University Press, Syracuse
2. Young RA, Giese RL (eds) (1990) Introduction to forest science, 2nd edn. Wiley, New York; see also 3rd edn. Introduction to forest ecosystem science and management, 2004
3. Anon., Wood handbook: wood as an engineering material. Centennial edn. General Technical Report, FPL-GTR-190, USDA Forest Service, Forest Products Laboratory, Madison, WI, 2010
4. Amidon TE, Wood CD, Shupe AM, Wang Y, Graves M, Liu S (2008) Biorefinery: conversion of woody biomass to chemicals, energy and materials. J Biobased Mater Bioenergy 2(2):100
5. Hood EE, Nelson P, Powell R (eds) (2011) Plant biomass conversion. Wiley, New York
6. Timell TE (1967) Recent progress in the chemistry of wood hemicellulloses. Wood Sci Technol 1:45
7. Browing BL (1963) The supply and uses of wood. In: Browing BL (ed) The chemistry of wood. Wiley, New York
8. Kubler H (1990) Properties and utilization of wood. In: Young RA, Giese GL (eds) Introduction to forest science, 2nd edn. Wiley, New York, pp 451–477
9. Young RA, McGovern JN, Rowell RM (1990) Wood for fiber, energy and chemicals. In: Young RA, Giese RL (eds) Introduction to forest science, 2nd edn. Wiley, New York, pp 478–507
10. FAO Yearbook (2003) Forest products, FAO Forestry Series No. 36. FAO-UN, Rome
11. McGovern JN (1980) Pulping. In: Casey JP (ed) Pulp and paper chemistry and chemical technology, 3rd edn. Wiley, New York
12. Johansson B, Mjoberg J, Sandstrom O, Teder A (1984) Modified continuous kraft pulping – now a reality. Sven Pepperstidn 87 (10):30
13. Kleppe PJ (1970) Kraft pulping. Tappi 53(1):35
14. Masura V (1984) Kraft pulping of spruce and beech wood. Cell Chem Technol 18:421
15. Lai Y-Z (1991) Chemical degradation. In: Hon D, Hon N-S, Shiraishi N (eds) Wood and cellulosic chemistry. Marcel Dekker, New York, pp 443–512
16. Ojanen E, Tulppala J, Virkola N-E (1982) Neutral sulfite anthraquinone (NS-AQ) cooking of pine and birch wood chips. Pap Puu 8:453
17. Kleinert T, Tayenthal K (1931) Cellulose. AT Patent 124,738
18. Aziz S, Sarkanen KV (1989) Organosolv pulping - a review. Tappi 72(3):169
19. Young RA, Akhtar M (eds) (1998) Environmentally friendly technologies for the pulp and paper industry. Wiley, New York
20. Hergert HL (1998) Development in organosolv pulping—an overview. In: Young RA, Akhtar M (eds) Environmentally friendly technologies for the pulp and paper industry. Wiley, New York
21. Patt R, Kordsacha O, Schubert H-L (1998) The ASAM process. In: Young RA, Akhtar M (eds) Environmentally friendly technologies for the pulp and paper industry. Wiley, New York
22. Sundquist J, Poppius-Levlin K (1998) MILOX pulping and bleaching. In: Young RA, Akhtar M (eds) Environmentally friendly technologies for the pulp and paper industry. Wiley, New York
23. Smook GA (2002) Handbook for pulp and paper technologies, 3rd edn. Angus Wilde Publications Inc, Vancouver, BC
24. Gellerstedt G (2009) Cellulose products and chemicals from wood. In: Ek M, Gellerstedt G, Henriksson G (eds) Pulp and paper chemistry and technology, vol 1. Walter Gruyter, Berlin
25. Stark NM, Cal Z, Carll C (2010) Wood-based composite materials. In: Wood handbook: wood as an engineering material. Centennial edn. General Technical Report, FPL-GTR-190, USDA Forest Service, Forest Products Laboratory, Madison, W. Chapter 11
26. Smulski S (ed) (1997) Engineered wood products: a guide for specifiers, designers and users. PFS Research Foundation, Madison, WI
27. Haygreen J, Bowyer J (1996) Forest products and wood science, 3rd edn. Iowa State University Press, Ames, IA
28. Clemons C (2008) Raw materials for wood-polymer composites. In: Niska KO, Sain M (eds) Wood-polymer composites. CRC Press, Boca Raton
29. Caufield DF, Clemons C, Rowell RM (2010) Wood thermoplastic composites. In: Rowell RM, Caldeira F, Rowell JK (eds) Sustainable development in the forest products industry. Universidada Femando Pessoa, Porto, Portugal
30. Rowell RM, Ibach RE, Nilsson T (2010) Wood durability and stability without toxicity. In: Rowell RM, Caldeira F, Rowell JK (eds) Sustainable development in the forest products industry. Universidada Femando Pessoa, Porto, Portugal
31. Meyer JA (1984) Wood-polymer materials. In: Rowell RM (ed) The chemistry of solid wood. American Chemical Society, Washington, DC, pp 257–290
32. Rowell RM (1984) Penetration and reactivity of cell wall components. In: Rowell RM (ed) The chemistry of solid wood. American Chemical Society, Washington, DC, pp 175–210
33. Rowell RM, Tillman AM, Simonson R (1986) J Wood Chem Technol 6(3):427
34. Lebow ST (2010) Wood preservation. In: Wood handbook: wood as an engineering material, Centennial edn. General Technical Report, FPL-GTR-190, USDA Forest Service, Forest Products Laboratory, Madison, WI, Chapter 15
35. Schultz TP, Nicholas DD, McIntyre CR (2008) Recent patents and developments in biocidal wood protection systems for exterior applications. Recent Patents Material Sci 1:128
36. Rowell RM (2006) Chemical modification of wood: a short review. Wood Mater Sci Eng 1:29

37. Rowell RM (1977) Nonconventional wood preservation methods. In: Wood technology: chemical aspects, American Chemical Society, Symposium Series No. 43, Washington, DC, pp 47–56
38. LeVan SL (1984) The chemistry of fire retardancy. In: Rowell RM (ed) The chemistry of solid wood. American Chemical Society, Washington, DC, pp 531–574
39. White RH, Dietenberger MA (2010) Fire safety of wood construction. In: Wood handbook: wood as an engineering material, Centennial edn. General Technical Report, FPL-GTR-190, USDA Forest Service, Forest Products Laboratory, Madison, WI, chapter 18
40. Arshadi M, Sellstedt A (2008) Production of energy from biomass. In: Clark JH, Deswarte FEI (eds) Introduction to chemicals from biomass. Wiley, New York
41. Naik SN, Goud VV, Rout PK, Dalai AK (2010) Production of first and second generation biofuels: a comprehensive review. Renew Sustain Energy Rev 14:578
42. Liu T, Lin L, Sun Z, Ruofei Hu R, Liu S (2010) Bioethanol fermentation by recombinant E. coli FBR5 and its robust mutant FBHW using hot-water wood extract hydrolyzate as substrate. Biotechnol Adv 28:602
43. Li H, Saeed A, Jahan MS, Ni Y, van Heiningen ARP (2010) Hemicellulose removal from hardwood chips in the pre-hydrolysis step of the kraft-based dissolving pulp production proces. J Wood Chem Technol 30:48
44. Al-Dajani WW, Tschirner UW (2008) Pre-extraction of hemicelluloses and subsequent kraft pulping. Part I: Alkaline extraction. Tappi J 7(6):3
45. Hassan A, Bujanovic B, Amidon T (2010) Strength properties of kraft pulp produced from hot-water extracted woodchips within the biorefinery. J Biobased Mater Bioenergy 4(1):46
46. Wayman M (1986) Comparative effectiveness of various acids for hydrolysis of cellulose. In: Young RA, Rowell RM (eds) Cellulose: structure, modification and hydrolysis. Wiley-Interscience, New York, pp 265–279
47. Eyal A, Baniel A. Process for the recovery of HCl from a dilute dilution thereof. US Patent 20100093995
48. Connor AH, Wood BV, Hill CG, Harris VF (1986) Kinetic modeling of the saccharification of prehydrolyzed southern Red Oak. In: Young RA, Rowell RM (eds) Cellulose: structure, modification and hydrolysis. Wiley-Interscience, New York, pp 281–296
49. Gollapalli LE, Dale BE, Rivers DM (2002) Predicting digestibility of ammonia fiber explosion (AFEX)-treated rice straw, Appl Biochem Biotechnol 98:23
50. Vaega E, Szengyl Z, Reczey K (2002) Chemical pretreatments of corn stover for enhancing enzymatic digestibility. Appl Biochem Biotechnol 98:73
51. Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, Ladisch M (2005) Features of promising technologies for pretreatment of lignocellulosic biomass, Bioresour Technol 96:673
52. Gamage J, Lam H, Zhang Z (2010) Bioethanol production from lignocellulosic biomass, a review, J Biobased Mater Bioenergy 4:3
53. Tian S, Luo XL, Yang XS, Zhu JY (2010) Robust cellulosic ethanol production from SPOL-pretreated lodgepole pine using an adapted strain Saccharomyces cerevisae without detoxification. Bioresour Technol 101:8678
54. Jin Y, Jameel H, Chang H-M, Phillip R (2010) Green liquor pretreatment of mixed hardwood for ethanol production in repurposed kraft pulp mills, J Wood Chem Technol 30(1):86
55. Singhania RR, Sukumaran RK, Pandley A (2007) Improved cellulase production by Trichoderma reesei RUT C30 under SSF through process optimization, Appl Biochem Biotechnol 142:60
56. Brandberg T, Sanandaji N, Gustafsson L, Francen CJ (2005) Continuous fermentation of undetoxified dilute acid lignocellulose hydrolysate by Saccharomyces cerevisiae ATCC 96581 using cell recirculation. Biotechnol Progr 21:1093
57. Martin C, Galbe M, Wahlborn CF, Hahn-Hagerdal B, Jonsson JL (2002) Ethanol production from enzymatic hydrolysates of sugarcane bagasse using recombinant xylose-utilising Saccharomyces cerevisia. Enzyme Microb Technol 31:274
58. Koutinas AA, Du C, Wang RH, Webb C (2008) Production of chemicals from biomass. In: Clark JH, Deswarte FEI (eds) Introduction to chemicals from biomass. Wiley, New York , pp 77–101
59. Mohan D, Pittman CU Jr, Steele PH (2006) Pyrolysis of wood/biomass for biol-oil: a critical review. Energy Fuel 20:848
60. Sakaguchi M, Watkinson AP, Ellis N (2010) Steam gasification of bio-oils and bio-oil/char slurry in a fluidized bed reactor. Energy Fuel 24:5181
61. Chhiti Y, Salvador S, Commandre J-M, Broust F, Couhert C (2011) Wood bio-oil noncatalytic gasification: influence of temperature, dilution by an alcohol and ash content. Energy Fuel 25:345
62. Leduc S, Schwab D, Dotzauer E, Schmid E, Oberstiner M (2008) Optimal location of wood gasification plants for methanol production with heat recovery. Int J Energ Res 32:1080
63. Appell HR (1977) The production of oil from wood waste. In: Anderson L, Tilman DA (eds) Fuels from waste. Academic, New York, pp 121–140
64. Pan H, Shupe TF, Hse C-Y (2009) Characterization of novolac type liquefied wood/phenol/formaldehyde (LWPF) resin. Eur J Wood Prod 67:427
65. Kunaver M, Jasiukaityte E, Cuk N, Guthrie JT (2010) Liquefaction of wood, synthesis and characterization of liquefied wood polyester derivatives. J Appl Polym Sci 115:1265
66. Demirbas A (2004) Current technologies for the thermo-conversion of biomass into fuels and chemicals, Energ Source 26:715
67. Binder JB, Raines RT (2009) Simple chemical transformation of lignocellulosic biomass into furans for fuels and chemicals, J Am Chem Soc 131:1979
68. Gandini A, Belgacem N (2007) Furan chemistry at the service of functional macromolecular materials: the reversible Diels-Alder reaction. In: Argyropoulos DS (ed) ACS symposium series 954: materials, chemicals, and energy from forest biomass. American Chemical Society. Washington, DC, pp 280–295
69. Saladino R, Pagliaccia T, Argyropoulos DS, Crestini C (2007) Production of chemicals from cellulose and biomass-derived compounds: advances in the oxidative functionalization of levulinic acid. In: Argyropoulos DS (ed) ACS symposium series 954: materials, chemicals, and energy from forest biomass. American Chemical Society, Washington, DC
70. Funaoka M (1998) A new type of phenolic lignin-based network polymer with the structure-variable function composed of 1,1-diarylpropane units, Polym Int 47:277
71. Gardner JAF, Hills WE (1962) The influence of extractives on the pulping of wood. In Hillis WE (ed) Wood extractives and their significance to the pulp and paper industry. Academic, New York, pp 367–403
72. Zinkel DF (1975) Naval stores: silvichemicals from pine. J Appl Polym Sci 28:309
73. Hajny GJ (1981) Biological utilization of wood for production of chemicals and foodstuffs, USDA Forest Service Research Report FPL 385. Forest Products Laboratory, Madison, WI
74. Henriksson G, Teer T (2009) Biotechnology in the forest industry. In: Ek M, Gellerstedt G. Henriksson G (eds) Pulp and paper chemistry and technology, vol 1. Walter Gruyter, Berlin
75. Young RA (1997) Processing of agro-based resources into pulp and paper. In: Rowell R. Young RA (eds) Paper and composites from agro-based resources. CRC Press, New York

76. Suchsland O, Woodson GE (1986) Agric. Handbook 640: Fiberboard manufacturing practices in the United States. Washington, DC, USDA

Selected References: Books

American Wood Preservers' Association (1971) American wood preservers' standards. American Wood Preservers' Association, Washington, DC

Argyropoulos DS (ed) (2007) Materials, chemicals, and energy from forest biomass. ACS symposium series, vol 954. American Chemical Society, Washington, DC

Bierman CJ (1993) Essentials of pulping and papermaking. Academic, New York

Browning BL (1967) Methods of wood chemistry, vols I and II. Wiley, New York

Casey JP (1980) Pulp and paper, 2nd edn., vols I–III. Wiley, New York

Clark JH, Deswarte FEI (eds) (2008) Introduction to chemicals from biomass. Wiley, New York

Ek M, Gellerstedt G, Henriksson G (eds) (2009) Pulp and paper chemistry and technology, vol 1 (Wood chemistry and wood biotechnology), vol 2 (Pulping chemistry and technology), vol 3 (Paper chemistry and technology), and vol 4 (Paper products physics and technology). Walter Gruyter, Berlin

Fengel D, Wegener G (1984) Wood: chemistry, ultrastructure, reactions. Walter de Gruyter, New York

Hood EE, Nelson P, Powell R (eds) (2008) Plant biomass conversion. Wiley, New York

Glasser WG, Sarkanen S (eds) (1989) Lignin: properties and materials. American Chemical Society Symposium Series, vol 397. American Chemical Society, Washington, DC

Glasser W, Northey R, Schultz T (eds) (2000) Lignin: historical, biological and material perspectives, ACS Symposium Series, vol 742. American Chemical Society, Washington, DC

Goldstein I (ed) (1981) Organic chemicals from biomass. CRC Press, Boca Raton, FL

Haygreen J, Bowyer J (1996) Forest products and wood science, 3rd edn. Iowa State University Press, Ames, IA

Heitner C, Dimmel DR, Schmidt JA (eds) (2010) Lignin and lignans—advances in chemistry. CRC Press, Boca Raton, FL

Hillis WE (ed) (1962) Wood extractives. Academic, New York

Hoadley R (1990) Identifying wood: accurate results with simple tools. Taunton Press, Newtown, CT

Hon DN-S, Shiraiski N (1991) Wood and cellulosic chemistry. Marcel Dekker, New York

Niska KO, Sain M (eds) (2008) Wood-polymer composites. CRS Press, Boca Raton, FL

Roser D, Asikainen A, Raulund-Rasmussen K, Stupak I (eds) (2008) Sustainable use of forest biomass for energy. Springer, Berlin

Rowell RM (ed) (1984) The chemistry of solid wood. American Chemical Society, Washington, DC

Rowell RM, Caldeira F, Rowell JK (eds) (2010) Sustainable development in the forest product industry. Universidade Femando Pessoa, Porto, Portugal

Rowell R, Young RA (eds) (1997) Paper and composites from agro-based resources. CRC Press, New York

Rydholm SA (1965) Pulping processes. Wiley, New York

Sarkanen KV, Ludwig CH (eds) (1971) Lignins. Wiley, New York

Schultz TP, Militz H, Freeman MH, Goodell B, Nicholas DD (eds) (2008) Development of commercial wood preservatives—efficacy, environmental, and health, ACS Symposium Series, vol 982. American Chemical Society, Washington, DC

Sjostrom E (1981) Wood chemistry: fundamentals and applications. Academic, New York

Smulski S (ed) (1997) Engineered wood products, A guide for specifiers, designers and users. PFS Research Foundation, Madison, WI

Toshihirona O (ed) (2004) The improvement of forest resources for recyclable forest products. Springer, Tokyo

Young RA, Giese RL (eds) (1990) Introduction to forest science, 2nd edn. Wiley, New York

Young RA, Giese RL (eds) (2004) Introduction to forest ecosystem science and management, 3rd edn. Wiley, New York

Young RA, Rowell RM (1986) Cellulose: structure, modification and hydrolysis. Wiley-Interscience, New York

Young RA, Akhtar M (eds) (1998) Environmentally friendly technologies for the pulp and paper industry. Wiley, New York

Articles and Chapters

Akhtar M, Blanchette R, Myers G, Kirk T (1998) An overview of biomechanical pulping. In: Young RA, Akhtar M (eds) Environmentally friendly technologies for the pulp and paper industry. Wiley, New York

Campbell MM, Brunner AM, Jones HM, Strauss SH (2003) Forestry's fertile crescent: the application of biotechnology to forest trees. Plant Biotechnol J 1:141–154

Chum HL (1989) Structural materials for the automobile of the future—composite materials. In: Chum HL (ed) Assessment of biobased materials. Solar Energy Research Institute, Report No. SERI/TR-234-3650, Department of Energy, Golden, CO, pp 1.1–1.19

Hegert H (1998) Developments in organosolv pulping—an overview. In: Young RA, Akhtar M (eds) Environmentally friendly technologies for the pulp and paper industry. Wiley, New York

Ince P, Schuler A, Spelter H, Luppold E (2007) Globalization and structural changes in the U.S. forest sector: an evolving context for sustainable forest management. General Technical Report, FPL-GTR 170, USDA Forest Service, Forest Products Laboratory, Madison, WI

Johansson A, Aaltonen O, Ylinen P (1987) Organosolv pulping—methods and pulp properties. Biomass 13:45

Lai Y-Z (1991) Chemical degradation. In: Hon DN-S, Shiraishi N (eds) Wood and cellulosic chemistry. Marcel Dekker, New York

Lai Y-Z (1995) Reactivity and accessibility of cellulose, hemicelluloses, and lignins. In: Hon DN-S (ed) Modification of lignocellulosic materials. Marcel Dekker, New York

Levan SL, Ross RJ, Winandy JE (1990) Effects of fire retardant chemicals on the bending properties of wood at elevated temperatures. USDA Forest Service Research Report FRL-RP-498, Forest Products Laboratory, Madison, WI

Mansfield SD, Esteghlalian AR (2003) Applications of biotechnology in the forest products industry. Applications of enzymes to lignocellulosics. ACS Symp Ser 855:2

Nelson P (1998) Elemental chlorine free (ECF) and totally chlorine free (TCF) bleaching of pulps. In: Young RA, Akhtar M (eds) Environmentally friendly technologies for the pulp and paper industry. Wiley, New York

Sanadi A, Caulfield D, Jacobsen R (1997) Agro-fiber/thermoplastic composites. In: Rowell R, Young RA (eds) Paper and paper composites from agro-based resources. CRC Press, New York

Winandy JE, Williama AW, Rudie AW, Ross RJ (2008) Opportunities for using wood and biofibers for energy, chemical feedstocks and structural applications. In: Pickering KL (ed) Properties and performance of natural-fibre composites. CRC Press, Boca Raton, FL

Young RA (1986) Structure, swelling and bonding of cellulose fibers. In: Young RA, Rowell RM (eds) Cellulose: structure, modification and hydrolysis. Wiley-Interscience, New York, pp 91–128

Pigments, Paints, Polymer Coatings, Lacquers, and Printing Inks

29

Dean C. Webster and Rose A. Ryntz

Introduction

Change is constant in the coatings market. As mergers, acquisitions, and partnerships take shape, consolidation and globalization remain prominent. The 80/20 rule (20% of the firms accounting for 80% of business) takes effect as the need for regulatory and environmental compliance continues to plague the market. In 1975, the United States alone supported about 2,000 coatings companies. Today, there are less than half that many.

The global paint and coatings market garnered approximately $90 billion in sales in 2009. The major participants included: Akzo Nobel Coatings BV (Sassenheim, The Netherlands); PPG Industries, Inc. (Pittsburgh, PA); The Sherwin Williams Co. (Cleveland, OH); DuPont Performance Coatings (Wilmington, DE); BASF Coatings AG (Ludwigshafen, Germany); RPM International Inc (Medina, OH); The Valspar Corporation (Minneapolis, MN); Kansai Paint Co., Ltd. (Osaka, Japan), Nippon Paint Co., Ltd. (Kita-Ku, Osaka, Japan). Smaller companies, being forced out of the market because they lack the capital required to meet compliance issues, are forced into consolidations. Major acquisitions also highlighted the paint and coatings industry and its suppliers. These included Akzo-Nobel's acquisition of ICI Paints, PPG's acquisitions of Ameron and Sigma Kalon, Perstorp's acquisition of Rhodia, BASF's acquisitions of Ciba and Cognis, Sherwin-Williams' acquisition of Becker Acroma, the merger of Hexion and Momentive, and Dow Chemical's acquisition of Rohm and Haas. The economic downturn in 2008–2009 and sluggish recovery has affected the coatings market significantly. Although forecast growth rates of the coatings markets in the developed world is slow, about 4% annually in North America and similarly in Europe, Asian coatings growth is estimated to be about 8% annually [1]. Thus, all of the major paint and coatings manufacturers and suppliers to the industry have developed a presence in Asia, with a particular focus on China.

The auto industry is also providing coatings companies with opportunities to form increasingly important relationships. PPG was awarded an exclusive coatings contract at the Nissan plant which opened in Canton, Mississippi in 2003, providing everything from pretreatment to clearcoat. In addition, PPG formed a joint venture with a local minority-owned business, Systems Consultants Associates (Jackson, Mississippi) to coat chassis frames and other underbody components, as well as the Nissan replacement parts. The joint venture, Systems Electro Coating, is minority-owned and operated at a facility adjacent to the Nissan plant. A Toyota plant in France, which opened in 2001, procures all of its coatings materials from DuPont Herberts Automotive Systems.

Over the past decade, automotive supplier parks have proliferated in South America and Europe as automakers pursue a manufacturing value chain that is as lean and efficient as possible. That trend has recently also gained popularity in North America. Ford Motor Company launched its first North American supplier park in Chicago in 2004, where painted exterior and interior components will be produced. Chrysler is planning a supplier co-location project at the Jeep Toledo North Assembly plant, where painted vehicle body and the chassis module will be sent on conveyer lines directly into the Chrysler Group's final assembly building.

The major coatings markets in the United States consist of architectural coatings (consumer and decorative), product finishes (maintenance/protective coatings for bridges, industrial machinery, etc.), and special-purpose coatings, such as automotive finishes ((OEM) and refinish) (Table. 29.1) [2]. The volume of coatings delivered declined in 2008 and 2009 and began to show signs of recovery in 2010.

D.C. Webster (✉)
Coatings and Polymeric Materials, North Dakota State University, Fargo, ND 58108-6050, USA
e-mail: dean.webster@ndsu.edu

R.A. Ryntz
International Automotive Components, Southfield, MI, USA

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_29,

Table 29.1 U.S. sales of paints and coatings (millions of gallons) [2]

	Year				
Coating type	2006	2007	2008	2009	2010
Architectural	768	777	682	643	652
OEM product	382	370	341	286	331
Special purpose	210	207	178	161	171

Despite years of environmental pressure to reduce the volatile organic compound (VOC) component in paints, solvent systems are still king, compromising 47% of the U. S. paint market [3]. Government pressure to reduce the contributions these systems can make to ground-level ozone formation (smog) has reached its limit. Most formulators have either substituted acceptable solvents or reduced solvent content to produce acceptable high-solids paint systems.

Not all government authorities believe that VOC reduction limits have been reached. California's South Coast Air Quality Management District (SCAQMD), which includes counties in the Los Angeles Basin, keeps squeezing paint makers for further reductions. The agency has now succeeded in regulating that liquid coatings such as wood stains, sealers, and roof coatings, contribute VOCs of only 50 g/L (approximately 0.4 lbs/gal), down by 50% from previous regulations and is proposing that additional products have reduced VOC limits in the coming years.

The coatings market is constantly being sought out by local, state, and federal governments to reduce environmental emissions even more. The key concerns of government regulators and consumers include: treating the earth as one biosphere, having an infinite capacity for cleaning itself of toxins; cleaning air from pollutants that are killing lakes and causing a dramatic increase in respiratory illnesses; ensuring that smog-enshrouded cities do not become industrial nightmares; and making the agricultural and livestock products that people eat safe for consumption over the long term.

These concerns have led to legislation over the last several years that cause manufacturers to place a greater emphasis on research that enables them to meet or exceed government and consumer demands for clean air.

VOC Regulations

The environmental demands have lead to a nationwide program to limit VOCs. The VOCs of paint are calculated as pounds of solvents per gallon of paint. Prior to 1970, the VOC content of most paints was well above 5 lb/gal. Current major industrial paints now are limited to approximately 3.5 lb/gal of VOC because of imposed environmental pressures. Stricter legislation in the future, however, will place requirements of less than 3.0 lb/gal of VOC on most markets. The issue of air quality compliance is complicated by different regulations in various states and regional areas. Although the federal government (through the U.S. EPA) has set minimum national standards, states and designated regional areas can adopt their own compliance rules.

All states and regions tend to restrict the solvent content in coatings. The following summaries indicate the wide variation in compliance requirements throughout the country.

West

California. This is where the regulations began. The regulatory activities of the SCAQMD have set a pattern that many other regulatory boards have chosen to follow. The district restricts application equipment and VOC content. The VOC requirements are continuously changing; so no attempt was made to list them here.

The SCAQMD in Southern California has mapped out a three-tier plan to bring the district into clean-air compliance. The plan will be revised periodically based on reports from over 40 monitoring stations throughout 13,350 square miles including Los Angeles, Orange, and Riverside counties and the urban portion of San Bernardino County. VOCs will not be monitored directly. Rather, ozone, which is formed when hydrocarbons (VOCs) react with nitrogen oxides in sunlight, will be measured for the presence of VOCs. The federal clean-air standard for ozone requires that ozone concentrations do not exceed 0.12 ppm more than 1 day a year.

Tier I of the three-tier plan involves 22 rules, which were adopted in 1993, to restrict the emission of VOCs from coatings. Control methods include the application of low-solvent coatings, use of high transfer efficiency (TE) techniques to apply coatings, and "to a certain extent" the use of add-on control devices (see Table. 29.2).

Tier II (adopted in 2000) reduced the VOC emissions by 50%. Reductions were achieved by further tightening the VOC content of coatings (at times to levels lower than those stated in the Federal Control Technique Guidelines (CTG)) and increasing application transfer efficiency requirements (at times to levels more stringent than those stated in the EPA "Blue Book").

Tier III (initially to be implemented by December 31, 2007) calls for applying new technology to improve waterborne, UV-cure, and two-component coatings, and prohibiting "certain high-emitting uncontrollable coating processes."

Washington. No rule prohibits any method of spray finishing. VOC guidelines are determined by subtracting

Table 29.2 Application equipment requirements of SCAQMD clean-air act[4]

	Rule									
Application equipment	1104	1106	1107	1113	1115	1124	1125	1136	1145	1151
Dip		X			X	X	X			
Electrostatic application			X			X	X	X		X
Flow coat			X			X	X	X		
Hand application methods					X	X	X			
HVLP[a] spray		X			X	X	X		X	
Roll coater			X			X	X	X		
Other methods that are proven 65% transfer efficient			X			X	X	X		X
No specified application method	X	X		X	X				X	

[a] High-volume low-pressure application equipment

the amount of solids from the gross weight of the material used and determining whether the amount of solvent used exceeds the acceptable limits.

The maximum VOC limits (lb/gal) are: clear-coat, 4.3; color, 3.5; performance coatings, 3.5; all other coatings, 3.0. Metric limits (g/L) are defined for aerospace: primers, 650; topcoats, 600 (equivalent to 5.4 and 5.0 lb/gal, respectively).

Southwest

Texas. In Texas, regulators restrict emissions rather than equipment. Emissions are limited across the board for certain industries. Limits (lb/gal) are specified for particular coatings in other industries. For example: large appliances, 2.8; furniture, 3.0; coil coatings, 2.6; automotive OEM—primers, 1.2 if submersed and 2.8 if sprayed, topcoats, 2.8, and repair, 4.8; miscellaneous metal—color, 2.5, top clear-coat, 4.3, and all others, 3.0.

South Carolina. VOC limits (lb/gal) include: metal furniture, 3.0; appliances, 2.8; miscellaneous metal—color, 3.5, clear-coat, 4.3, and all others, 3.0; coil coating, 2.6.

North Carolina. The rules are applicable only in areas designated as nonattainment zones. If air pollution control equipment is not used, emission standards are based on the weight of solids used. High-solids coatings are automatically compliant regardless of spray equipment used. The limits (lb/gal) for automotive shops with air pollution control equipment include: primers, 1.4 if high solids, 1.2 if low solids; topcoats, 4.5 if high solids, 2.8 if low solids; repair, 13.8 if high solids, 4.8 if low solids. For automotive shops without air pollution control equipment, limits are: primers, 1.2; topcoats, 2.8; repair, 4.8. Air pollution control equipment is not precisely defined. Other limits include: metal furniture, 5.1; appliances, 2.8.

Georgia. VOC emissions are limited (lb/gal) by industry and coatings used: automotive—primers, 1.2 if dipped and 2.8 if sprayed, topcoat, 2.8, and repair, 4.8; metal furniture, 3.0; appliance, 2.8; and miscellaneous metal-performance coatings, 3.5, high-performance coatings, 6.2, clear-coat, 4.3, and all others, 3.0; coil coating, 2.6.

Midwest

Indiana. Rules in Indiana are among the most complex in the nation. They require a baseline TE, specify coatings formulations, and limit VOC emissions (lb/gal) for automotive, metal furniture, and large appliances but then only specify acceptable equipment choices for wood. Limits are: automotive—30% TE at 62% solids; metal furniture—60% TE at 59.2% solids; large appliances—60% TE at 62% solids. No TE test procedure is defined for the following VOC limits: automotive—primer, 1.9, topcoat, 2.8, and repair, 4.8; metal furniture, 3.0; large appliances, 2.8; miscellaneous metal—color, 3.5, specialty coats, 3.5, clear-coats, 4.3, coil coatings, 2.6, and all others, 3.0. In the wood industry, the state departs from these guidelines and identifies acceptable equipment choices: airless, heated airless, and air-assisted airless spray; manual and automatic electrostatic spray; and rollers brushes, or wipes. Significantly, no conventional air spray method, including HVLP, is recognized as compliant.

Illinois. A 30% minimum TE is required for automotive. Automotive VOC limits (lb/gal) are: primer/surface, 2.8; topcoat, 2.8; repair, 4.8. The rules allow higher VOC limits for automotive topcoats if the finishing method can be demonstrated to achieve a TE of over 65%. In such cases, topcoats may be used with a VOC of up to 3.6. For the wood industry, a minimum TE of 65% is required with the following VOC limits: washes, 6.5; stains, 4.7; top-stains, 6.6; sealers, 5.6; clear-coats, 5.6. No test procedure for TE has been defined.

Michigan. Although VOC guidelines are suggested, they are not stated clearly. In general, the finisher is expected to achieve minimal VOC levels when using metallic topcoats of waterborne coatings and performing touchup or repair work. Although conventional air spray is generally acceptable, some areas of Michigan require electrostatic systems with a TE of at least 40% and define VOC limitations.

East

New Jersey. A minimum of 65% TE is required for all industries except wood. Shops using waterborne coatings are automatically compliant regardless of TE. VOC limits (lb/gal) for all others are: automotive—primers, 2.8, topcoats, 2.8, specialty coats, 5.0, and repair, 4.8; automotive refinishing—basecoats, 6.0, clear-coats, 4.4, and all others, 5.0; coil coating, 2.6; miscellaneous metals—color, 4.3, performance coatings, 3.5, and all others, 3.0; leather, 5.8; urethane users, 3.8. Wood industry rules specify equipment by the shop size. Wood shops using less than 50 tons of coatings a year must achieve a minimum TE of 40%. Approved spray-finishing methods include airless, heated airless, and air-assisted airless. Wood shops that use 50 or more tons of coatings a year are required to achieve a TE of 65% or higher. Approved spray-finishing methods for these shops include airless, heated airless, air-assisted airless, electrostatic spray, or flat line finishing. VOC limits (lb/gal) are: wash-coat, 6.1; stain, 4.7; pigment, 5.0; sealer, 5.6; clear-coat, 5.6.

Pennsylvania. The rules specify only that wood cabinet and furniture finishers must use air-assisted airless or electrostatic spray.

New York. New York follows the same rules as Pennsylvania.

Ohio. Ohio has adopted VOC limitations and will require a TE of at least 30%.

Hazardous Waste Regulations

Included in regulatory compliance is the issue of hazardous wastes. On March 29, 1990, the U.S. EPA promulgated a final rule revising the procedure used to determine whether wastes are hazardous because of their potential to leach significant concentrations of toxic constituents to landfills. The extraction procedure (EP) toxicity characteristic procedure was replaced with the toxicity characteristic (TC), utilizing the toxicity characteristic leaching procedure (TCLP). The U.S. EPA also added to the list 25 organic chemicals commonly used in surface-finishing operations (i.e., methyl ethyl ketone, tetrachloroethylene, etc.).

Table 29.3 Maximum concentration of contaminants allowable from TCLP[a][5]

EPA Hazardous waste no.	Contaminant	Regulatory level (mg/L)
D004	Arsenic	5.0
D005	Barium	100.0
D018	Benzene	0.5
D006	Cadmium	1.0
D019	Carbon tetrachloride	0.5
D021	Chlorobenzene	100.0
D022	Chloroform	6.0
D007	Chromium	5.0
D023	*o*-cresol	200.0
D024	*m*-cresol	200.0
D025	*p*-cresol	200.0
D026	Cresol	200.0
D027	1,4-dichlorobenzene	7.5
D028	1,2-dichloroethane	0.5
D029	1,1-dichloroethylene	0.7
D030	2,4-dinitrotoluene	0.13
D032	Hexachlorobenzene	0.13
D033	Hexachlorobutadiene	0.5
D034	Hexachloroethane	3.0
D008	Lead	5.0
D009	Mercury	0.2
D014	Methoxychlor	10.0
D035	Methyl ethyl ketone	200.0
D036	Nitrobenzene	2.0
D037	Pentachlorophenol	100.0
D010	Selenium	1.0
D011	Silver	5.0
D039	Tetrachloroethylene	0.7
D040	Trichloroethylene	0.5
D041	2,4,5-trichlorophenol	400.0
D042	2,4,6-trichlorophenol	2.0
D043	Vinyl chloride	0.2

[a] Excluding D012–D017 and certain pesticides

Although the maximum allowable concentrations of the regulated eight metals (and pesticides, herbicides, and insecticides) do not change from those listed under the EP toxicity, the new TCLP analytical procedure may make it more difficult for listed wastes to stay under the allowable concentrations (Table. 29.3).

Technical Trends in Coatings

Through the years, improved paints have been developed because of competition from environmental and consumer activists. The objective of improving overall quality and reducing emissions of VOCs has led to a wider selection of

Table 29.4 Coatings markets served by leading paint and powder coating producers[6]

Market segment	Average percent of total coatings output
General industrial	34.4
Automotive	18.0
Metal furniture/business equipment	8.6
Metal containers/closures	9.0
Wood furniture/wood products	12.5
Appliance/HVAC	3.3
Lawn and garden	2.4
Aircraft	4.0
Other	7.9

Table 29.5 Evolution of coatings development[7]

1923	Low viscosity nitrocellulose
1923	Alkyds
1924	Titanium dioxide
1928	Phenolic resins (oil-soluble)
1930	Urea formaldehyde–alkyd mixtures
1933	Vinyl chloride copolymer
1934	Oil-based emulsion paints
1936	Thermoset acrylics
1937	Polyurethanes
1939	Melamine formaldehyde–alkyl mixtures
1944	Silicon-based paints
1947	Epoxy resins
1950	Polyvinylacetate and acrylic emulsions
1955	Powder coatings
1958	Latex house paint
1960	Water-based coatings
1962	Anodic electrodeposition
1963	UV- and electron-beam cure
1971	Cathodic electrodeposition
1974	Clear over color topcoats
1989	Supercritical CO_2

paint types. The user, therefore, must select a coating type that meets the cost of required application, equipment, energy costs for using, pollution and waste collection expenses, and application and cured performance requirements of the coating.

Table 29.4 lists types of coatings markets served by leading paint and powder coating producers. By far, the leading type of coating manufactured for the OEM market produced by the suppliers surveyed is waterborne (87.4%), followed by low-solids solvent-borne (78.7%), high-solids solvent-borne (74.0%), and two-component coatings (69.3%). The percentage of powder coatings produced by the surveyed producers accounted for only 26.0%, while ultrahigh solids and UV/EB coatings added an additional 22.0 and 22.8%, respectively. The majority of those producers surveyed suggested that most of their R&D spending would be on waterborne coatings. While powder coating research continues to grow, E-coat and radiation cure research continues through suppliers in niche markets. Lower co-solvent, lower cost, and shortened cure times continue to be the hottest areas of research.

A solvent-borne coating generally consists of seven major constituents:

- Resin (binder).
- Cross-linker.
- Rheology control agent.
- Catalyst.
- Pigment.
- Flow additive.
- Solvent(s).
- Weathering stabilizers (if required).

The resin (often referred to as binder) is responsible for the main attributes of the coating when cured with the cross-linker (or cross-linking agent).

The curing, or network formation, often is accelerated through the use of a catalyst. Rheology control agents, often referred to as sag control agents, are added to impart shear-thinning behavior when under high shear (i.e., spray application) and rheology or anti-sagging behavior when under low shear (applied on substrate). Pigments, discussed later in this chapter, are introduced to impart color.

The amount of pigment utilized depends on the color and the hiding power required of the coating. The flow additive often is introduced to relieve surface tension variations between the coating and substrate, to eliminate pinholes or crater formation. Solvents are added as necessary to achieve flow under application conditions.

In order to achieve a VOC-compliant coating, one must consider the molecular weight of the main binder utilized in the coating formation, as well as the effect of additives on the final viscosity. After all, application viscosity, which in turn is affected by the rheology of the coating constituents, is most often controlled by freely adding additional solvent to achieve necessary flow and leveling upon application. This in turn affects the VOC emission of the coating.

If one considers the molecular weight dependency of resins utilized in the various coating types, it is readily evident why VOC compliance is more easily achieved in radiation-curable and powder coatings. The molecular weight of nonaqueous dispersion resins, for example, exceeding 10^5 g/mol, results in very viscous materials, whereas the typical molecular weight of a radiation cure oligomer ranges from 10^2 to 10^3 g/mol, for a very low viscosity material, which requires little if any additional solvent to attain application specifications.

If we consider the evolution of coating technology over the last 70 years (see Table. 29.5) we can see that by the mid-1950s, waterborne and higher solids coatings already were under development. Powder coatings, 100% reactive liquid coatings (electron/beam and UV-cure), electro-coatings, aqueous or waterborne (water-base) coatings, and high solids coatings can all be considered as novel "non-polluting" VOC-compliant coating systems.

Powder Coatings

Today, powder coatings are one of the fastest growing finishing technologies in North America, with a growth rate (in pounds sold) of 7–9% per year over the next 5 years. The major impetus for growth can be attributed to the virtual "pollution-free" finishing system as well as the dramatic leap forward in materials and equipment technology.

Powder coating resin systems have been developed to meet specific needs of the metal finishing operations. Epoxies were utilized almost exclusively during the early years of thermosetting powder coatings and continue to be utilized in a range of applications. However, because of the poor outdoor durability of epoxy resins, acrylics, and polyester-urethane hybrids are finding increasing use.

Powder coating manufacturers continue to work on perfecting flow and leveling of the coatings, which often are prone to "orange peel." The other major disadvantage of thermosetting powder coatings is the high temperature (300–350°F) required to cure these systems. Research seeks to lower the temperature of cure.

Electron Beam and Ultraviolet Curable Coatings

Radiation-curable coatings were developed by W. J. Burlant (of the Ford Motor Company) in 1962. Acceptance of finishes designed for radiation curing (high-solids materials formulated to be cured by UV and EB systems) is growing at roughly 8% per year with a market size of approximately 85 million lb.

The major advantages of radiation-curable coatings are speed of cure, reduced emissions, 100% solids systems, and elimination of ovens, or other sources of thermal energy. The most rapidly growing areas in UV are UV-curable powder coatings and new applications for UV-curable coatings for automotive parts. The UV coating usually is comprised of unsaturated monomers that polymerize with UV light (in the presence of a photo-initiator, which absorbs radiation and initiates polymerization). Unsaturated polyesters copolymerized with vinyl monomers or multifunctional acrylates often are utilized. Most recently, however, vinyl ether monomers for cationic curing with epoxies have shown promise due to reduced shrinkage of cured films and improved adhesion to substrates. Waterborne UV-curable coatings have also recently been introduced.

Major uses for UV curing (room temperature) include use of flat stock-fillers in particle boards and hardwood flooring, and for high-gloss (no-wax) protective coatings for floor tiles, wood finishes, and other specialty uses.

Current Automotive Coating Trends

Until 1965 automotive topcoats consisted of nothing more than acrylic lacquer technology. As technology improved to include solvent-borne nonaqueous dispersions in the mid-1960s and medium solids thermoset coatings in the 1970s, two competing factors were becoming more important: control of emissions (VOC) and the quality of the appearance of the topcoat. Today, the quality of the topcoat includes chip resistance, environmental etch resistance, and metallic glamor; so virtually all automotive procedures (except in truck plants) now utilize basecoat/clear-coat technology. The basecoat is the pigmented film that is responsible for providing color. It is covered, in a wet-on-wet spray process, with a clear-coat that provides chemical resistance, UV resistance, chip resistance, and distinctness of image (DOI) (often referred to as the "wet look").

The total automotive coating composite, however, consists of more than just the basecoat/clear-coat. The metal sheet stock that is formed into the automotive car body is treated with a corrosion-resistant primer applied by cathodic electrophoresis. This coating, often referred to as "ECOAT" or "uni-prime" is a hydroxy functional epoxy/blocked isocyanate dispersion that deposits onto the cathode of an electrolytic cell at a voltage of between 250 and 425 V. Film thickness, which is dependent upon the residence time, temperature, and coating voltage of the electro-coat bath (electrolytic cell), varies between 0.8 and 1.2 mils (1 mil = 25 μm). The curing temperature required to deblock the isocyanate and yield a cross-linked network varies from 275 to 375°F for 20 min. The function of this coating is to provide a corrosion-resistant film to the metal sheet stock.

A primer surfacer, composed most often of a polyester to be cross-linked with a blocked isocyanate or melamine cross-linker, is applied next at a dry film thickness of approximately 1.0 mil by conventional spray application. The function of this coating is to provide filling and leveling properties as well as some stone-chip resistance. Cure conditions range from 250 to 325°F for 20 min.

The solvent-borne basecoat, consisting of polyester or acrylic binder to be cross-lined with melamine or isocyanate, is applied at a dry film thickness of between 0.6 and 1.5 mils, depending upon the color. Poorer-hiding colors, such as red and white straight shades, require the higher film builds, whereas metallics and darker colors are applied at the lower film builds. The solvent-borne basecoat is allowed to flash-off solvent for approximately 3–5 min at ambient conditions prior to application or the solvent-borne clear-coat. This process is referred to as a weton-wet application because the basecoat is not in its cured or cross-linked state.

Optionally, a waterborne basecoat can be utilized to achieve better aluminum orientation in metallic films. Several papers [8, 9] have described metallic orientation and the advantages of waterborne basecoats, so this concept will not be thoroughly discussed here.

The waterborne basecoat consists of a urethane or an acrylic dispersion with passivated aluminum for metallic glamor. The aluminum must be passivated to prevent reaction with water and the evolution of hydrogen gas. Passivating agents can include polymers, hexavalent chromium (now mandated for removal), or phosphate esters. The basecoat is pre-baked (with infrared or thermal ovens) to release the majority of the water prior to clear-coat application.

The solvent-borne clear-coat is spray-applied over the basecoat at a dry film thickness of 1.8–2.0 mils. The clear-coat, responsible for UV protection, chemical resistance, chip resistance, and so on, is most often composed of a hydroxy functional acrylic binder together with a melamine or an isocyanate cross-linker. UV light absorbers and hindered amine light stabilizers also are added to achieve better free-radical oxidative degradation resistance when the coating is exposed to sunlight.

The clear-coat is referred to as a two-component (2K) clear-coat when isocyanate is used as the cross-linker. The isocyanate must be added to the hydroxy functional acrylic just prior to spray application to prevent premature cross-linking (or gelation). For this reason, the two-part mixing is referred to as 2K. Cure conditions for isocyanate-cured clear-coat range from room temperature (car refinish) to 250°F for 30 min.

One component, or 1K clear-coats are equivalent to 2K clear-coats with the exception that they are cross-linked with melamines. Cure conditions for 1K clear-coats are 250°F for 30 min. Blocked isocyanate versions of 1K clear-coats also are available, but cure requirements necessitate 275°F or higher.

Significant research is being conducted on waterborne clear-coats in attempts to achieve better VOC emission levels. Waterborne clear-coat research is limited to water-dispersible acrylics because high gloss and good weatherability can be attained. Removal of water from the film, however, is a problem, as coalescence often begins prior to through-cure, often resulting in “solvent-popping.” Much work remains to be performed on optical formulation parameters (i.e., solvent blend, anticratering surfactants, rheology control, etc.) in order to attain defect-free coatings.

Coatings for Plastics

The use of plastics in automotive applications has expanded considerably over the past 20 years. In 2001, approximately 250 lb of plastics were used per automobile compared to less than 100 lb five years earlier. With the increased use of plastics, coating research has concentrated on the painting of these often difficult substrates. Among the numerous prerequisites for coatings on plastic are: wetting of the substrate; uniform film development; a cure response below the heat distortion temperature of the plastic; adhesion to the substrate; inhibition of substrate imperfections; and often a conductive nature to increase the TE of subsequent topcoat application.

Plastic types are variable, ranging from flexible to rigid and from thermoplastic to thermoset. Because of these variations, the type of coating applied to them also varies, depending on cure response, solvent sensitivity, and modulus. The type of coating also depends upon the end use of the coated plastic. For example, topcoats can be applied directly to the plastic without the use of a conductive primer if two-toning or blackout areas are not part of the styling latitude. Adhesion promoters also are often utilized if a particularly difficultto-adhere-to plastic is encountered (i.e., polypropylenes or thermoplastic polyolefins (TPOs), a blend of elastomer and olefin).

Conductivity in plastic coatings is achieved through the use of a conductive pigment, most often conductive carbon black. Conductivities of the order of 1 MΩ/cm^2 are required for effective transfer efficiencies of subsequent topcoat applications. Without the use of a conductive primer, plastics in themselves would not be conductive, and hence would be very unreceptive to topcoat transfer efficiencies of greater than 20%.

The adhesion of a coating to plastic is related to the substrate wettability and also to physical–chemical forces. Both covalent (bond formation) and dispersive (London or van der Waals) forces are responsible for the adhesion of the coating due to physical–chemical forces.

Substrate wetting often is measured by contact angle measurements: the smaller the contact angle, the better the wetting of a coating to the plastic. The wettability of a coating on a plastic substrate is related to the surface tension of the substrate: the lower the surface tension of the plastic, the more difficult it is to wet.

Polypropylene, in particular, has a very low surface tension (less than 32 dynes/cm^2). Because of the low surface tension, it is very difficult to wet out; and because of the apolar nature of the polypropylene backbone, it is a very difficult material to adhere to. Therefore, copolymers of ethylene propylene diene monomers (EPDM) with polypropylene have been developed. The higher the level of the EPDM in these TPOs, the more paintable is the substrate.

Conventional coating approaches to TPO, which rely on the use of substrate pretreatments such as flaming, corona discharge, plasma or chemical pretreatment, have led to a substantial oxidation of the surface, which enhances adhesion. These pretreatments, however, have a limited service life before the surface reverts to its unoxidized, apolar condition.

Chlorinated polyolefins (CPOs) also are a mainstay of adhesion promotion to TPO. The CPO generally is applied as a dilute solution (5–35 wt. %) in aromatic solvents to the TPO to act as a "tie-coat," to which the topcoat is applied. These materials are often humidity sensitive and only adhere to TPOs with EPDM levels of greater than 5%.

New Cross-Linking Technologies

The traditional cross-linking technologies utilized in the coatings market range from hydroxyl functional binders (acrylics or polyesters) cross-linked with melamines (1K) or isocyanates (2K) to drying oils that cure through oxidation of unsaturation in the backbone, to epoxy resins cross-linked with amino resins (amido amines, amines, etc.).

New resin and cross-linking technologies, however, have been introduced with hopes of achieving more environmentally acceptable coatings. However, most of these have not achieved widespread success with the possible exception of areas such as automotive coatings where they can provide properties such as chip resistance and resistance to "acid rain."

Bio-Based Resins

As companies begin to focus on issues such as sustainability, and "green" chemistry, interest in the use of bio-based resins in coatings has been rapidly increasing. Procter & Gamble Chemicals introduced a line of sucrose ester resins of vegetable oils for industrial applications under the brand name SEFOSE. In collaboration with Cook Composites and Polymers (CCP), they demonstrated that SEFOSE resins could be used as a reactive diluent in alkyd paints to reduce the VOC from 440 g/L to 250 g/L and won the 2009 Presidential Green Chemistry Challenge Award [10]. Other companies have also introduced new bio-based products, including the Solus line of cellulose ester resins by Eastman, Albecor-Bio, a bio-based polyester for powder coatings by Hexion, and Dow Chemical introduced the Renuva line of bio-based polyols. Sherwin-Williams received the 2011 Presidential Green Chemistry Challenge Award for their new paint technology based on soybean oil and recycled PET.

Printing Inks

Printing inks, utilized in printing and packaging applications, are colored, liquid materials consisting of essentially two components: the colorant, an insoluble solid or a dye, and the vehicle, a liquid in which the colorant is suspended or dissolved. Recent years have seen significant progress in water-based inks, but compared with the phenomenal growth in water-based paints, the advances in water-based inks have had a much smaller impact on the industry. The upsurge of interest in water-based inks, which began in the mid-1970s, was stimulated by the need to conserve expensive petrochemical-derived organic solvents, attention to environmental pollution problems, and increasing pressures from health and safety legislation.

Inks can be applied to a wide variety of substrates, including plastics, paper, textiles, metals, and glass. In contrast to paint films, inks are applied very thin, varying from 2 to 30 μm depending on the print process. The print process consists of application of the ink to the substrate by a printing press of varying design. As such, inks are classified by the type of process through which they are applied. There are five classifications: flexo-graphic, gravure, offset lithographic, letterpress, and screen.

The mechanisms of letterpress, offset lithographic/flexographic, and gravure printing processes are outlined in Fig. 29.1 [11]. In letterpress and offset lithographic processes, a thin film of ink is carried to the printing plate by means of a train of rotating rubber rollers. Both of these processes preclude the use of volatile organic solvents in the inks because of the swelling effect or attack on the rubber rollers and blankets and because of their high evaporation rates, which could lead to premature drying or tack development of the inks on rollers, blanket cylinders, and printing plates. Water, too, evaporates more rapidly than the high-boiling aliphatic-type solvents normally used in letterpress and offset lithography, and the evaporation is accelerated by the heat that can be developed (temperatures of 30 °C or more) in the rollers after a period of running.

Several different types of drying mechanisms may be operative in printing of letterpress and offset lithographic inks (see Table. 29.6) [12]. As shown in Table. 29.6, the ink distribution systems of flexo-graphic and gravure presses are quite different from those of letterpress or offset lithography. There are no long roller trains, and because the drying mechanisms depend mainly upon evaporation of volatile solvent components, it is not surprising that most of the development work has been on inks for these processes.

Inks that are printed by the flexo-graphic and gravure printing processes are characterized by their extremely fluid nature, and are generally termed "liquid inks." When

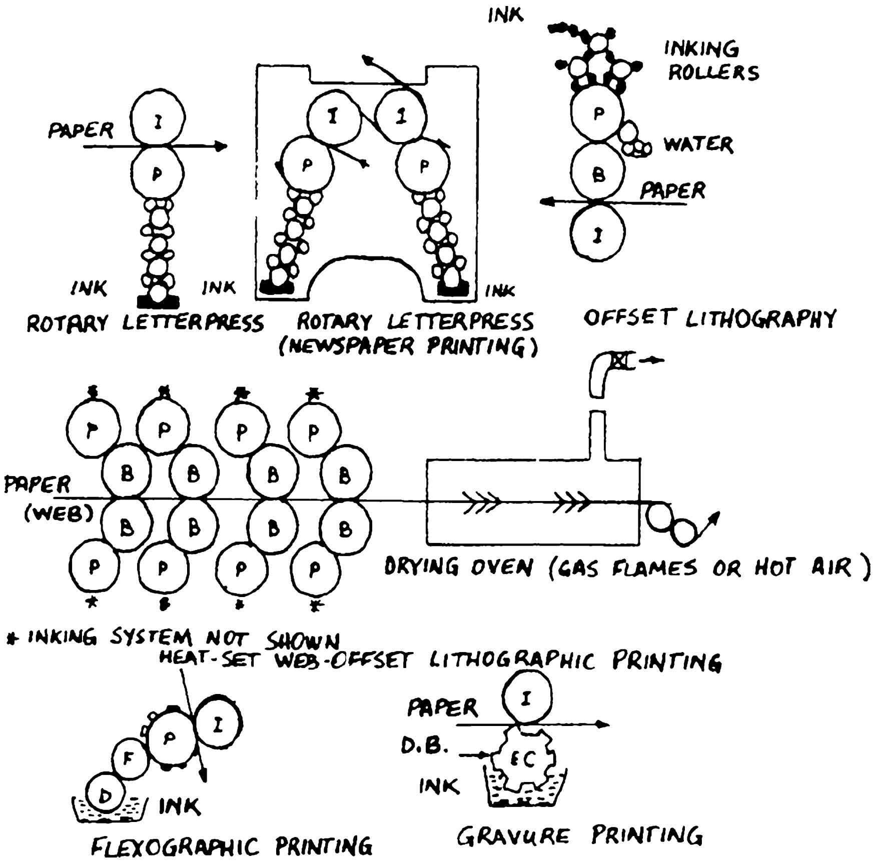


Fig. 29.1 Mechanisms of letterpress, offset lithographic/flexo-graphic, and gravure printing processes[12]

Table 29.6 Letterpress and lithographic inks—dry mechanisms[12]

Type ink or process remarks	Principal ink Vehicle	Typical substrates	Drying mechanism
Letterpress W	Mineral oil	Newsprint A	Non-volatile mineral oil
Letterpress S	Oleoresinous and solvent	Papers, boards A, O	High-boiling range solvent
Letterpress S	Resin and glycol	Papers, boards P	High-boiling glycol
Letterpress S	Polyfunctional (acrylated) compounds	Papers, boards, R foils	No solvent ultraviolet curing
Offset litho W	Mineral oil and resin and solvent	Newsprint A	High-boiling solvent
Offset litho S	Oleoresinous and solvent	Papers, boards A, O	High-boiling solvent
Offset litho S	Oleoresinous and solvent	Tinplate O, E (stoving)	High-boiling solvent
Offset litho S	Polyfunctional (acrylated) compounds	Papers, boards, R foils, tinplate	No solvent ultraviolet curing
Offset litho W	Resin/solvent	Papers A, E (heat-set)	Gas flame hot air, high-boiling solvent

Notes: *A* Absorption, *O* Autoxidative/polymerization, *P* Precipitation/gelation, *R* Photochemical/polymerization, ultraviolet cure, *S* Sheet-fed, *W* Web-fed

compared to the fluid nature of flexo-graphic and gravure inks, lithographic and letterpress inks are much more viscous and paste-like.

In screen printing, the ink is forced through the open areas of a stencil supported on a mesh of synthetic fabric, stretched across a frame. The ink is forced through the mesh onto the substrate underneath by drawing a squeegee across the stencil. The inks have to remain fluid and coherent on the mesh before printing, transfer through the mesh without its sticking to the substrate, flow to make the ink film continuous and so avoid the pattern of the mesh but maintain the image, and then dry rapidly once applied to the substrate. Inks that are able to print by this method are of household paint consistency; they are intermediate in viscosity between the fluid flexo-graphic and gravure inks and the paste-like lithographic inks.

An ink must be transformed to a solid from its liquid supply form once on the substrate. This change of state is referred to as an ink drying, and can be physical (absorption or evaporation) or chemical (oxidative or photochemical radiation induced). There are several forms of radiation used to dry ink, including UV, infrared, EB, microwave, and radio frequency.

Pigments

Pigments are generally incorporated into paints to affect properties associated with appearance such as color, opacity, gloss, metallic look, and depth. In addition, pigments commonly are used to protect the substrate against corrosion, attack by microbes or barnacles, and to retard flammability. Other desirable properties can be obtained by the addition of pigments. For example, pigments can be used in paints to control flow and leveling as well as either spray viscosity or brush-ability, depending upon the type of system. In addition, pigments are added to paint to control the cost of raw materials.

One of the most important properties of pigments, however, is the ability to cover or "hide" the surface being painted. This property is commonly referred to as "hiding power" or opacity. The fewer coats required to hide the surface, the higher the hiding power or the pigment's opacity. Pigments of different colors have different degrees of opacity. In general, darker colors have a higher hiding power than lighter colors, in part because of their relative opacities. In practice, this trend is also dependent upon the refractive index of both the pigment and the binder; that is, the difference in refractive index between the pigment and the binder determines the hiding power of the paint. Hence, the larger the refractive index difference, the higher the hiding power. In addition, the particle size of the pigment critically affects the hiding power of the pigment. There is an optimum particle size range, 0.2–0.4 μm (which is approximately one half the wavelength of light in air), which affords maximum light scattering and hiding.

Pigments ideally should be non-soluble solids with low oil absorption to facilitate dispersion and allow maximum pigment loading, which would also reduce the VOC content of the paint. In addition, pigments should be chemically inert, thermally stable, pure in their composition, and nontoxic.

Pigments can be classified into two main categories: inorganic pigments and organic pigments. Among the more common inorganic pigments are white hiding pigments, extender pigments or fillers, blacks, colors, and metallic pigments. Organic pigments consist of colors and the carbon blacks.

Inorganic Pigments

Inorganic pigments generally hide better than organic pigments when dispersed in paint. Inorganic pigments impart flowability to paint because they do not tend to swell in the presence of common paint solvents. In addition, inorganic pigments do not bleed inorganic solvents, and they afford excellent color permanence, especially in exterior paints. Inorganic pigments are thermally stable and generally impart heat resistance to the coating.

White pigments are important not only as the principal pigment in white paint but also as hiding pigments in many shades and the tinting colors. The most commonly used of the white pigments are the nonreactive titanium dioxides and the reactive zinc oxide. These pigments are designed reactive or nonreactive, depending on whether they interact chemically with the binder or vehicle portion of the paint.

Titanium dioxide is available in two crystalline forms—anatase and rutile. The anatase form has about 75% of the hiding power of the rutile form and is less resistant to chalking than the rutile form. Because of its high refractive index (2.76), titanium dioxide imparts excellent opacity to the paint. Because the hiding power of the titanium dioxide is so high, good opacity can be achieved by the addition of small amounts of the pigment to the paint, thus not adding significantly to the cost of the paint. The rutile form is available in grades designed specifically for enamels, flats, and solvent-based and water-based coatings [13].

Although zinc oxide has about 15% of the hiding power of titanium dioxide and has a fairly high cost, it contributes a number of important properties to paint. Zinc oxide can form zinc soaps in some formulations, which can aid in pigment wetting and grinding. It also can be utilized to improve film hardness and color retention of the film on aging. The most common use of zinc oxide is in exterior finishes to reduce chalking and the growth of mildew in house paints [14].

Extender pigments, because of their relatively low refractive index, exhibit poor hiding power and must be used in conjunction with pigments of high hiding power. The principal function of extender pigments in paint is to reduce the raw material cost. They also function to reduce gloss, as well as to control texture and viscosity. Calcium carbonate, more commonly known as whiting pigment, is available in a variety of particle sizes and surface treatments and is used to control sheen, degree of flow, degree of flatting, and tint retention. Magnesium silicate or talc generally is used because of its low cost and excellent durability. Talc imparts good antisettling properties and low sheen. Aluminum silicate (China clay) is used mainly in water-based systems. China clay is easily dispersed with a high-speed dispenser and may even add to the hiding power of water-reducible paints. Other extenders such as diatomaceous silica, regular silica, barites, and mica also are frequently used to impart specific properties to coatings, which are determined mainly by the particle size and shape of the pigment.

Most of the black pigment used in coatings is organic carbon black. There are, however, a few inorganic black pigments such as black iron oxide, which is used as a tinting black having brownish tones and also is used in primers. In addition, some inorganic black pigments such as mineral or thermal blacks are used as low-cost black extender pigments.

Iron oxide pigments are low-cost, highopacity pigments that impart excellent properties to coatings. These pigments are easy to disperse in both water and oil, are alkali- and chemical-resistant, and hence can be utilized in both water-based and solvent-based systems. Although both the red and yellow iron oxides afford rather dull colors, they possess high tinting strength and are popular shading colors. In addition, iron oxide pigments are extremely durable, color-fast, and non-bleeding.

Lead chromate pigments are mainly used as corrosion inhibitive pigments. These pigments are of low opacity and can be combined with higher opacity pigments for use in primers. Zinc chromate, also an anti-corrosion pigment, is used in metal pretreatments such as "wash primer," which also promotes adhesion to aluminum and steel.

Iron blue, one of the most widely used pigments, combines low cost, good hiding power, and high tinting strength. This blue pigment exhibits good exterior durability and good heat resistance. Iron blue should not be used in water-based paints because it has very poor alkali resistance.

Ultramarine blue has poor opacity but high tinting strength, high heat resistance, and good alkali resistance. It is primarily used as a tinting color, and when added to whites makes them look whiter by giving them a bluish cast.

Chrome greens range in color from light yellow greens to dark blue greens. They are bright, durable, economical, and have high hiding power. They are used less frequently of late because of their lead content. Chrome greens are not suggested for use in aqueous systems because of their poor alkali resistance.

Chrome oxide green is a dull pigment having excellent exterior durability and very good chemical as well as alkali resistance. Thus, this pigment is suitable for use in either water-based or solvent-based systems. Chrome oxide green has moderate hiding power and is easy to emulsify, and its infrared reflectance properties make it valuable for use in camouflage paints.

Organic Pigments

Organic pigments are insoluble solids, and thus are differentiated from dyes, which are colored organic compounds that are soluble in the media in which they are used. Pigments require a binder in order to be used as colorants. Organic pigments that are insoluble in their pure form are referred to as "toners." Organic pigments that require precipitation by an inorganic base such as a metal to be insoluble are referred to as "lakes."

Organic pigments have greater decorative value than inorganic pigments; that is, organic pigments have a greater esthetic appeal due to higher chroma. The higher tinting strength of organic pigments partially offsets their higher cost. Organic pigments have a smaller particle size than inorganic pigments and a higher oil absorption. In addition, organic pigments are generally less toxic than inorganic pigments.

Hansa yellows are insoluble azo pigments having a wide range of shades. They are generally excellent in masstone lightfastness although not good enough in tint lightfastness to warrant outdoor use. Compared with the chrome yellows, they have less hiding power and greater tint strength. The Hansa yellows are most commonly used in water-based coatings and air-dried coatings. Benzidine yellows are no longer used because they have been shown to be carcinogenic. Nickel azo yellows are transparent pigments used in greenish yellow enamels and tinting enamels. These pigments have better permanency than the benzidine yellows although they tend toward darker transparent masstones. Nickel azo yellows impart excellent alkali resistance and fairly good bleed resistance. The tinting strength of these pigments increases upon aging in the formulation. Vat yellows are comprised of flavanthrones and anthrapyrimidines. Of these, the flavanthrones are the more red in hue. The vat yellows have very good lightfastness in weak tints, and they impart very good heat resistance and bleed resistance to the coatings. Tetrachloroiso-indolinone-based pigments, although relatively expensive, are of very high quality, imparting good lightfastness to the coatings.

Benzidine orange pigments are bright, strong pigments that exhibit good alkali, heat, and bleed resistance. There are both opaque and transparent versions of these pigments, and they are fairly lightfast although they are not recommended for exterior use. Benzidine orange is no longer used because of its potential carcinogenic nature. Dianisidine orange pigments are more bleedresistant than the benzidine oranges but less colorfast. Dinitroaniline orange pigments have better permanency than the two orange pigments described above. Although the dinitroaniline orange pigments bleed in most solvents, they are very stable in aqueous systems and exhibit superior lightfastness. Tetrachloroiso-indolinone, brominated anthanthrone, and perinone oranges are expensive, but very high-quality orange pigments, characterized by their excellent masstone and exterior lightfastness.

Copper phthalocyanine greens afford very clean, semi-transparent masstone, non-bleeding tints ranging from bluish green to yellowish green. These pigments have excellent heat and chemical resistance, are very lightfast in both masstone and tint, and have high tint strength. Pigment green B is an iron salt of nitro-sobetanaphthol that affords a semitransparent masstone having a dull yellowish tint. Although this material is not lightfast enough for exterior applications, it does have excellent alkali resistance and is suitable for use in aqueous pigment dispersions.

Phthalocyanine blues are very clean, semitransparent, non-bleeding pigments yielding high tint strength. As with the phthalocyanine greens, the blues have excellent heat and chemical resistance and are very lightfast in both masstone and tint. Also as with the greens, these blue pigments at high concentrations may bronze upon exterior exposure although this property is vehicle dependent. These characteristics apply to the crystallization and flocculation grades. These pigments are highly recommended for use in water-based systems. Indanthrone blues are expensive vat-type, non-bleeding, semitransparent masstone pigments having excellent masstone and tint lightfastness. Although they exhibit very good heat and chemical resistance, these pigments find limited use in coatings.

Carbazole dioxazine violets are semitransparent masstone pigments having high tint strength, excellent heat and bleed resistance, and good lightfastness. Isoviolanthrone violets afford clean tints exhibiting excellent chemical resistance and very good lightfastness, but they bleed badly in paint solvents.

Toluidine reds are insoluble azo pigments ranging from light to deep reds having excellent chroma and masstone lightfastness. Although these pigments hide very well and exhibit very good alkali resistance, they are not suitable for exterior use and bleed in paint solvents. Toluidine reds have been replaced to a large extent by a combination of molybdate orange and quinacridone or rubine toners. This combination affords better gloss and gloss retention, better bleed resistance, and less haze. Compared with the toluidine reds, this combination does sacrifice some cleanliness and brilliance of tint as well as some masstone lightfastness. Para reds are insoluble azo pigments that are deeper, less bright, and not as masstone-lightfast as the toluidine reds. They have good opacity and good alkali resistance, but they bleed very badly in paint solvents. These pigments can be utilized in inexpensive air-dried water-based systems. Quinacridones possess hues from gold to deep maroons and violets. These reds afford excellent bleed, heat, and alkali resistance, as well as very good tint and masstone lightfastness.

Carbon black is probably the most widely used pigment in the coatings industry. It is popular both as a straight shade and as a tinting color. It comes in a wide variety of particle sizes and surface treatments, the choice depending upon required jetness, tone, ease of dispersion, and of course, cost. Carbon black is characterized by excellent hiding power of high opacity, high tint strength, excellent durability, and very good lightfastness. Large-particle-size carbon blacks are chosen for their ease of dispersion and low resin demand, as well as for being less susceptible to floating and flooding than the finer particle sizes. They should, however, be used in applications that do not demand high levels of jetness. The finer particle sizes, higher surface area carbon blacks afford extremely jet, very durable, high-gloss coatings. These carbon blacks would be very suitable in automotive topcoat applications. The fine-particle-size carbon blacks also are used frequently in conductive coatings. Although all carbon blacks are good conductors, differences exist based on surface area, structure, and volatile content. In general, however, conductivity of the carbon black increases as the surface area and structure increase and as the volatile content of the carbon black decreases. The structure of the carbon black is defined as the agglomeration of the fine particle into treelike or “dendritic” shapes.

Pearlescent Pigments

Pearlescent pigments are comprised mainly of mica platelets coated with a thin layer of titanium dioxide. Co-precipitating titanium dioxide and other metal oxides onto the mica results in pigments exhibiting a colored luster effect. These pigments have a very high refractive index, indicating high hiding power or high opacity. In addition, because of their light-reflective properties, they impart some exterior durability to the coatings; and because of their small, thin, flat structure and high refractive index, they give the coated object a three-dimensional effect, that is, the appearance of depth. This is quite different from the apparent surface gloss due to the use of metallic pigments. Pearlescent pigments are available in a range of particle sizes and characteristics, and provide a variety of effects. These include a stain-luster appearance resulting from fine-particle-size pigments. These pigments are alkali- and acid-resistant and extremely durable when coated with rutile titanium dioxide. Pearlescent pigments are readily dispersible in water and afford stable metal-like water-based coatings.

Aluminum Pigments

Aluminum pigments generally are provided in paste form to eliminate the hazards of dust explosion or ignition during handling and storage, as occurs with the aluminum powders. In addition, an aluminum paste is very economical because of its high hiding power, excellent durability, and low specific density. Aluminum paint films provide excellent protection of substrates because the aluminum flakes align in many parallel layers. The layers of aluminum flake are staggered, thus yielding a tortuous path for the transmission of gases, chemicals, light, and moisture. Therefore, it takes longer for these materials to penetrate the film to get to the substrate. Aluminum paints made up of the leafing-type of aluminum paste reflect 75–80% of light and radiant heat, and are commonly used as exterior maintenance-type coatings or decorative coatings. In fact, the minute flakes in aluminum paste impart a smooth, brilliant appearance to the film. The nonleafing type of aluminum flake can be tinted with other various colored pigments to effect brilliant finishes such as those used in automotive metallic topcoats.

Lacquers

Some time before 200 BC., the Chinese used an exudate from the conifer *Rhus verniciflua* (later known as the varnish tree) as a coating. The process was further developed in Japan, where by AD. 400, the product was known as Japanese lacquer. These lacquers are a type of oleoresin that dry by oxidation in a damp atmosphere [14].

Lacquers are defined as non-convertible coatings; that is, lacquer film formation does not depend upon any chemical reaction. Film formation is dependent only upon the evaporation of solvents, and the formed film remains soluble in the parent solvent blend. Lacquers generally are spray-applied, taking advantage of the solvent's ability to dissolve the previously applied coat. Chlorinated rubbers and nitrocellulose esters are typical examples of lacquers. Nitrocellulose lacquers can form glossy coatings having excellent adhesion when other suitable resins are added that are compatible in all proportions. Two types of cellulose nitrates are available, a low-viscosity type that affords high solids, high-film-build coatings and a high-viscosity, low-solids type. The primary use of pigmented nitrocellulose lacquers has been in the car refinish and the wooden domestic furniture markets. The chlorinated rubber lacquers are used principally as protective coatings for steel and other vulnerable materials in particularly harsh environments.

Vinyl lacquers are used mainly where a high degree of chemical resistance is required; these lacquers are based on vinyl chlorides and vinyl acetates. Acrylic lacquers are based on methyl methacrylate and methyl acrylate polymers and copolymers. Other esters of acrylic and methacrylic acid also may be used to make non-convertible film formers. Judicious selection of these acrylic acid or methacrylic acid esters allows one to produce film formers with specifically designed properties such as hardness, flexibility, gloss, durability, heat, and chemical resistance. Acrylic lacquers, however, are not noted for their water resistance. The principal uses of acrylic-type lacquers are fluorescent and metallic paints, car refinish applications, clear lacquers and sealers for metals, and protective coatings for aircraft components and for vacuum-deposited metals, as well as uses in pigmented coatings for cabinets and appliances.

Vinyl emulsion paints that contain a non-convertible resin dispersed in water along with pigment can be considered as waterborne lacquers if film formation results from evaporation of the water followed by coalescence of the binder in which the pigment is dispersed. As environmental regulations become more strict, the role of traditional lacquers will essentially go the way of the dinosaur. The only allowable lacquers or non-convertible coatings eventually will be the water-based type. Indeed, as environmental regulations become more restrictive, the challenges facing the coatings chemist and the formulator become ever more demanding.

References

1. Detiveaux, S., Tsaur, A., "The Asia-Pacific Paint and Coatings Market," Coatings World, 16 (8), 28 (2011, August)
2. www.census.gov/manufacturing/cir/historical_data/ma325f/index.html
3. Reisch, M. S., "Paints and Coatings," Chemical and Engineering News, 23 (2003, November 3).
4. Huberfield, D., "VOC Regs Vary Widely," Industrial Finishing, 31 (1991, March).
5. Schulte, S., "EP Toxicity is Out, TCLP is In," Products Finishing, 28 (1990, June).
6. Greissel, M., "Coatings R&D 2001: Increasing Demands and Higher Budgets," Industrial Paint and Powder, 18 (2001, February).
7. Krajewski, J.J., "Photofabrication," J. Coating Technology, 62 (780), 83 (1990).
8. Chapman, D.P., "Aluminum pigment technology for waterborne and powder coatings in the 1990s," 18th Annual Waterborne and Higher Solids Symposium Proceedings, New Orleans, LA, 1991, February 6–8.
9. Akkerman, J. M., "Effect coatings: the future paint?" Prog. Org. Coatings, 17, 53–68 (1989).
10. Tullo, A.H., "Paints from plants," Chemical and Engineering News, 88(15), 16–29 (2010 April 12).
11. Hutchinson, G. H., "Advances in the technology of water-based inks and coatings for the printing and packaging industries," JOCCA, 68 (12), 307 (1985).
12. Martens, C. R., "Pigments, Paints, Varnishes, Lacquers and Printing Inks," in Riegel's Handbook of Industrial Chemistry, p. 787, J. Kent (Ed.), Van Nostrand Reinhold, New York, 1983.
13. Singer, E., "Raw Materials," Paint Handbook, pp. 3-23–3-32, McGraw-Hill, New York, 1981.
14. Skolnik, H., "History of Naval Stores inCoatings," in Organic Coatings: Their Origin and Development, p. 15, B. Seymour and H. F. Mark (Eds.), Elsevier, New York, 1990.

Industrial Biotechnology: Discovery to Delivery

30

Gopal K. Chotani, Timothy C. Dodge,
and Michael V. Arbige

Introduction

Fermentation products have penetrated almost every sector of our daily lives. They are used in ethical and generic drugs, clinical and home diagnostics, defense products, nutritional supplements, personal care products, food and animal feed ingredients, cleaning and textile processing, and in industrial applications such as fuel ethanol production. Even before knowing about the existence of microorganisms, for thousands of years ancient people routinely used them for making cheese, soy sauces, yogurt, and bread. Although humans have used fermentation as the method of choice for manufacturing for a long time, it is only now being recognized for its potential towards sustainable industrial development.

Since the discovery of fermentative activity of microorganisms in the eighteenth century and its proof by the French scientist Louis Pasteur, fermentative production of alcohols, amino acids, enzymes (biocatalysts), organic acids, vitamins, and natural polymers for food, feed, and other industrial applications has become well established. Subsequently, microbes have been used as production workers in industry. The production of bakers' yeast in deep aerated tanks was developed towards the end of the nineteenth century. The German scientist Buchner discovered that active proteins, called enzymes, are responsible for ethanol fermentation by yeast. During World War I, Chaim Weizmann used a microbe to convert maize mash into acetone, which was essential in the manufacture of the explosive cordite. In 1923, Pfizer opened the world's first successful plant for citric acid fermentation. The process involved fermentation utilizing the mold *Aspergillus niger* whereby sugar was transformed into citric acid. Other industrial chemicals produced by fermentation were found subsequently, and the processes were reduced to commercial practice. These processes included production of butanol, acetic acid, oxalic acid, gluconic acid, fumaric acid, and many more.

The serendipitous discovery of penicillin in 1928 by Alexander Fleming, while researching agents that could be used to combat bacterial infections, opened the whole new world of antibiotics. American pharmaceutical companies such as Merck, Pfizer, and Squibb were the first in mass-producing penicillin in the early 1940s. Initially, *Penicillium notatum* was surface cultured in flasks. Later, a new strain, *Penicillium chrysogenum*, was cultured in deep aerated tanks in the presence of corn steep liquor medium, and gave 200 times more penicillin than did Fleming's mold. Streptomycin was next, an antibiotic that was particularly effective against the causative organism of tuberculosis. Today, the list of these antibiotics is long and includes among many others, such important antibiotics as chloramphenicol, the tetracyclines, bacitracin, erythromycin, novobiocin, nystatin, and kanamycin.

Besides the booming antibiotics industry, fermentative syntheses of amino acids such as L-lysine and L-glutamic acid became billion dollar businesses. Despite the long history of microbial fermentation processes, understanding of the molecular basis of biological systems has developed starting from the discovery of the double helix DNA only decades ago. The practice of modern biotechnology started when Boyer and Cohen recombined DNA, and Boyer and Swanson founded Genentech, the first biotechnology company. Soon, it became clear that novel biotechnological manipulations could create new cell systems capable of producing new molecules and modifying existing products and processes. By the 1980s, a number of new biotechnology companies such as Amgen, Biogen, Cetus, Genencor, and others started to develop products for healthcare, agricultural, and industrial applications. The promise of biotechnology has been high in terms of delivering new products through partnerships with established pharmaceutical, agricultural, and chemical companies such as Eli Lilly, Roche, Johnson & Johnson, Monsanto, Shell, and others.

G.K. Chotani (✉) • T.C. Dodge • M.V. Arbige
DuPont Industrial Biosciences, 925 Page Mill Road,
Palo Alto, CA 94304, USA
e-mail: Gopal.Chotani@dupont.com

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_30, © Springer Science+Business Media New York 2012

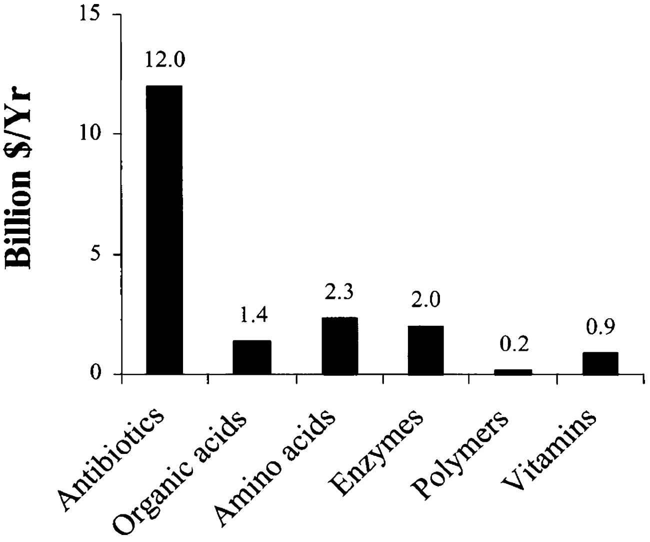


Fig. 30.1 Worldwide industrial biotechnology products market (except ethanol)

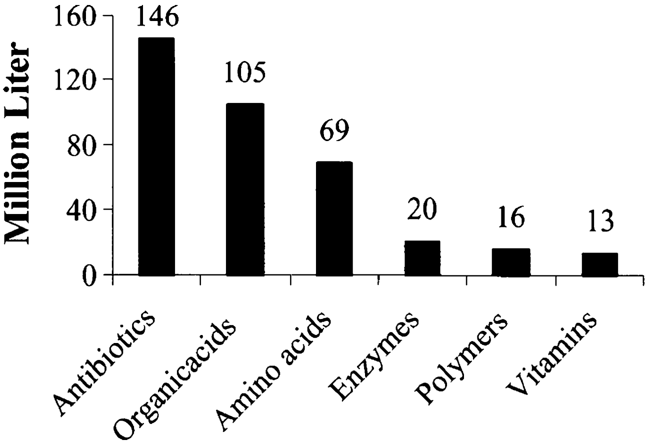


Fig. 30.2 Worldwide industrial biotechnology fermentation capacity distribution

By the end of the twentieth century, industrial applications of biotechnology started gaining momentum and proven laboratory techniques started to move into potentially huge markets. The field of industrial biotechnology (called "White Biotechnology" in Europe) was deemed the third wave of biotechnology. By definition, it is the application of scientific and engineering principles to the processing of materials by biological agents to provide goods and services. Industrial biotechnology continues to supply not only unique products, but is also considered synonymous with sustainable manufacturing processes. Some distinct advantages of industrial biotechnology include the use of renewable feedstocks (agricultural crop materials), generation of less industrial waste, lower cost for cleanup and disposal, and less pollution. The global market and fermentation capacity distribution (Figs. 30.1 and 30.2) for industrial biotechnology products, excluding ethanol, was estimated in 2002 at $17 billion. McKinsey & Company estimates that in the first decade of this century, biotechnology will affect up to 20% of the worldwide chemical market [1]. The major volumes of industrial biotechnology goods, such as alcohols, organic acids, amino acids, biopolymers, enzymes, antibiotics, vitamins, colorants, biopesticides, alkaloids, surfactants, and steroids, are expected to increase driven by innovative products, lower costs, renewable resources, and pollution reduction.

Discovery of Organisms and Molecules

Microbial Diversity

Microorganisms are chemically similar to higher plant and animal cells; they perform many of the same biochemical reactions. Generally, microorganisms exist as single cells, and they have much simpler nutrient requirements than higher life forms. Their requirements for growth usually are limited to air, a carbon source, generally in the form of sugars, a nitrogen source, and inorganic salts. Fermentation originally meant cellular activity/process without oxygen, but industrial biotechnology includes both aerobic and anaerobic fermentation processes. The nutrient source, complex (microbial, plant, or animal derived) or defined, not only provides carbon, nitrogen, and other elements (P, S, K, Mg, Ca, etc.) but also provides reducing and oxidizing agents [2]. Of course, in most of the cellular pathways, O_2 from air provides oxidation, releasing abundant metabolic energy from redox reactions. Depending on the physiological state of the cell, part of the metabolic energy escapes as heat that must be removed by the fermentor design. Sucrose, derived from sugarcane/beet and glucose derived from starch are often used as the carbon as well as energy sources for fermentation. Use of other carbon sources such as cellulose, hemicellulose, fats, etc. might increase in the near future. Proteins and fats are more reduced than saccharides and affect metabolism in terms of oxidation–reduction fluxes. The fluxes also depend on the endproduct(s) being more oxidized or reduced than the starting substrate(s).

Microbes occur in four main groups [3]:

- Bacteria and actinomycetales
- Viruses
- Fungi, including yeast
- Protozoa and algae

The chemical composition of microorganisms can be quite varied depending upon such factors as the composition of the growth medium, the age of the culture, and the cell growth rate. Table 30.1 lists the composition of a typical microbial cell [4]. All organisms contain the genetic information to produce a wide variety of enzymes and hence produce a great number of chemicals. Individual genes for

Table 30.1 Typical composition of microbial cell (about 70 % water)

Molecule	% Dry cell weight
DNA	3
RNA	20
Protein	55
Lipid	9
Liposaccharide	3
Peptidoglycan	3
Glycogen	3
Metabolites	3
Metal ions	1

specific enzymes can be organized into metabolic pathways that will define the biochemical metabolites that can be produced by the organism, which usually depends upon the environment in which the microorganism is growing. Inherent regulatory control processes allow cells to regulate their enzyme content in direct response to the environment. They prevent the formation of excess endproduct and superfluous enzymes. In this postgenomic era, the sequencing of a genome can be used for metabolic reconstruction of the microbial strain. Individual genes can be identified through homology to previously sequenced genes. Transcriptional analysis is another recent tool that allows the simultaneous measurement of all actively transcribed genes in a given organism at any given time. When combined with genomic information, attempts can be made to predict phenotypes or identify the organism.

For industrial processes, microbial strains with faulty regulation, altered permeability, enhanced enzymatic activities, or metabolic deficiencies are used to accumulate products. Such mutants have been changed so that their genetic mechanism is no longer sensitive to a particular controlling metabolite. Modern genetic engineering tools typically focus on creating such strains by making directed changes to the DNA of the organism.

Screening and Selection

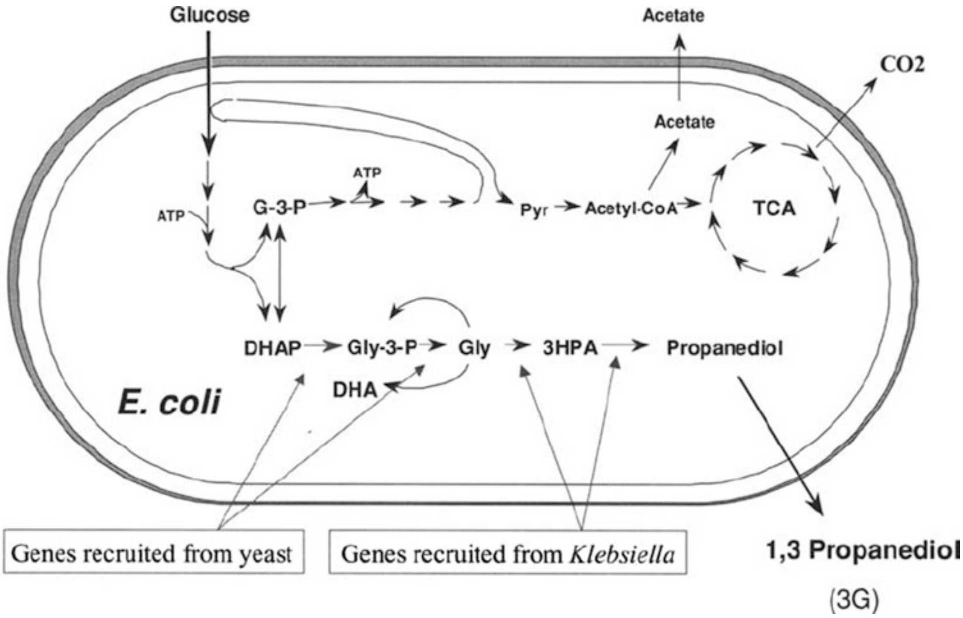
Product discovery means identifying the molecular target (enzyme, pathway, metabolite) of the intended application. A very early phase in the discovery of a microbial product involves searching for microorganism(s) that synthesize the product of interest [5, 6]. In the past this meant screening for living microorganisms. However, by applying molecular biology tools [7], today it is possible to screen for the gene (s) of interest without culturing the organisms. Of course, the rich diversity of microorganisms in nature is often the starting point for the screening/selection process. This process requires a well-designed growth medium, catalyst function, and assay. For example, a novel enzyme screening and selection strategy is based on an understanding of the real-world conditions under which the enzyme must function. Many enzymes are used in applications, for example, laundry, cooking, and solvent, that are far removed from the natural physiological conditions. Such applications expose enzymes to pH, temperature, and/or chemicals that might inactivate or inhibit them.

The classical screening method has its roots in antibiotic discovery. It involves collection of a variety of samples of soils and organic matter, and isolation of their microbial population. Being laborious at times (i.e., finding a needle in a haystack), random screening methods have been replaced in some cases by more efficient selection techniques. These techniques subject a microbial source sample to a selection pressure (i.e., pulling a needle from a haystack selectively with a magnet). For example, the growth conditions are selected such that only microorganisms expressing the desired enzyme/pathway/metabolite are able to survive. Therefore, selecting an environment (pH, T, growth nutrient) that is optimal for the growth of the desired microorganism increases the success of such a discovery. Very often, for the source of enzymes, industrial biotechnology companies look for microbes called extremophiles that live in extreme conditions, because they generally contain enzymes that perform optimally under extreme conditions.

In the last decade, tremendous progress has been made in sequencing DNA from various organisms. It is possible to extract DNA directly from the microbial community present in a sample taken from the environment. As a result, thousands of gene sequences have been deposited in public databases. This information can be used to screen for proteins/enzymes by searching for similarity at the gene level. Once a homologous gene is identified, it can be transferred to a laboratory host organism such as *Escherichia coli* or *Bacillus subtilis*. This allows for the expression of the identified gene product without making the source organism overproduce the molecule of interest. However, application of this technique is limited to microorganisms that have been isolated, characterized, and deposited in culture collections. Ironically, a very small number of microorganisms that exist in nature have been isolated and those characterized are even fewer. There may be many useful reactions in organisms yet to be discovered. Faster sequencing methods have started to exponentially expand the genomic database. Several hundred genomes are in the database and the tools of bioinformatics are vastly improving for the discovery of genes of interest.

According to J. Craig Venter Institute (http://www.venterinstitute.org), it should be possible to reduce the cost of sequencing a human individual DNA to less than $1,000. At this low cost and high speed of sequencing, we expect to learn more about human genes by comparing the human

Fig. 30.3 Rational biocatalyst design strategy for 1,3-propanediol



genome sequence to genome sequences of other species that have been thoroughly studied. Such diverse information has the potential to aid in the development of new products, processes, and applications in medicine, agriculture, energy, and the environment.

Cell Engineering

Classical strain development was traditionally dependent on the evolutionary process of random mutation. In a normal bacterial population, one mutant arises in about 10^6 cells. This low rate of spontaneous mutation is unsatisfactory for strain improvement. However, mutagens, such as *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine (NTG) and Ethyl Methane Sulfonate (EMS), are available that markedly increase the rate of mutation. Although it is relatively easy to produce mutants, strain improvement requires painstaking effort and ingenuity in devising screening tests. Knowledge of microbial physiology is essential in the development of many of these screens. Screening for resistance to metabolic analogues is probably the most important screening tool for devising means around normal metabolic control mechanisms. Although truly exemplary production organisms for important medical and industrial products were created using this process, the timeframe for many of these developments can be measured in decades.

In this era of genetic and metabolic engineering, the timeframe for strain development has been reduced, although it still is a major effort requiring substantial investment. As in classical strain development, knowledge of microbial physiology remains essential to the practice of metabolic engineering. Understanding the regulation of genes, in particular the interaction of promoters and regulatory proteins, is also essential. The five key strategies for metabolic engineering can be summarized as follows [8, 9].

1. Enhance committed step from central metabolism or branch point.
2. Eliminate transcriptional and allosteric regulation.
3. Identify and relieve rate-limiting step(s).
4. Prevent carbon and energy loss to competing pathways.
5. Production of foreign enzymes or metabolites.

Sophisticated metabolic models teamed with experimental data are finding widespread use in identifying strategic sites for metabolic engineering. For example, DesignPath® [10] is an approach based on the use of mathematical models and specifically generated algorithms that can be used to predict the right combination of genes that would enable a cell to overproduce a desired biomolecule, for example, 1,3-propanediol [11], and minimize the synthesis of undesired by-products, for example, acetic acid. Using such models, scientists have been able to combine in a single host (*E. coli*), a natural biological pathway present in yeast for the production of glycerol from dihydroxyacetone phosphate, and a pathway present in the bacterium *Klebsiella pneumoniae* to convert glycerol into 1,3-propanediol, to create a novel and unique process (for details see the section, “Delivery of Products”; Fig. 30.3).

Production strains, established through classical mutagenesis and screening to produce metabolites at a commercial level, were often started from strains known to overproduce at least some amount of the desired product. These strains had already overcome some of the natural regulation to keep from overproducing metabolites. However, this often led to strains that could overproduce more than one endproduct. To maximize production of the desired endproduct alone, strains were selected that could not overproduce a sideproduct. This resulted in the development of an auxotrophic strain, which required the particular sideproduct in order to grow. The need for specialized growth media often resulted in added expense and downstream difficulties as the required component(s) were added as complex mixtures from animal or plant extracts. By using enzymes from different organisms, it is now possible to change gene regulation, prevent carbon loss to competing pathways, and eliminate auxotrophies.

The ability to use heterologous enzymes in a production host is one of the hallmarks of modern metabolic engineering. This can be accomplished by specific directed changes to the genome or by generating a high level of diversity. The goal of the effort is to improve a specific phenotype of the strain. This could be to produce a new metabolite not found in the original host or change the energy balance in the host or some other critical attribute. In any case, deep knowledge of the physiology of the organism is used to develop the appropriate screen to find the desired phenotype.

Molecular Engineering

Despite the diversity of enzymes in nature, it is rare that a screened or selected protein/enzyme might have all of the desired characteristics for a targeted application. Therefore, it may be necessary to engineer the starting protein by modifying its structure through mutation (deletion, insertion, or substitution) of one or more amino acids. Protein engineering can be carried out by genetic modification or by chemical means (acetylation, amidation, oxidation, covalent attachment of ligands). Compared to the chemical approach, site-directed genetic manipulation is more predictable. The specific changes in the gene sequence are made through the use of recombinant DNA technology. Because the structure of a protein determines its function, rational protein engineering starts with studying the three-dimensional structure of the protein either by X-ray diffraction, NMR, or molecular modeling. Often such studies take time and require access to special expertise; therefore, alternative approaches such as random mutagenesis and directed evolution have emerged.

Random mutagenesis has been used extensively as a tool for increasing the genetic variability and improving microbial strains as described earlier. However, random mutagenesis also means less control over the number and types of mutations made. A large number of clones (called a library) are first generated and then screened or selected for the targeted properties. Therefore, random mutagenesis methods tend to use high throughput or rapid and efficient screening techniques.

Directed evolution, a term that covers a wide range of techniques designed to quickly improve strain performance, uses repetitive rounds of mutagenesis along with screening/selection, until a series of mutations accumulate to give an enhanced phenotype. Once again, the key factors for success are understanding of the target gene, the method for generating variability, selection, and screening of the host. The technique makes use of the fact that enzymes retain their relevant tertiary structures and thereby function. The evolution through a range of acceptable molecular structure variations leads to better enzymes with which the organism can tolerate specific environmental stress.

Several in vivo as well as in vitro methods of molecular evolution have been developed based on the recent advances in high-throughput screening, functional genomics, proteomics, and bioinformatics. Molecular evolution technologies provide efficient tools for creating DNA libraries (random, directed, recombinational) and processes of selection for a desired function or characteristic of a target molecule. For example, “DNA shuffling” is a variation of the molecular evolution approach and involves recombination of closely related DNA sequences, for example, random fragmentation and reassembly of the fragments into genes. The method is useful to combine two or more properties, which have evolved separately, into the target molecule. Molecular engineering, optimization, evolution, and other tools enable the design of completely novel biomolecules and proteins for applications in agriculture, medicine, and industry.

Development of a Production Process

Strain

Strain development begins with identification or isolation of an organism capable of producing the molecule of interest. Invariably, the initial strain is improved by mutagenesis, screening, selection, and genetic engineering to meet the production process and economic requirements. The process requirements include handling of the strain, fermentation conditions, product recovery, formulation, and application. Not surprisingly, strain and process development take place in an integrated manner. For example, processes that are sensitive to the cost of fermentation raw materials look for carbon-efficient strains. Likewise, processes that aim to achieve high purity in product need to have fermentations that require the use of minimal medium and strains that produce minimal sideproducts. New genetic techniques allow the rational engineering of a production strain. A rational approach means that the genes defining the production of a particular metabolite are specifically engineered through mutagenesis, deletion, or over-expression. Controlling expression of such genes determines the rate (productivity), yield (carbon conversion efficiency), and titer (concentration) of the product. The result is a production organism that is optimally tailored for the process and the product.

Some bacterial and fungal strains are capable of differentiation (sporulation, filamentation). This property can affect product formation and the physical properties of the fermentation broth. Some enzymes are synthesized as secondary products and their production does not appear to be growth associated. Production strains are distinguished on the basis of their fermentation behavior such as viscosity or recoverability. They can also be distinguished by patent or proprietary status. The complexity of product synthesis ranges from a relatively simple and well-understood induction and repression to very complex regulation mechanisms. Process development must thus deal with such complexity either by changing the genetic make-up or by optimizing process conditions.

Because classical approaches to strain development are labor intensive and offer limited knowledge, new strain development methods now start with a production host that is capable of rapid genetic manipulation and efficient product synthesis. Generic hosts and techniques allow for the construction of multiple strains for multiple products utilizing a baseline fermentation and recovery process. The so-called toolbox strategy of building strains applies to production of enzymes/proteins as well as biochemicals. Generic host-based strains (denoted GRAS, Generally Recognized As Safe) have been proven for safe use in industrial processes and have been approved by regulatory agencies such as the FDA, EPA, and USDA.

Once a desired production strain is developed, it can be maintained as a stock culture. In maintaining stock cultures, genetic changes must be minimized. This is best achieved by preventing nuclear divisions as most mutations occur as errors during DNA replication. The method of choice is to store cells or spores (if these are produced) in sealed ampoules at very low temperatures (≤130°C) in liquid nitrogen. This method offers the great advantage that the culture can be stored almost indefinitely, thawed, and used immediately as an inoculum without loss of viability or diminution in metabolic rate. Cultures kept at −20 to −60°C are satisfactory but less active than those kept in liquid nitrogen. Although storage at 0–4°C allows some growth, this is better than storage at room temperature. Lyophilization (freeze-drying) is widely used and is very convenient because freeze-dried cultures retain viability without any genetic changes for years when stored at room temperature. It may be noted that all of these methods are, in effect, techniques to immobilize intracellular water and yet retain viability.

Fermentation Process

Analysis of many industrial fermentation processes shows that they are common reactions from a chemical, as well as a physical, viewpoint. Fermentation processes can be classified by the reaction mechanisms involved in converting the raw materials into products. These include reductions, simple and complex oxidations, substrate conversions, transformations, polymerizations, hydrolyses, complex biosyntheses, and the formation of cells.

Fermentation processes, except for sterilization, have in common many of the familiar chemical engineering unit operations. For example, aerobic fermentations involve the “mixing” of three heterogeneous phases: microorganisms, medium, and air. Other unit operations include “mass transfer” of oxygen from the air to the organisms and “heat transfer” from the fermentation medium.

Analysis of fermentations by the unit operation technique has added greatly to the understanding of their behavior. Of the operations auxiliary to those in the fermentor, engineers have made a major contribution towards the design of equipment to provide large volumes of sterile medium and air. Close cooperation between biologists and engineers has resulted in devising logical methods for screening large numbers of strains and translating the results of shake-flask and pilot-plant experiments to production vessels. The scale-up of fermentations, in some instances, is still empirical although sensitive oxygen probes and gas analysis techniques now available have enabled a rational approach to scaling-up aerated, Newtonian as well as non-Newtonian, fermentation processes.

Sterilization

In all fermentation processes, it is necessary to have contamination-free fermentation media and seed cultures. Liquid sterilization of the fermentation medium is conducted by two means, physical removal or inactivation by heat [12, 13]. Contaminating microorganisms can be removed from fluids by filtration. With improvements in membrane technology, sterile filtration is finding wider use, but can only be used with completely soluble media. The contaminants can also be destroyed by heat. Heat sterilization of media is the most common method used for sterilization of liquid media. This can be accomplished in either a batch or a continuous fashion.

Interest in continuous methods for sterilizing media is increasing, but for the successful operation of a continuous sterilizer, foaming of the media must be carefully controlled and the viscosity of the media must be relatively low. The advantages of continuous sterilization of media are as follows.

1. Increase in productivity because the short period of exposure to heat minimizes damage to media constituents.

2. Better control of quality.
3. Leveling of the demand for process steam.
4. Suitability for automatic control.

Design and operation of equipment for heat sterilization of media are based on the concept of thermal death of microorganisms. Consequently, an understanding of the kinetics of the death of microorganisms is important to the rational design of sterilizers.

The destruction of microorganisms by heat implies loss of viability, not necessarily destruction in the physical sense. The destruction of organisms by heat at a specific temperature follows a first order rate of reaction [14, 15]:

$$\frac{dN}{dt} = -KN = -(Ae^{-E/RT})N, \tag{30.1}$$

where K is the reaction rate constant with units of time^{-1}, N is the number of viable organisms/unit volume, t is the time, T is the absolute temperature, E is the energy of activation for death, R is the gas law constant, and A is the Arrhenius constant.

This equation can be integrated to give the design equation:

$$\ln\frac{N_0}{N_f} = A\int_0^{t_s} e^{-E/RT}\,dt, \tag{30.2}$$

where N_o is the number of contaminating organisms in the total fermentation medium to be sterilized, N_f is the level of contamination that must be achieved to produce the desired degree of apparent sterility, and t_s is the sterilization time.

In estimating the sterilization time for the medium, one must define the contamination, the desired degree of apparent sterility, and the time–temperature profile of the medium; that is, $T = f(t)$. For typical bacterial spore contaminants, the constants used in most designs have the following values.

$E = 68{,}700$ cal/g mol
$R = 1.987$ cal/g mol, K
$A = 4e^{+87.82}$, min^{-1}

Aerobic fermentation processes also require a continuous supply of large quantities of air, typically on the order of one volume of air per volume of liquid per minute, VVM. Sterilization of this air is mandatory in almost all fermentations. Absolute filter cartridges of polymeric membranes are now used almost exclusively in the fermentation industry. Relatively small units have replaced the large depth filters used in the past. Still, water and particulates pose a major problem for filters thus requiring the use of prefilters and traps to remove these contaminants before they reach the absolute filter. Parallel installation of the filters prevents a total shutdown of the fermentation process in the event of filter clogging.

Microbial Kinetics

Microbial kinetics [16, 17] can be separated in four distinct levels: at the molecular or enzyme, the macromolecular or cell component, cellular, and population level. Because each level has its own unique characteristics, different kinetic treatments are needed. Moreover, the environment in which these reactions take place also affects the kinetics. For example, reactions at the molecular/enzyme level involve enzyme-catalyzed reactions. When these reactions occur in solution, their kinetic behavior is similar to that of homogeneous catalyzed chemical reactions as described in Chap. 31. However, when enzymes are attached to inert solid supports or contained within a solid cell structure, their kinetics is similar to that of heterogeneous catalyzed chemical reactions.

As discussed in regard to microbial diversity, all microorganisms have basic requirements, those being water, a source of energy, carbon, nitrogen, salts, and trace metals, and possibly growth factors. However, the media used to isolate and screen production hosts are not necessarily those used in production fermentors. Statistical methods (Plackett–Burman, Box–Benkhen) are used for media screening and optimization. The media and conditions used may change from shake flask through the fermentor stages. The objective in developing a production medium is to maximize productivity and product quality, minimize sideproducts, and meet economics. Generally, large-scale fermentation media are made up of complex natural materials, supplemented with inorganic/organic salts [18, 19]. Development of a production medium combines understanding the physiology of growth and product formation developed through plate, shake flask, and fermentor studies.

For example, for deciding on the type of carbon source, the phenomenon of catabolite repression must be considered. Similarly, regulation of nitrogen and sometimes phosphorus metabolism are important factors to consider. Sometimes high concentrations of salts and free amino acids necessary for high cell density are inhibitory to product formation. One area often overlooked is the elimination of unwanted impurities, derived either from the raw material or produced via metabolism. These impurities represent not only a waste of the carbon source, but their accumulation at high levels may lead to inhibition of growth or even cell death.

At a population level, a material balance on cells growing by binary fission can be described by:

$$\frac{dX}{dt} = \mu \cdot X, \tag{30.3}$$

where X is the cell concentration and μ is the specific growth rate of the microbial culture.

Monod [20] suggested that cells growing on a limiting substrate were controlled by a limiting enzymatic reaction. Enzymatic reactions had been modeled using Michaelis–Menten kinetics and Monod applied the same methodology to microbial growth. He described the specific growth rate by:

$$\mu = \mu_{\max} \cdot \frac{S}{K_s + S}, \tag{30.4}$$

where $\mu_{\max}$ is the maximum growth rate of the particular strain being studied, s is the concentration of the limiting substrate, and K_s is the so-called affinity constant for the limiting substrate and is equal to the concentration of the substrate resulting in one-half the maximum growth rate. Combining the above equations results in the cell balance equation:

$$\frac{dX}{dt} = \mu_{\max} \cdot \frac{S \cdot X}{K_s + S}. \tag{30.5}$$

A similar material balance on the limiting substrate requires knowledge of the relationship between cell growth and substrate utilization. The simplest relationship would assume that a fixed amount of cells could be produced from a given amount of substrate, or the yield of biomass on substrate:

$$Y = \frac{dX/dt}{-(dS/dt)} = -\frac{dX}{dS}. \tag{30.6}$$

The linearity between cell growth and substrate consumption does not always hold true, especially at low specific growth rate. Pirt [3] first suggested the concept of maintenance. He described maintenance as the substrate required to generate energy for cell functions independent of growth rate. Examples of cell functions include maintenance of ion gradients and the turnover of macromolecules. The resulting material balance on the limiting substrate is:

$$\frac{dS}{dt} = -\left(\frac{\mu}{Y^{\max}} + m\right)X, \tag{30.7}$$

where $Y^{\max}$ is the maximum growth yield and m is the maintenance coefficient. The term in parentheses is often called the specific substrate consumption rate, q_s.

Because substrate can also be consumed to make a microbial product, the concept of specific rate can also be used for product formation,

$$\frac{dS}{dt} = -(q_s + q_p)X, \tag{30.8}$$

where q_p is the specific product formation rate. The simplest expressions relate product formation to either cell growth or substrate consumption:

$$q_p = Y_{p/s} \cdot q_s \tag{30.9a}$$

$$q_p = Y_{p/x} \cdot \mu \tag{30.9b}$$

The last equation is an example of growth-associated product formation, often observed for primary metabolites. However, for many secondary metabolites, much of the product formation takes place when growth rate slows. This is described as nongrowth-rate associated kinetics. Leudeking and Piret [21] first described a mixture of the two by the following equation, which takes into account both growth-rate associated and nongrowth-rate associated forms of production,

$$q_p = \alpha \cdot \mu + \beta. \tag{30.10}$$

Ideal Types of Fermentors

There are a large number of different types of fermentation processes that are used commercially, which are selected based on several different factors [22–24]. Depending on the strain to be used, the fermentation could be aerobic or anaerobic, and the desired product could be either the biomass itself or a metabolite or polymer produced by the biomass. The kinetics of product formation, whether growth associated or nongrowth associated, also influences the process. Often procedures downstream of the fermentation unit operation have a major control of the overall process and determine how the fermentation is conducted.

Although there are a multitude of possible different fermentation process designs, most of them can be lumped under the following four [23], shown in Fig. 30.4a–d,

- Batch
- Fed-batch
- Continuous
- Continuous with recycle

The batch process is the simplest fermentation to perform. It is a closed system with constant volume. The only addition after the start of the fermentation is air, if it is needed, resulting in a constant volume. All components are in a constant state of change as the substrate is consumed and biomass and products are produced. Figure 30.4a shows the typical batch fermentor characteristic: nutrient is present from the start of the process; nothing is fed and no steady state can be reached. In other words, characteristic volume of culture (V) is constant, denoted by $V = c$. But specific growth rate (μ) and cell mass (X) are not constant. Batch

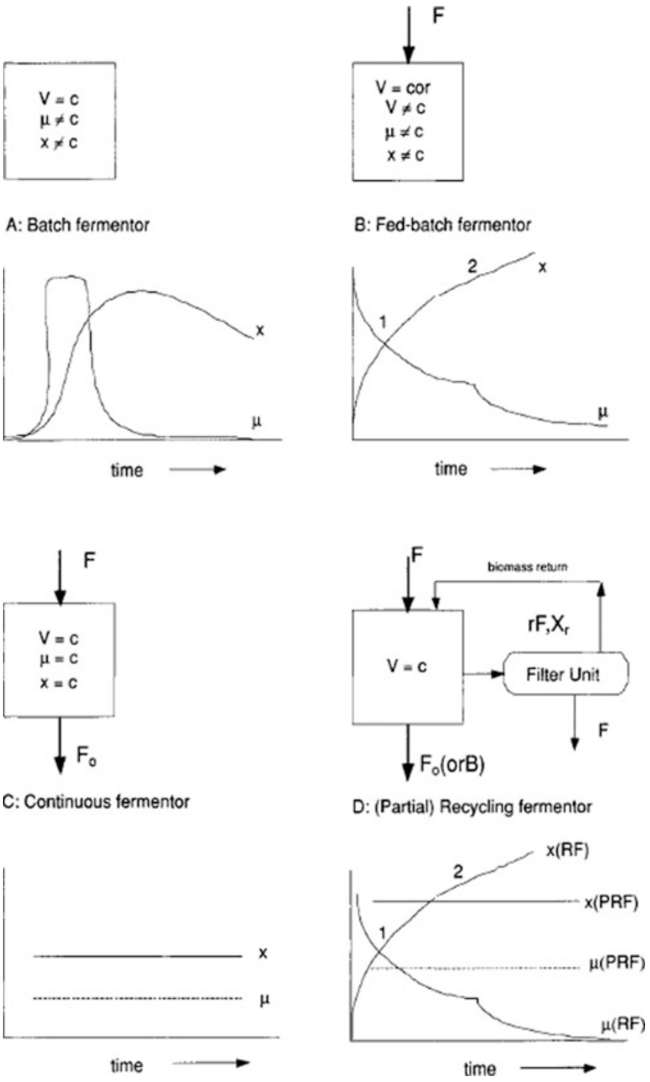


Fig. 30.4 Characteristics of major fermentor types

fermentations are often used for highly mutated strains that are developed by classical mutagenesis and selection. Such strains are easily taken over by faster-growing, less-productive strains in prolonged cultures. Batch processes are still quite common in the antibiotic, organic acid, and ethanol production industries.

In the fed-batch fermentation, nutrients can be added making it an open system for substrates, but still a closed system for biomass and biomass-derived products. As shown in Fig. 30.4b, in this type of process, the volume is not constant. The flow rate of nutrients can vary during the course of the fermentation. Fed-batch operation is very commonly used in industrial processes for the production of baker's yeast, enzymes, amino acids, and many other metabolites.

In the continuous system, as shown in Fig. 30.4c, by constantly removing contents from the fermentor, the process is now open with respect to all components. Typically, the addition of feed and the removal of broth are equal, resulting in a constant volume. Under these conditions, a steady state is achieved wherein all parameters become constant. Such a system is widely used for physiological studies. However, industrial uses are typically limited to the production of biomass (single cell protein) for food or feed, and waste treatment plants.

The final system, shown in Fig. 30.4d, is the continuous system with a partial (PRF) or complete (RF) cell recycle. It is similar to the continuous system, but cells are returned to the fermentor by means of a biomass separation device. Tangential flow-filtration units, centrifuges, and settling tanks have all been used for biomass separation [25]. In the partial cell recycle fermentor, a steady state is achieved as in the continuous system. This process is typically used to increase the productivity of the system and is used commonly in wastewater treatment and ethanol production type applications.

The differential equations shown in the Microbial Kinetics section are used to describe the batch fermentation mode. These can be solved to describe the increase in biomass and product concentrations and the decrease in substrate concentration. The fed-batch system takes into account the addition of substrate during the fermentation. The mass balance on substrate is described by

$$\frac{\mathrm{d}}{\mathrm{d}t}(V \cdot S) = F(t) \cdot S_{\mathrm{f}} - V(q_{\mathrm{s}} + q_{\mathrm{p}})X. \qquad (30.11)$$

A total mass balance on the reactor is:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\rho \cdot V) = \rho_{\mathrm{F}} \cdot F(t), \qquad (30.12)$$

where ρ is the density of the fermentation broth, ρ_{F} is the density of the feed solution with S_{f} as the substrate concentration, and $F(t)$ describes the rate of addition of feed as a function of fermentation time. As in the batch fermentation, these equations can be solved to show the concentration–time profiles of substrate, biomass, and product. The last equation assumes a negligible change in the reactor mass due to air addition (humidity, oxygen) and gas removal (carbon dioxide, water). Such assumptions, in practice, must be checked.

For a typical single-stage continuous fermentor at steady state, a biomass material balance around the reactor yields the following relationship,

$$\frac{\mathrm{d}X}{\mathrm{d}t} = 0 = -\frac{F}{V}X + \mu X, \qquad (30.13)$$

from which:

$$\mu = \frac{F}{V} = D. \qquad (30.14)$$

This means that the dilution rate in the fermentor sets the growth rate of the biomass, and a change in the dilution rate will cause a change in the growth rate.

Because most single-stage continuous fermentors are used to produce biomass, they are usually operated to optimize biomass productivity. The unit volume biomass productivity of such a reactor is defined as DX. This unit volume productivity can be expressed as

$$DX = DY\left(S_0 - \frac{K_s D}{\mu_{\max} - D}\right). \qquad (30.15)$$

By taking the first derivative of the productivity expression with respect to the dilution rate and setting it equal to zero, the dilution rate of maximum productivity, D_m, can be found as

$$D_m = \mu_{\max}\left(1 - \sqrt{\frac{K_s}{K_s + S_0}}\right) \qquad (30.16)$$

and the maximum productivity, $D_m X$, as

$$D_m X = \mu_{\max}\left(1 - \sqrt{\frac{K_s}{K_s + S_0}}\right) \cdot Y \cdot \left(S_0 - \frac{K_s D_m}{\mu_{\max} - D_m}\right). \qquad (30.17)$$

Biomass recycle as sketched in Fig. 30.4d is frequently used in fermentors as a way of increasing the biomass productivity. The increased biomass productivity obtained with a recycle fermentor is a function of the recycle ratio, r, and the cell concentration factor, $C = X_r/X$, achieved in the concentrator. Equations expressing the recycle system behavior are derived from material balances around the reactor. For the cell biomass balance at steady state:

$$\frac{dX}{dt} = 0 = \frac{F}{V}(0) + \frac{rF}{V}X_r - \frac{F}{V}(1 + r)X + \mu X, \qquad (30.18)$$

from which:

$$\mu = D\left(1 + r - r\frac{X_r}{X}\right) = D(1 + r - rC) \qquad (30.19)$$

Because $(1 + r - rC) < 1$, it is possible to operate the system at dilution rates greater than the maximum growth rate. It is this stability imparted by the cell recycle fermentor system that makes it useful, especially in waste treatment applications.

Oxygen Transfer Considerations

In aerobic fermentations, oxygen is a basic substrate that must be supplied for growth. As described earlier for other substrates, the relationship between oxygen concentration and growth is of a Michaelis–Menton type and can be described by a specific rate relationship (see (30.5)). The K-value for oxygen typically ranges from 0.5 to 2.0 ppm for well-dispersed bacteria, yeast, and fungi growing at 20–30°C. For growth at temperatures greater than 30°C; the specific oxygen uptake increases only slightly with increasing oxygen concentration [14].

Under steady-state conditions, the oxygen transfer rate must be equal to the oxygen uptake rate:

$$Q_{O_2}X = k_L a(C^* - C)_{\text{mean}}, \qquad (30.20)$$

where C^* is the concentration of oxygen in the liquid that would be in equilibrium with the gas-bubble concentration. k_L is the oxygen mass transfer coefficient and a is the bubble interfacial area. In small, well-mixed systems, the gas-bubble concentration can be assumed to be equal to that in the gas escaping from the fermentor. However, in large reactors, the log-mean average between the inlet and outlet air concentrations is more appropriate. The oxygen mass transfer coefficient and interfacial area are typically lumped together as $k_L a$. It is not practical to try to separate these two terms.

Utilizing this relationship, the lumped oxygen mass transfer coefficient can be estimated as

$$k_L a = \frac{Q_{O_2}X}{(C^* - C)_{\text{mean}}}. \qquad (30.21)$$

Scale-Up/Down and Control

There are usually problems scaling up new fermentations as well as with translation of process-improvement data for well-established fermentations from laboratory operations to existing plant equipment [26]. In general, fermentations are scaled up on the basis of achieving similar oxygen transfer capabilities in the plant equipment that proved to be optimal at the bench scale.

The oxygen transfer capability required for processes can vary greatly. Organisms with a low specific growth rate may require as little as 25 mmol/L/h whereas high specific growth rate cultures could require ten times that amount. The range of airflow required can vary from as little as 0.1 VVM to >1 VVM. Likewise, mechanical power requirements can vary from less than 0.5 kW/m^3 to >5 kW/m^3.

There are various correlations between $k_L a$ and power inputs. Some design engineers prefer to scale-up/down on the following basis,

$$k_L a = k\left(\frac{P_m}{V}\right)^{\alpha}(v_s)^{\beta}, \qquad (30.22)$$

where P_m is the motor power and v_s is the superficial gas velocity. The coefficients α and β are typically in the range of 0.5 for large-scale plant equipment. Mixing is one of the most critical factors of large-scale fermentors and among agitator designs, Rushton turbine type impeller has most often been used and studied [27].

Unaerated mechanical power input can be estimated from:

$$P_u \propto \rho N^3 D^5, \tag{30.23}$$

where N is the rotation speed of the impeller shaft (rpm) and D is the diameter of the impeller.

For scale-up, however, it is usually the removal of heat that causes design problems. With the previously mentioned mechanical agitation power inputs, up to 5 kW/m^3 of energy are needed to remove mechanical heat. The peak metabolic heat load for aerobic fermentation of glucose at ~250 mmol O_2/L h uptake rate is generally greater than 35 kW/m^3. Even with a fermentation temperature as high as 37°C and cooling water temperature as low as 18°C, it is difficult to remove the heat in large fermentors without external heat exchange or extensive cooling coils in the fermentor. Internal spiral cooling coils can be undesirable because they could interfere with the mixing patterns. Thus, vertical coils that also act as mixing baffles are finding more widespread use. Chilling of the cooling water can also be used to increase heat removal.

Once a plant is built, the conditions of agitation, aeration, mass (oxygen) transfer, and heat transfer become set. Therefore, those environmental conditions achievable in plant scale equipment should be scaled down to the pilot plant and laboratory-type equipment (shake flask, micro-reactor) to ensure that the earlier studies are carried out under conditions that can be duplicated.

With the rise of the genomic and postgenomic era, high-throughput gene sequencing, proteomics, metabolomics, and systems biology have created a wealth of biological data. This has given rise to the need for higher throughput fermentation technology to screen libraries of natural and in vitro generated compounds. Initially, screens of new organisms and metabolic pathways were performed in shake flasks and throughout the last decade increasingly in microtiter plates. Although these methods are straightforward and well proven, they are laborious and do not allow complex manipulation of growth conditions. Furthermore, they do not provide a full complement of data for the knowledge-based selections of organisms or desired physiological properties.

Within the past 5 years, novel technologies have been developed that allow more directed and controlled experiments to screen organisms, fermentation, and media conditions at small scales. A number of academic and private institutions are developing machines that are capable of growing cells at the microliter and sometimes nanoliter scale. They are able to supply aeration, oxygen transfer, pH control, and online data output similar to a pilot-scale fermentor. These devices require minute amounts of cells and media and thus offer time, cost, and environmental advantage and ultimately enhance productivity. Very small scales have been described for cell growth and cell-based assays in microfluidic devices, such as the compact disk (CD) format technology by Gyros Microlabs (http://www.gyrosmicro.com). Intermediate scale microfabricated devices include the bioprocessors platform (PCT WO 2002/083852), the Micro Reactor device (http://www.gener8.com), and a number of devices by academic laboratories, most prominently, from the University of Maryland at Baltimore (http://www.umbc.edu/cbe/rao.html). For instance, the Micro Reactor device is based on a 24-well plate format with several milliliters fermentation capacity in each well. Airflow, pH, temperature, and agitation can be controlled electronically. The machine provides online readouts for pH, dissolved oxygen, and other parameters. Another larger intermediate scale machine offered by DASGIP AG, Fedbatch-Pro®, is a modular system that delivers a high degree of flexibility to perform a number of microbiological applications (http://www.dasgip.com). It remains to be seen if these machines will eventually be capable of replacing the classical scale-up versions of stainless steel fermentors.

Instrumentation and Control

In successfully scaling up/down any fermentation, knowledge of the regulatory mechanisms of metabolic pathways that synthesize the desired product and the instrumentation to detect the regulatory metabolites is necessary. To optimally run a fermentation process, it is essential to perform the initial fermentation research on fully monitored environmental systems, then correlate the environmental observations with existing knowledge of cellular control mechanisms, and finally reproduce the desired environmental control conditions through continuous computer monitoring, analysis, and feedback control of the fermentation environment.

As a result of advances made in sensor development, today more so than in the past, it is possible to rely on environmental control in order to gain economical fermentation results. Until recently fermentation control was limited to that of temperature, pH, and aeration. With the development of numerous sensors and inexpensive computing systems, the engineer can think in terms of sophisticated control systems for fermentation processes. Figure 30.5 shows how a highly instrumented fermentor is designed to secure basic information on almost all the parameters of the fermentation process. For advanced information, both

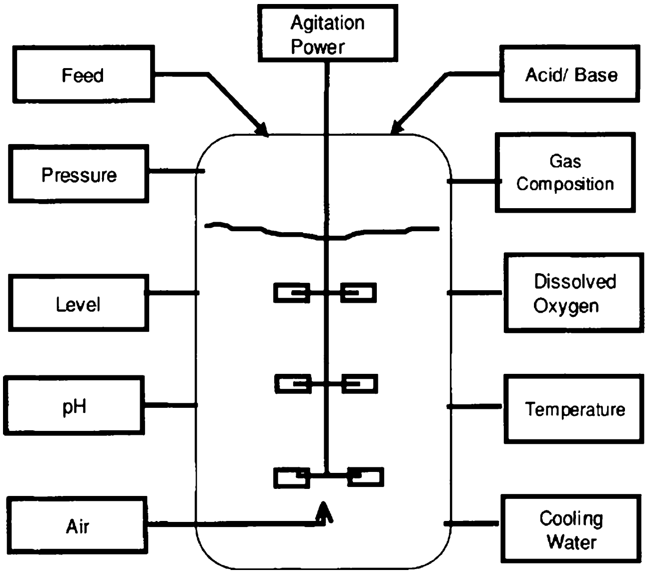


Fig. 30.5 Highly instrumented industrial fermentor

offline intracellular (e.g., messenger RNA and protein arrays, metabolites) and extracellular (e.g., metabolites, proteins/peptides by LC/GC) analyses methods are readily available.

One of the most important sensors needed is one that reliably monitors cell density. An IR fiber-optic cell density probe has been used for this because it can directly monitor cell growth (without dilution) in high-cell-density bacterial fermentations. The ability to do an online sample filtration through the use of hollow fibers or rotating filters has made possible continuous, online measurement of glucose, lactate, and other metabolites. However, glucose, nitrogen substrate, and phosphate sensors that can withstand repeated system sterilization are still needed.

Repeated sampling for measurement during a fermentation process can be tedious and thus indirect measurement via computers appears to be a viable alternative. Combined with other information obtained from sensors, these measurements make possible the calculation of several fermentation parameters, as given in Table 30.2. Computer simulation can also be used to indirectly measure a given component based on a mass balance equation of that component for the fermentation process. Besides these uses, the computer has applications in fermentation processes for continuous automated monitoring and feedback control and dynamic optimization of the process [16, 28].

Recovery of Fermentation Products

The isolation and purification of fermentation products is often collectively referred to as downstream processing. The early part of the separation of a bioproduct is the primary recovery process, whereas the elements further downstream may include purification, concentration, and formulation. The overall goal of downstream processing and formulation is to recover the product of interest cost effectively at high yield, purity, and concentration, and in a form that is stable, safe, and easy to use in a target application.

Table 30.2 Gateway measurements for fermentor monitoring and control

Parameter	Calculation
pH	Acid/base uptake or formation
Air flow rate, in and out O_2 concentration	O_2 uptake rate
Air flow rate, in and out CO_2 concentration	CO_2 production rate
O_2 uptake rate, CO_2 production rate	Respiratory quotient, specific metabolic rate
Power input, air velocity	O_2 transfer rate
Integrated CO_2 produced	Cumulative metabolic activity

Fermentation products include the cells themselves, solvents and chemicals, organic and amino acids, antibiotics, polysaccharides, lipids, RNA and DNA, vaccines, bulk and fine diagnostic and therapeutic proteins, food and feed ingredients, and enzymes. These biomolecules can differ substantially in nature and require a large variety of methods to separate and purify them. They are often produced in low concentrations in complex media together with many other components.

In the development of manufacturing processes for the production of biological molecules, much emphasis is placed on fermentor design and scale-up, thus one might assume that the recovery processes of fermentation products are rather straightforward and relatively simple. Nothing could be further from the truth. A point in case is an example of an antibiotic production plant. The investment for the recovery facilities is claimed to be about four times greater than that for the fermentor vessels and their auxiliary equipment. In organic and amino acid fermentations, as much as 60% of the fixed costs of fermentation plants is attributable to the recovery unit operations.

Figure 30.6 shows a typical recovery process for antibiotics, and a schematic overview of a typical downstream process in an enzyme plant is given in Fig. 30.7. From these diagrams it is apparent that most recovery processes involve combinations of the following procedures.

- Fermentor harvest
- Mechanical separations of cells from fermentation broth (clarification)
- Optional disruption of cells
- Extraction of the compound of interest
- Preliminary fractionation procedures
- High-resolution separation steps
- Concentration
- Formulation
- Optional drying

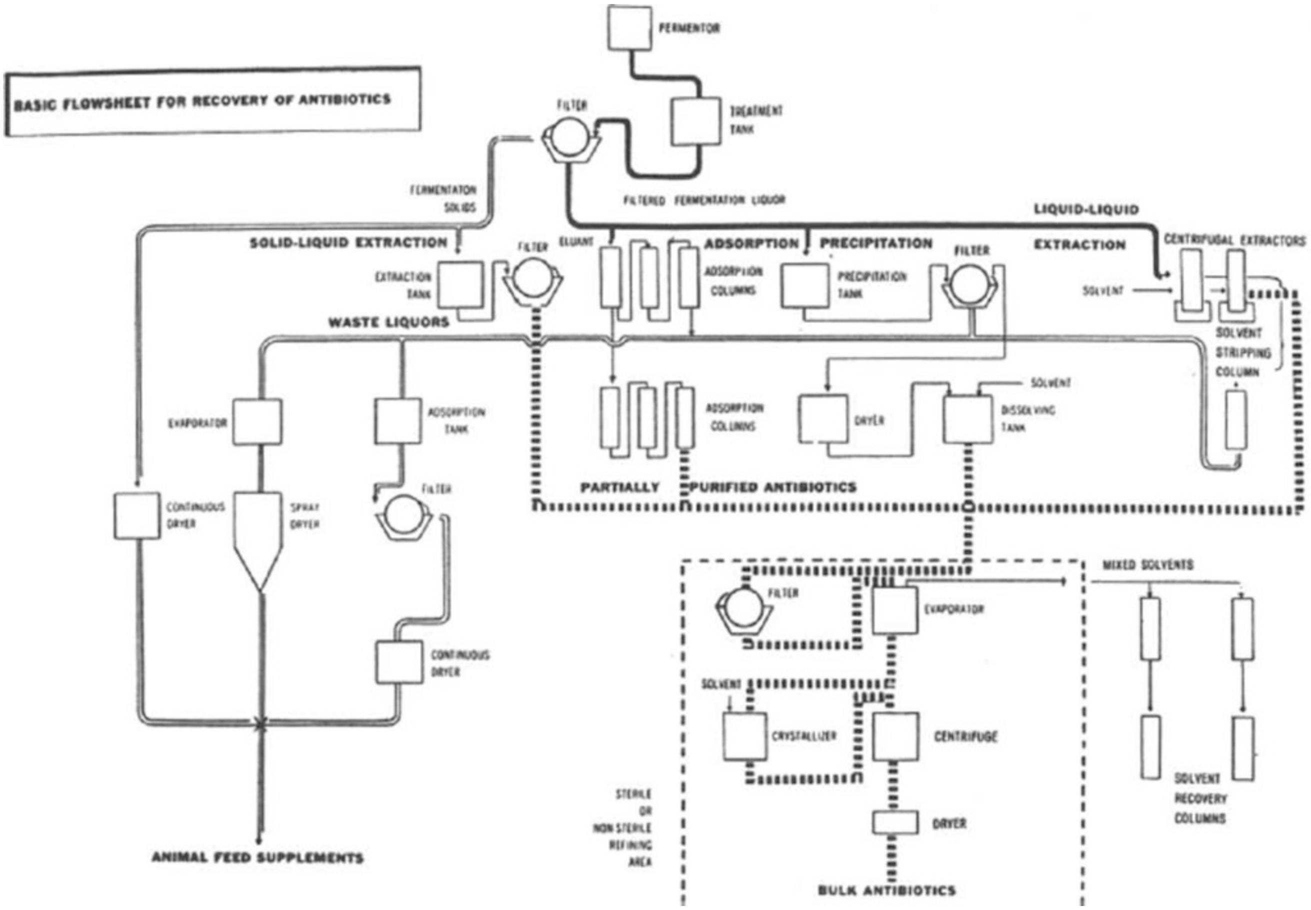


Fig. 30.6 Basic flow sheet for the recovery of antibiotics (Ind Eng Chem 49:1494, 1957. Copyright American Chemical Society and used with the permission of the copyright owner)

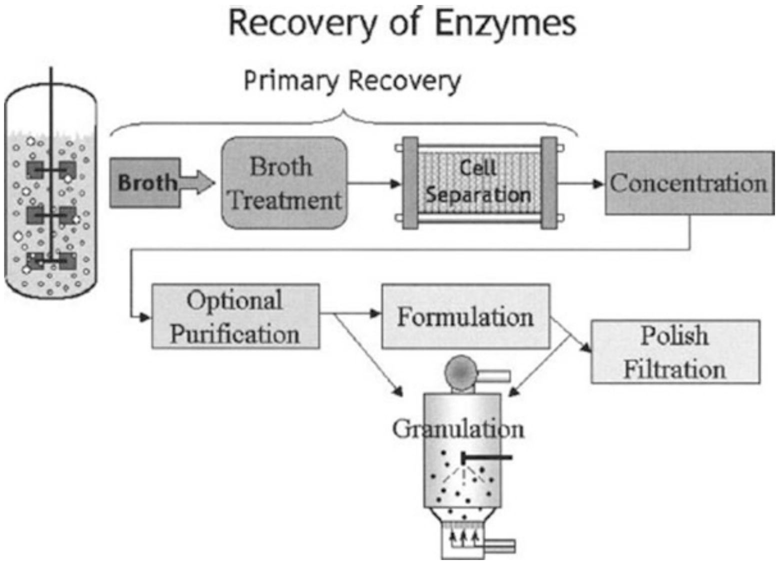


Fig. 30.7 Typical recovery process for industrial enzymes

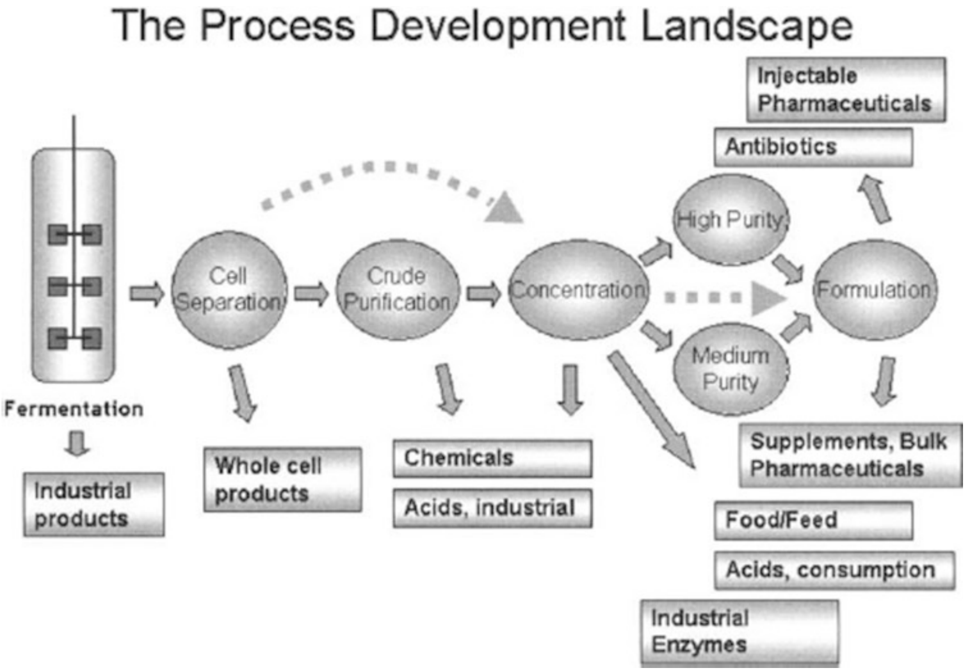


Fig. 30.8 The downstream processing cascade arranged by final product requirement

In contrast to fermentation processes, where typically one format (i.e., a fermentor vessel) is used, a number of widely differing techniques are necessary to accomplish purification and formulation of biological products. The variety of available separation technologies is large and the order and permutations of sequential processing steps are countless. Although this may hamper standardization of downstream processes, it also provides the benefit of flexibility because different unit operations can be tailored to the greatest efficiency in a cost-effective manner. Possibly a better way to illustrate a downstream process design is shown in Fig. 30.8. Here it is shown that recovery is not a linear sequential process but rather the downstream process development follows a palette of options that can be applied depending on the required specifications of the endproduct and as such, it is a modular approach.

Several general considerations limit the range of practical choices for a process design. The type of production organism (bacteria vs. fungi or mammalian cells) has an impact on the downstream process design. The fermentation media for these organisms can differ substantially and the media themselves can play an important role for the recovery strategy. Defined media often deliver a purer product but lower concentrations, whereas complex media contain more impurities but also generates more product. The physical, chemical, and biochemical properties of the desired product have an impact on the recovery strategy, as do the microbial cell's physiological states and properties. Proteins and enzymes require more "gentle" conditions to avoid unfolding or chemical denaturation than those required for small molecules. General parameters to observe are large swings in or extreme changes in pH and temperatures. Both proteins and small molecules can have hydrophobic, hydrophilic, or limited solubility properties that need to be considered when devising a recovery and purification strategy.

The principles and equipment discussed below are mostly focused on proteins and enzymes. Several other products, such as small molecules, metabolites, vitamins, and acids can require more specialized methods. However, many of the unit operations described here are also used in the recovery of these latter compounds.

Separation of Proteins and Peptides

Biological products can be inside the cells (intracellular), loosely associated with the cells, or secreted (extracellular). Each of these products requires a different approach towards the purification strategy. Intracellular and membrane-bound proteins and enzymes are more difficult to recover than the secreted ones, requiring physical or chemical disruption of cells and the challenge of separating the desired compound from viscous or entraining substances such as nucleic acids and cell wall debris. Fermentations are typically not harvested until the cell mass and the product are concentrated, often requiring some postharvest dilution to avoid entrainment losses and low equipment throughput. Continuous fermentations, on the other hand, generate a more dilute stream with a hydraulic load that can add significantly to the cost of subsequent concentration.

The simplest downstream processes include direct use of the entire fermentation broth, cells and all. This is suitable for low-cost industrial products. Another cost-effective method is the lysis of cells in the crude broth and direct use of the lysate. Yet another method that avoids lengthy purification steps is in situ extractive fermentation, whereby a product is directly extracted from the broth by use of solvents. However, many industries require more sophisticated purification methods due to safety and environmental considerations. For instance, it is imperative that live genetically modified organisms (GMO) and their DNA are not present in a released product.

As can be seen from Fig. 30.8, the nature of its final application will dictate how pure the product needs to be and which raw materials may be used in its manufacture. Typically, high purity is not required of industrial enzymes. In contrast, therapeutic proteins and small drug molecules require the highest level of purification. The purity range for food and feed compounds is somewhere in the middle. Raw materials must meet regulatory requirements for the application and must often be tested for their toxicological properties. Finally, economic considerations constrain both the range of raw materials and separations processes that can be used. Although multiple chromatographic steps are routinely included in the downstream processing of human therapeutic proteins, few industrial proteins include even a single chromatographic step [29]. Purification, if any, of industrial products is accomplished by less expensive techniques such as extraction or crystallization.

Fermentor Harvest and Primary Recovery

Extracellular products. In order to prevent cell lysis, removal of cells from a fermentation broth is usually started within hours after harvest. After cell separation, the clarified fermentation broth is more stable and can be stored refrigerated for days. Upon harvest, most primary recovery steps involve some pretreatment of the fermentation broth, which can range from simple cooling and dilution for fungal broths to pH adjustments, addition of salts, stabilizers, or flocculation agents in order to minimize degradation of the desired product and facilitate further processing. For labile compounds, such as protease enzymes, control of temperature, pH, oxidants, inhibitors, and activators is essential at harvest and throughout the recovery process. For example, it is important to maintain a molar excess of calcium ion to ensure the thermal stability of certain *Bacillus* proteases and amylases that contain calcium-binding sites.

Industrial products are mostly of the extracellular variety. They can be recovered directly from the fermentation broth. The primary recovery involves removing the cells from the broth, aptly called cell separation. Three different techniques are commonly used to achieve this goal: filtration, microfiltration, and centrifugation.

Filtration. Filtration can include filter presses, rotary drum vacuum filters (RDVFs), belt filters, and variations on synthetic membrane filtration equipment, such as filter cartridges, pancake filters, or plate and frame filter presses. These processes typically operate in a batch mode: when the filter chamber is filled up or the vacuum drum cake is exhausted, a new batch must be started. This type of filtration is also called dead-end filtration because the only fluid flow is through the membrane itself. Due to the small size of cells and their compressible nature, typical cell cakes have low permeability and filter aids, such as diatomaceous earths, perlite, or other mined materials are added to overcome this limitation. Moreover, the presence of high solids and viscous polymeric fermentation by-products can limit filtration fluxes without the use of filter aids.

In dead-end filtration, a cake forms on the surface of the pad as the filtration proceeds. The cake permeability is the most important physical property of a porous medium and the hydraulic properties of the flow and the specific cake resistance are described by Darcy's Law:

$$J = \frac{\mathrm{d}V}{A\ \mathrm{d}t} = \frac{\Delta p}{v_0 (R_\mathrm{m} + R_\mathrm{c})}, \tag{30.24}$$

where R_m is the resistance of the membrane, R_c is the resistance of the filter cake, Δp is the transmembrane pressure, A is the area available for filtration, and v_0 is the permeate viscosity. Cake resistance can be expressed in terms of the specific cake resistance:

$$R_\mathrm{c} = \alpha \rho_\mathrm{c} \frac{V}{A}, \tag{30.25}$$

where α is the specific cake resistance and ρ_c is the mass of dry filter cake per unit volume of permeate. The combination of equations (30.24) and (30.25) and integration gives:

$$\frac{tA}{V} = \frac{\rho_\mathrm{c} \alpha v_\mathrm{o}}{2\Delta_\mathrm{p}} \left(\frac{V}{A} \right) + \frac{v_\mathrm{o} R_\mathrm{m}}{\Delta p}, \tag{30.26}$$

where specific cake resistance can be obtained from the slope of a plot of tA/V vs. V/A.

Fungal fermentations, such as those of *Trichoderma* or *Aspergillus* sp., lend themselves particularly well to cell separation by filtration through a RDVF because of the ease with which the fungal mat can be shaved off by the drum's knife, renewing the filter cake surface to maintain high filtration flux.

Bacterial fermentation broths can be processed either by filtration or centrifugation, but the much smaller size of bacteria generally requires the addition of a polymeric flocculant. Most flocculants are cationic and function by bridging the negative surface charges on neighboring cells to increase the particle size and facilitate either sedimentation rate or filtration flux and clarity. The choice of flocculant and optimization of dosage is a delicate balance among obtaining good separation quality, yield, cost, and minimization of residual excess polymer in the product.

Microfiltration. Microfiltration, the use of tangential flow anisotropic membranes to permeate the product of choice while retaining solids, can be an attractive cell separation technique because it does not require the use of flocculants or filter aids. It is, in principle, a more technically sophisticated version of classic dead-end filtration processes. Microfiltration yields can be low due to progressive fouling of membranes. Advanced engineering has overcome many of the early problems and microfiltration is now used widely in the recovery of therapeutic proteins and in the food, beverage, dairy, and water treatment industries.

Microfiltration units can be configured as plate and frame flat sheet equipment, hollow fiber bundles, or spiral wound modules. The membranes are typically made of synthetic polymers such as polyethersulfone (PES), polyamide, polypropylene, or cellulosic mats. Alternate materials include ceramics, stainless steel, and carbon. Each of these comes with its own set of advantages and disadvantages. For instance, ceramic membranes are often recommended for the filtration of larger particles such as cells because of the wider lumen of the channels. However, it has been shown that spiral wound units can also be used for this purpose, provided appropriate spacers are used.

Together with related technologies such as ultrafiltration, nanofiltration, and reverse osmosis, microfiltration relies on membranes with a defined pore size. The operation mode is as "cross flow" or "tangential flow" in which filtrate is pumped parallel to the membrane surface. Cells or cell debris and large molecules are retained on one side of the membrane and the biological product solution passes through the membrane. The primary advantage of tangential flow vs. dead-end is that membrane fouling only affects the membrane pores themselves, whereas particulate matter on top of the membrane is continuously swept away. In that manner, no filter cake is created. Microfiltration is complex and much research has been done to describe the theoretical underpinnings of the process. Equipment and membrane replacement costs are high. Capital and operating costs of filters and centrifuges of equal production capacity can be more advantageous. Nevertheless, microfiltration equipment, when designed and operated properly, has the potential to deliver high throughputs due to its continuous operation, reduction of process steps, and reduction of the waste generated. It also enables the use of recycle streams to increase product yields and decrease the cost of raw materials.

Centrifugation. Centrifugation is another viable method for removing microbes from fermentation broth. Three types of centrifuges are commonly used: (a) disk-stack centrifuges (clarifiers), (b) tubular and basket centrifuges, and (c) decanter centrifuges that contain solid removing scrolls. All three can be operated in a continuous mode for higher throughput. Less common are tubular centrifuges, which generate the highest centrifugal forces but have limited throughput. The main challenge in centrifugation is to identify a suitable flocculant or process conditions whereby the cell mass is easily conveyed or discharged from the centrifuge bowl without breakup or carryover into the centrate. Although centrifugation can handle a high concentration of cell solids, filtration can provide more complete removal of trace solids that can interfere with downstream concentration or purification steps. The presence of flocculants and filter aids in cell wastes can cause added disposal costs. Often, the centrate cannot be completely clarified, thus requiring combinations of a final filtration step and centrifugation to achieve a nonturbid product.

Intracellular products. Intracellular production of bioproducts is less preferable but sometimes the only way to produce certain compounds in appreciable amounts. In this case, cell disruption is required for recovery. High-pressure homogenization, bead mills, and chemical or enzymatic disruption of the cell wall with lysozyme or similar enzymes can be used to achieve cell breakage. In the case of small molecules, organic solvent extraction has also been described. If cell debris remains in the centrate, it must be removed by methods described earlier, thus adding extra steps to the process.

Purification. Purification of the product can follow or precede the step of product concentration. Biological product purification is often and quite mistakenly equated with column chromatography. Undoubtedly, this technology is widely applied and has advanced rapidly in recent decades. Chromatography is a separation technique that includes various separation principles, such as gel filtration, size exclusion, ion exchange, hydrophobic interaction, and affinity binding/elution. A major advance has been the development of cross-linked pressure-sustaining support matrices for high volume applications. Continuous moving bed technologies, expanded bed methods, and displacement chromatography have all contributed to greater versatility and cost effectiveness of the technique and have greatly expanded the tools available for high-resolution separation. More recently, high-pressure liquid chromatography (HPLC), originally an analytical tool, has been refined for preparative separations.

Although chromatography offers the greatest potential and diversity of mechanisms for separation and purification, with some notable exceptions, it is not cost effective in the manufacture of industrial commodities due to the expense, low throughput, and low binding capacity of chromatography resins. Furthermore, the complexity and control required for reproducible operation can be cost prohibitive for very large-scale production systems. It is also uneconomical for concentration of the product. A large body of literature exists for the chromatography of biological molecules bolstered by its extensive use in the biopharmaceutical industry.

Purification can also be achieved in other ways, through precipitation with salts, crystallization, and through aqueous two-phase extraction. Some of these methods are associated with substantial capital cost, low throughput, low yields, or waste issues. Fractional precipitation, one of the oldest protein separation technologies, can be surprisingly effective to separate a compound of interest from a complex broth. For instance, the process of fractional solvent precipitation of blood plasma components has been used since World War II.

It is sometimes necessary to remove an undesirable side activity from proteins and enzymes, particularly when it is not feasible to delete the side activity genetically. A case in point is the presence of co-secreted proteases. Even if minor protease contamination from the host organism does not cause significant degradation of the product during processing, it is likely to damage the product in the final formulation over several months of storage. In other cases, side activities must be removed because they interfere in the final application. Incremental additions of chaotropic salts, such as ammonium sulfate, water-binding polymers, such as polyethylene glycol, or organic solvents are added to a solution containing the enzyme, and the relative amount of each enzyme vs. other protein impurities in each fraction is monitored. Similar to the pooling of chromatographic elution fractions, the fractional "cuts," which have the most favorable balance of net enzyme purity and recovered yield, are combined.

Crystallization is related to precipitation in that it is governed by a compound's physicochemical properties. For example, the second virial coefficient, B_{22}, which characterizes the two-body interactions between protein molecules in dilute solutions:

$$B_{22} = \frac{2\pi}{M^2} \int_0^\infty r^2 \left(1 - \mathrm{e}^{-u(r)/kT}\right) \mathrm{d}r \tag{30.27}$$

where M is the protein molecular weight, r is the intermolecular separation distance, $u(r)$ is the interaction potential, k is the Boltzmann constant, and T is the absolute temperature. The interaction potential $u(r)$ describes the interaction forces between the two protein molecules.

Proteins have long been crystallized for commercial production. This process is highly scalable, reproducible, and can be cost effective. As is the case with precipitates, crystals can be recovered via centrifugation or filtration processes. The purity achieved can be quite high, even though it is often not comparable to multiple chromatographic steps due to crystalline entrapment of impurities. Overall, it often meets the requirements for a number of industrial and food applications. A major challenge can be the time-consuming screening of appropriate precipitants, pH conditions, and temperatures and the mapping of the solubility phase diagrams. Recently developed high-throughput techniques for screening crystallization parameters have lessened the laborious nature of identifying appropriate conditions. Both precipitation and crystallization are excellent methods to concentrate the molecule of choice. A further advantage is that proteins in crystalline or precipitated form display good storage stability. Industrial biotechnology companies, such as former Genencor International, have developed large-scale crystallization methods for enzymes [30].

In the last one to two decades, aqueous two-phase extraction has become an attractive purification and concentration technique. It provides the selectivity of classical solvent extraction without its denaturing potential. Utilizing incompatible two-polymer and polymer–salt combinations and adjustments in pH and ionic strength, this technique separates proteins based on differences in hydrophobicity, surface charge, and molecular weight. Segregation of the separated molecule is governed by the partitioning coefficient (α):

$$\alpha = \frac{C_\mathrm{s}}{C_\mathrm{w}}, \tag{30.28}$$

where C_s is the concentration in the solvent phase and C_w is the concentration in the water phase. One application of the technique has been purification of genetically engineered chymosin from multiple side activities produced by the fungal host background.

An important step in the manufacture of biopharmaceutical injectables is the removal of endotoxin and viral contamination. Numerous methods have been described and include chromatographic steps and the use of filter cartridges [31].

Purification of bioproducts from fermentation, although necessary, comes with a major drawback. Biological compounds evolved in an environment of complex components are adapted to be most stable in these conditions. Highly pure forms of bioproducts have a tendency towards lower chemical and physical stability. For instance, proteases are prone to autolysis in the absence of impurities that inhibit degradation. This is one of the major challenges faced by the formulator of biological products (see section "Formulation").

Concentration. Clarified filtrates, centrates, or column eluates are usually too dilute for use in their specific applications, thus, substantial amounts of water must be removed. This can be achieved by evaporation or by ultrafiltration. Concentration methods used in industrial settings, such as evaporation, which is done under vacuum, and solvent extraction, may or may not be suitable for dewatering proteins because of their potential for thermal or chemical denaturation, and due to high energy costs associated with evaporation. The benefit of evaporation is that nonvolatile compounds that may stabilize the proteins are retained.

Ultrafiltration (UF), which can be a viable alternative, has membrane pore sizes that are much smaller than those in microfiltration, allowing only water and small molecules to travel through the membrane. Thus, proteins, peptides, and other large molecules can be retained as water is forced through the membrane. Typical molecular weight cut-off sizes for UF membranes are 5–100 kDa. This technology has emerged to become one of the most common methods for concentration of large biomolecules. UF is an attractive method because of its low energy consumption. It can be combined with diafiltration whereby repeated additions and removal of water lead to a cleaner product. UF fluxes and yields are often significantly enhanced by upstream removal or omission of potential membrane foulants, such as polysaccharides or antifoams. In the case of smaller molecules, nanofiltration can be used.

Nanofiltration membranes have a molecular weight cut-off ranging from 100 to 1,000 Da and with such small pore sizes are able to retain molecules such as humic acids and certain salts. This allows for production of parasite- and solids-free water without the need of chemicals.

Precipitation, crystallization, and extraction can also be used for concentration, but are more typically utilized as purification techniques. The latter methods also may cause problems with their environmental impact in as much as they require substantial additions of chemicals.

Decolorization and finishing. Decolorization is sometimes required for certain applications, mostly as an aesthetic preference. It is always desirable to solve these issues upstream. For example, color can be minimized by choice of fermentation medium components and control of the sterilization cycle so as to lessen the Maillard reactions between nitrogen and sugars. Color can also be reduced by treatment with activated carbon, use of antioxidants, and by diafiltration with membranes. Carbon-impregnated filter pads can be used to combine polish filtration with a decolorization step.

"Finishing," sometimes also called "polishing" refers to one or more filtration steps at the end of the downstream process that have the goal of clarifying and reducing the microbial load of the product. Solids are most conveniently removed using a filter press loaded with cellulosic, synthetic polymer pads, or disposable submicron gauge filter cartridges. Precoating pads with filter aids prevent premature fouling. All these methods can be used to remove microbes to the point of sterility if required. Often, formulation ingredients are added to the product concentrates before the final decolorization and finishing steps are performed.

Formulation

Formulation of a product depends on how the endproduct will be used and the ability to satisfy safety requirements. Many bioproducts are supplied in liquid form, as it is convenient to meter and use. If the compound is slated for dry products, the concentrate can either be stored for later use or applied directly for granulation. In many industries, proteins, peptides, and other bioproducts are supplied as encapsulates or granules in order to meet safety standards.

As with fermentation and downstream processing, precise formulation recipes are often kept as closely guarded trade secrets. Comparatively few publications on new approaches or new technologies are available in the literature [32]. However, the intellectual property landscape is filled with composition of matter patents.

Proteins are large molecules comprised of tens to hundreds of amino acid subunits. Their unique properties are associated with specific conformations of the amino acid chain. The active conformations are thermodynamically quite fragile, typically only a 5–15 kcal/mol difference in free energy between the stable and the unfolded conformations. Additionally, the amino acid subunits that define the protein contain a variety of reactive groups that are subject to chemical degradation pathways resulting in the loss of structure and/or activity. Thus, ensuring the physical and chemical integrity of a protein's native structure is a primary concern when using proteins or enzymes in applications where its structure is important. Many proteins are only soluble up to 20–30% (w/w) and therefore considerations for keeping the compounds well solubilized are of importance.

Dry formulations. Because of the long-term instability of proteins in aqueous solution, enzyme producers and formulators have attempted to produce stable solid formulations since enzymes were first used. Proteases tend to degrade via autolysis and can be incompatible with surfactants. These problems are easily overcome by storing the enzyme in the solid state. Initially, commercially produced enzymes for laundry detergents were spray-dried or sprayed onto salt cores without coating. Spray-drying provides a fast and cost-effective way of compartmentalizing

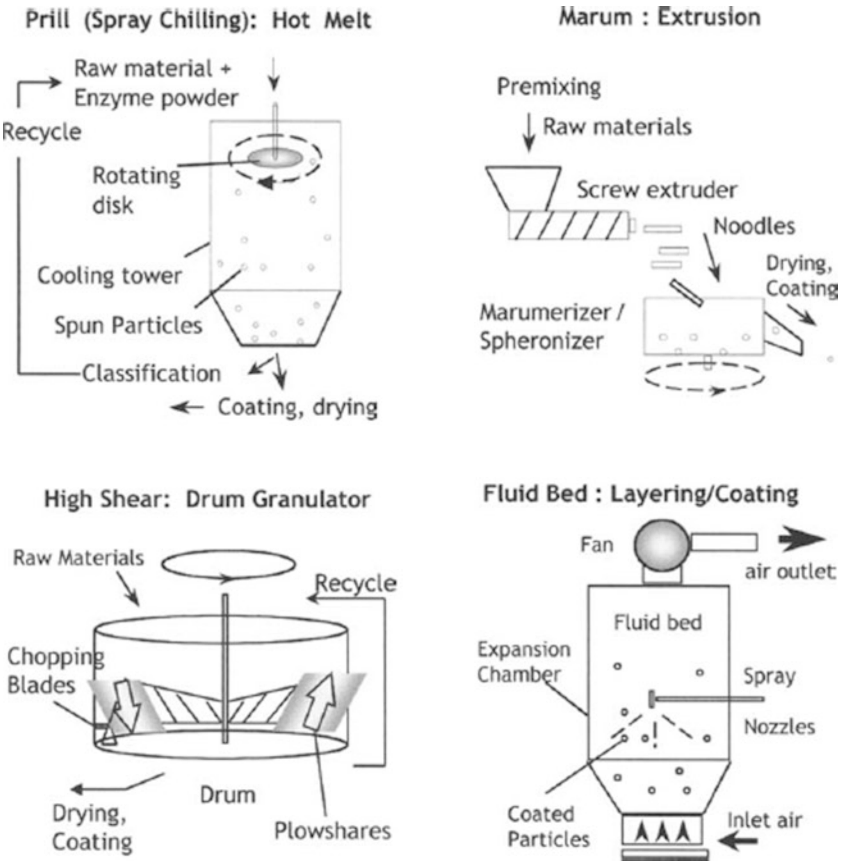


Fig. 30.9 Common granulation method in enzyme manufacturing

two incompatible ingredients and maintaining enzyme activity over a product's typical shelf life. However, noncoated granules have led to cases of allergic reactions in formulating plant personnel and resulted in a setback for the enzyme industry in 1970s. Thus, the technologies used to produce modern dry products have evolved far beyond simple spray-drying.

Product safety is an important consideration as many bioproducts can cause adverse responses in humans and many therapeutic fermentation products are very potent bioactive molecules. Negative effects can include irritation and allergic reactions, which are due to the possible immunogenic nature of many biological compounds. Lyophilization, or freeze-drying, which is used in therapeutic applications results in a powdered product that is contained in vials as an additional safety measure to prevent direct exposure. It offers a gentle way of protecting a compound's fragile nature because the drying temperature is kept below freezing. However, these technologies are expensive.

Many of today's bioproducts are supplied in a granular form. Granulation is a generic term for particle size enlargement. Granules in consumer products need to provide a tough barrier to prevent release of the bioactive molecules in airborne dust, while providing quick release once used in the final application. Enzyme granules contain coatings that are designed to withstand the physical impact and shear forces typically encountered during powder processing, maintaining excellent flowability throughout the process.

Many ways of producing enzyme granules have evolved, with just a handful of methods being currently in common use. These include prilling (spray-chilling), marumerization/spheronization, drum and high shear granulation, and fluid-bed coating (Fig. 30.9). The latter two techniques are superior in producing low dust granules. An overview of the different technologies is provided in Table 30.3. In general, fluid-bed technology (Fig. 30.10) is the most flexible approach, giving granules with the most uniform appearance and smooth coatings (Fig. 30.11). Other technologies may have an edge in either cost or throughput, as they are more amenable to continuous operation. The advantage of the fluidized-bed method is that the entire granulation process can be carried out within a single, contained piece of equipment. The spray-coating process allows the sequential application of layers of different thicknesses and compositions with almost infinite flexibility. In addition, spray coating allows one-step application of controlled release coatings.

A less commonly used method is spray-chilling, whereby the product is incorporated into meltable cores. This method is straightforward and inexpensive but has the drawback of being limited by the meltability of its ingredients. Moderate

Table 30.3 Comparison of powdered enzyme production methods

Technology	Equipment	Advantages	Disadvantages
Prilling (hot melt)	Prilling tower	Continuous process	High product dust
		High capacity	Low melting point
		Easy rework	Requires predried enzyme
Extrusion/spheronization (marum)	Screw extruder, Marumerizer, fluid bed coater	Inexpensive raw materials	Complex multistep process
		Continuous process	High dust during process and in product
			Requires highly concentrated enzyme
Drum granulation (high shear)	Lödige mixer	High capacity	Dusty, multistep process
	Littleford mixer	Continuous process	Requires highly concentrated enzyme
	Fluid bed or drum coater	Tough granules	Wide particle size distribution
Fluid bed (layering/coating)	Fluid bed coater	Single, contained reactor	Batch process
		Flexible formulation	Difficult rework
		Tough uniform granules	

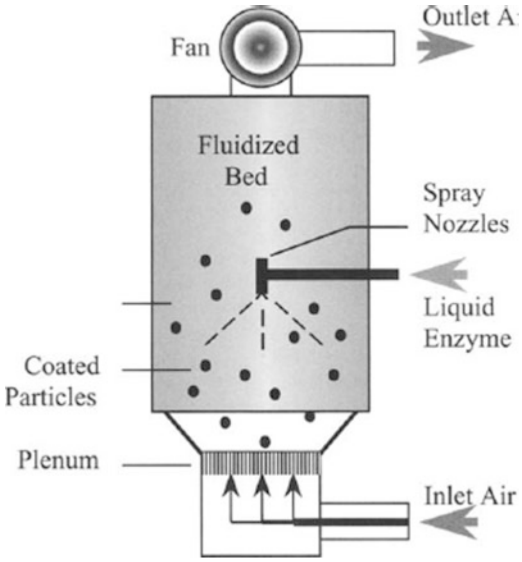


Fig. 30.10 Schematic of a fluidized-bed coater

to high temperatures during the shipping or during use of the product can lead to agglomeration of the capsules.

Solid formulations can provide some significant advantages, such as enhanced stability, delayed or controlled release, and protection against deactivation during harsh applications. One example of the latter is the use of granules to encapsulate cellulase enzymes against deactivation in the steam-pelleting process that is used in producing animal feed pellets. The use of effective stabilizers and coatings can prevent or retard exposure to moisture under high heat, yet allow release of the enzyme in the subsequent feed application.

Many therapeutic molecules can be encapsulated using other techniques, such as coacervation, micro- or nanospheres, microcapsules, or liposomes. Coacervation, often used for flavor encapsulation, can be a simple aqueous phase separation of immiscible droplets, or in the case of complex coacervation, the use of opposite electrostatic charges. It is one of the basic processes of capsule wall formation. Other biochemicals, such as flavors, vitamins, or citric acid are sold directly in the crystallized forms and can be incorporated directly into the final product, such as tablets, capsules, granules, and so on.

Liquid formulations. The successful application of enzymes in liquid formulations presents several technical obstacles that are not encountered in powders and granules. These problems stem from the fact that liquid products are complex aqueous solutions, and physical separation of enzymes from other potentially inactivating ingredients is impractical. Necessary ingredients for a viable endproduct can affect the physical and chemical stability of enzymes. In addition, as other types of enzymes are added to formulations that already contain proteases, proteolytic degradation of enzymes is a concern.

Liquid formulations can be used as pastes or slurries, dispersions or emulsions (in aqueous or other solvents) to physically separate the compound to be protected. Dispersions, pastes, or slurries are not commonly used but are a feasible approach when highly concentrated products are needed. For instance, many early analytical enzymes were sold as crystal slurries to preserve stability. Dispersions can be prepared from dry materials. Most of the dispersions are visually cloudy in appearance due to the particle size of the dispersed compounds and are thus undesirable in most endproducts.

Many catalytically active proteins are hydrolases and are subject to three principal means of deactivation: denaturing or unfolding, catalytic site inactivation, and proteolysis. Denaturation is best minimized by controlling temperature and pH and by avoiding the presence of chemical denaturants. Catalytic site inactivation is prevented by supplying sufficient levels of cofactor, typically a metal cation, and preventing oxidation of the active site, for example, by formulating with antioxidants. Alternatively, oxidative

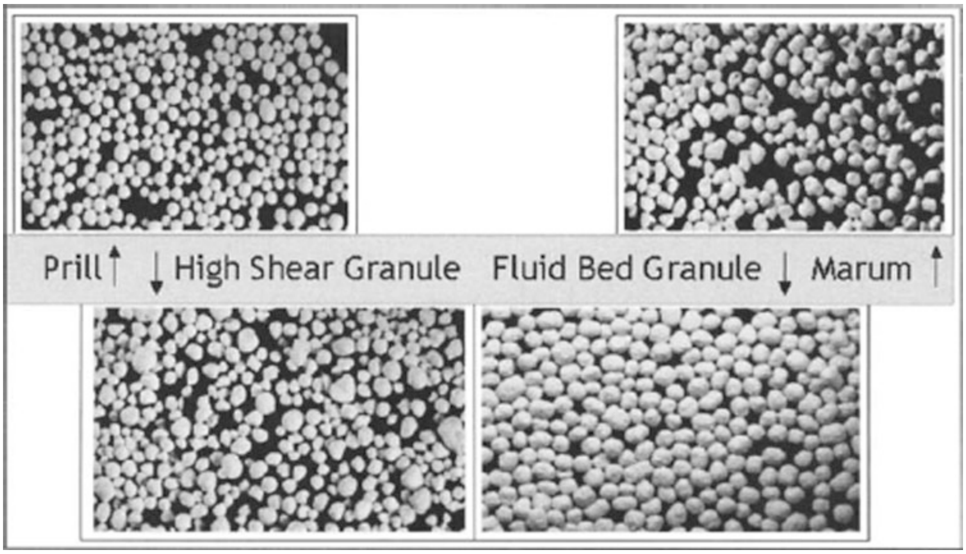


Fig. 30.11 Commercially available enzyme granules (Courtesy Genencor International)

resistance can be engineered into the protein structure. Finally, in the case of proteases, proteolysis or autodigestion can be minimized by reducing water activity or by addition of inhibitors. Thus, the final liquid formulation includes stabilizers, antimicrobial substances, and osmolytes to reduce the water activity and increase thermodynamic stability. Useful water-sequestering compounds include sugars and other polyols, such as glycerol, sorbitol, and propylene glycol. Useful inhibitors include substrate analogues such as peptides and acid salts. Once the formulation excipients are added, the last step includes a final polish filtration to provide a clear liquid.

Different industries pose different challenges to the protein formulator. Many feed enzymes are sold as formulated liquid concentrates. In this case, the major requirements for a liquid formulation are enzymatic stability and preservation against microbial growth. It is sometimes not appreciated that the dominant factor affecting enzyme stability is the intrinsic stability of the enzyme itself; formulation can do very little to correct for a structurally labile protein. Therefore, it is advisable to make stability an important criterion of the initial enzyme screening process.

An important case is the application of enzymes in laundry detergents. Market trends in the United States show that consumers prefer liquids to powder detergents by a ratio of 2 to 1. These products are stored with no temperature control on shelves in the presence of harsh surfactants, such as linear alkylbenzyl sulfonate (LAS) and require extraordinary measures for stabilization. LAS, by its nature as an effective cleaning agent, causes surfactant-induced unfolding in proteins. There are countless examples of the development of stabilization systems in the intellectual property space. A common theme is to reduce the water activity and to use borate/glycol stabilizers that bind to the active site of proteases.

Microbial growth is fairly easy to control by the addition of food grade antimicrobials such as sodium benzoate, methylparaben, and various other commercial preparations. In addition, reduction of the water activity through use of sequestering agents aids in controlling microbial growth. The prevention of precipitation and hazes is often a highly empirical challenge and one that is very specific to the enzyme and the specific raw materials used in the process. It is best to use ingredients with high solubility and to screen potential formulations by extended storage at high and low temperatures. In the case of proteins, an additional objective is to keep the pH value as far away as possible from the isoelectric point of the protein, at which the protein is least soluble.

Emulsions are the most commonly used type of formulation in the food, paint, and personal care industries. The typical emulsions are “oil in water” or “water in oil” preparations. Proteins, peptides, metabolites, and other fermentation products can be incorporated into emulsions. Covalent attachment of hydrophobic compounds, such as fatty acids or polyethylene glycols, to proteins or peptides lends these hydrophilic molecules to reside in the lipid phase.

Whole Cell Recovery

The recovery of whole cells is best explained by the manufacturing procedure for baker’s yeast. This process is almost identical to the early stage of protein recovery, except that the final product is the cell instead of the filtrate. After fermentation, the cells are spun out with a centrifuge, washed with water, and recentrifuged to yield a yeast cream with a solids concentration of approximately 18%. Cream yeast can be loaded directly into tanker trucks and delivered to customers equipped with an appropriate cream yeast handling system. Alternatively, the yeast cream can be pumped to a plate and frame filter press or an RDVF and dewatered to a cake-like consistency with 30–32% yeast solids content. The press cake yeast is crumbled into pieces and packed or spray-dried for dry products. After packaging, the yeast is ready for shipping to retail.

Separation of Small Molecules and Metabolites

Metabolites, vitamins, organic and amino acids, and specialty chemicals are commonly referred to as small molecules, particularly in the therapeutic world to differentiate these compounds from proteins or peptides, commonly named biopharmaceuticals or biologicals. Many of the same unit operations are applied for the recovery and purification of small molecules as described earlier. Purification of bulk products has cost constraints, whereas pharmaceuticals are subject to strict purity requirements.

Possibly the largest class of molecules produced in microorganisms are acids. This group includes acetic acid, citric acid, gluconic acid, lactic acid, and, of course, the amino acids. Citric acid is recovered (Fig. 30.14) after the cells have been removed by RDVF. It is then solvent extracted or precipitated as a calcium salt by adding hydrated lime [$Ca(OH)_2$]. The precipitate is recovered by a second RDVF and then acidified with sulfuric acid where it forms citric acid and calcium sulfate (gypsum). A third RDVF removes the gypsum. Lactic acid is usually produced as an alkali or earth alkali salt.

Another excellent example is given by penicillin recovery, a weak acid. Penicillin can be recovered by solvent extraction and its partitioning coefficient into solvent can be steered by the pH value. In the dissociated form it is more hydrophilic, whereas in the protonated form it is more hydrophobic. Thus, at low pH, penicillin partitions into the solvent phase.

A number of relatively new methods are being investigated to improve the recovery of small molecules. These methods include electrokinetic separators with bipolar membranes, simulated moving-bed chromatography, and supercritical fluid extraction. The latter is practiced for food components. It has also been described for proteins but has not yet found wide acceptance in this field. A fast-growing field is the production of bioethanol via fermentation processes either from milled corn or from recycled biomass. The fermentation and saccharification processes can occur simultaneously in the fermenting tank by means of saccharification enzymes (amylases, cellulases).

Once the mash is fermented, it is transferred to a distilling unit. The distilling system is designed to produce greater than 90% combustible bioethanol from fermented grain mash. The system can be arranged in an array of columns: mash columns, aldehyde-extraction columns, and rectifying columns. The remaining alcohol-free liquid is drawn off from the bottom of the rectifying column. This liquid can be blended with the watery meal from the mixing tank before cooking.

For the dehydration step, molecular sieves made of zeolite can be used. The alcohol is extracted from the molecular sieve at 99.75%, dehydrated as vapor and captured in condensers. New research focuses on alternate aspects of this separation process to improve efficiency such as developing membranes for ethanol recovery. Low-cost pervaporation, an energy efficient combination of membrane permeation and evaporation, for ethanol extraction and bioethanol production holds much promise.

The liquid left after distillation is subjected to centrifugation, where most of the suspended solids are separated. The clear liquid can be recycled by adding it back to the starch conversion stage. The moist cake released by the centrifuges is mixed with the syrup produced by the evaporator to form a homogenous mixture and is dewatered in dryers. The dryers produce a Distillers Dried Grains with Solubles (DDGS) meal, which can be formed into pellets. These are used in many applications, most prominently in animal feed.

It is evident from the discussion earlier that downstream processing and formulation depend heavily on the industry and the product. There are countless possibilities on how a molecule can be recovered, purified, and formulated. More importantly, the fermentation, recovery, and formulation processes are intimately intertwined. Depending on the type of production organism and fermentation process used, the downstream processing has to adjust accordingly. It is thus imperative that process development is highly integrated. In the modern fermentation industry, this is usually achieved with process development and formulation departments working closely to deliver the best product at a competitive cost. Moreover, optimal integration of these departments with the strain development teams is also seen. Current industry trends also include reduction in the number of process steps because each step results in a yield loss. Sustainable technologies such as reuse of water, ingredients, and minimal by-product generation are also being adopted. Ultimately, all these efforts are beneficial for improving the production economics.

Regulatory Considerations

Bioproducts are regulated according to their application, for example, whether enzymes are used in food, feed, detergents, textile processing, and so on. Regulations also differ country by country. Irrespective of legislation, it is prudent for the product manufacturer to ensure that the production process, the potential product, and its intended use are safe prior to introduction.

Along with transfer of the strain, fermentation, recovery, and formulation process to the manufacturing plant, validated quality control assays, which are suitable for ensuring product quality, must also be provided. The EPA, FDA, and other agencies throughout the world are all involved in regulation of products derived from naturally occurring organisms and GMO. Essentially the regulations require: (1) the use of nonpathogenic organisms, (2) the absence of endotoxin or other toxins, (3) a well-documented and reproducible fermentation and downstream process, and (4) use of safe and well-characterized organisms and processes. In response to the latter, not only are the GMO well characterized, they are designed to barely survive in the general environment. Because recombinant organisms can be grown under contained good large-scale industry practices, safety issues relating to environmental release are minimized.

For control and containment of waste, regulatory requirements dictate appropriate containment that is related to the risk presented by the organism. Most of the processes are contained and hence subject to guidelines for industrial applications of recombinant DNA. Containment may be achieved biologically on the basis of inherent properties of the organisms, for example, their survival in the general environment is limited. Also, operation and design of the manufacturing facility minimizes potential releases of the recombinant organism to the environment. The degree of physical containment is matched to the risk presented by the organism. The same considerations apply to products made by recombinant organisms. Air monitoring is performed in production environments to evaluate levels of protein. This is important for assessing the efficiency of containment, respiratory protection requirement, and ensuring that environment levels meet occupational exposure limits.

Delivery of Products

Figure 30.12 represents a cell factory model to produce several categories of products via fermentation and biocatalysis processes. Carbon conversion in a cell factory into product(s) is essential to maintain the necessary energy (ATP, NAD(P)H, FADH) flow for the cell to replicate and maintain itself. The production of biochemicals is strictly linked to cell functions like cell synthesis (reduction–oxidation balance, cofactor regeneration, charge neutralization), and substrate turnover. Maximization of conversion yield, therefore, requires minimization of wasteful functions. Likewise, maximization of conversion rate requires maintenance of maximum driving force, e.g., ATP per conversion, key intracellular metabolite concentration, biomass per ATP, sufficient enzymatic activities, etc. Such principles continue to evolve in developing synthetic biology based processes, but it is too early to call them "laws of biotechnology." Synthetic cell engineering continues to significantly improve beyond what nature has already made possible, reaching 85–95% of what's theoretically possible. Accordingly, most favorable hosts, enzymes, raw materials, and process steps should be optimally integrated to achieve the best product economics and production efficacies.

Synthetic cell engineering is primarily the integration of four components: mining of genetic diversity and sequencing, cell physiology analysis and system modeling, synthetic pathway design and optimizing construction, integrated process design and engineering (see Fig. 30.13). Such integration requires several upfront tools outlined below:

1. Expression vectors to clone genes and adjust their copy numbers, and multiple markers to clone multiple genes
2. Transformation methods for above, gene deletion and integration systems

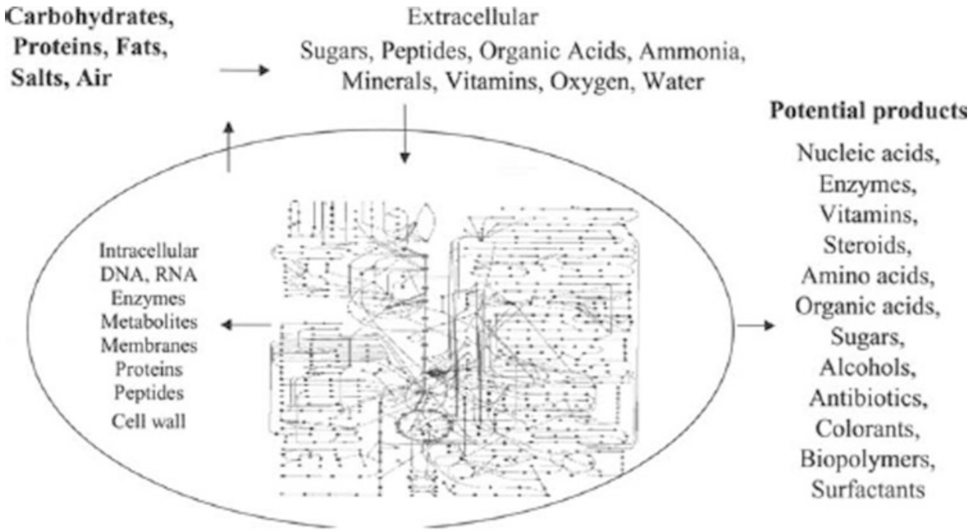


Fig. 30.12 A cell factory model

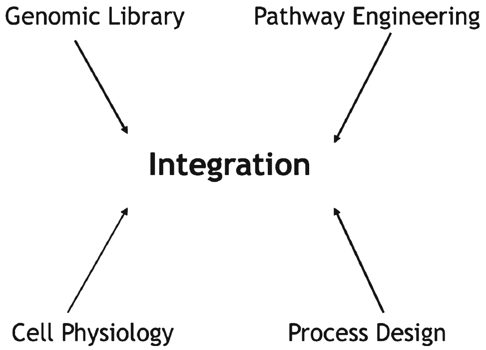


Fig. 30.13 A cell factory model depicting the four components that need to be integrated to produce several categories of products via fermentation and biocatalysis processes

3. Codon optimization of heterologous genes for higher expression
4. Mutagenesis of genes for in vivo regulation of enzymes
5. Transcriptional control, i.e., change promoter sequences to achieve different mRNA levels
6. Engineering of ribosomal sequences for translation modulation
7. Mutagenesis of pathway genes to change enzyme activity, specificity of substrate, product or inhibitor

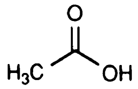
To develop cost-effective products for the biobased industry, the field of synthetic biology is continually developing a holistic understanding of microbial metabolism, and the effect of end-product on cell growth and maintenance. For example, functional genomic techniques now can monitor cell-wide transcription, translation, levels of metabolites and their interactions. By integrating the cell system-wide data and models, cell metabolism can be analyzed. The system knowledge can then be used to design/assemble cell components for specific functions and eventually build a cell factory platform.

Also, the size of the worldwide industrial chemicals business is ever increasing from $1 to $2 trillion, and global companies like DuPont, Genencor, Goodyear, BASF, DSM, Monsanto, Dow, Cargill, ADM, Tate & Lyle, Roquette, Mitsubishi, Braskem, Rohm & Haas and even new companies like Gevo, Amyris, Genomatica, Verdezyne, Ceres, and more continue to invest in industrial biotechnology to create high-value biochemicals, biomaterials, and biofuels. These companies are all interested in commercializing the biomaterial building blocks like propane diol, isoprene, succinic acid, acrylic acid, levulinic acid, isosorbide, butane diol, butylenes, ethylene glycol, ethylene, lysine, caprolactem, terephthalic acid, adipic acid, furan dicarboxylic acid, and of course natural sugar polymers like starch, cellulose, and xylose and their derivatives (see Fig. 30.14). But sustained growth, improved economics, and renewable resources will have to go hand in hand to enable replacement of petroleum derived chemicals and processes. Industrial-scale manufacturing aspects such as uses, synthesis methods, and costs of some of these bioproducts are reviewed below as well as in Chap. 31.

Organic Acids and Polymers

There are many organic acids that can be produced by microbial or biochemical means. However, at present, only acetic acid (as vinegar), citric acid, itaconic acid, gluconic acid, 2-keto-gulonic acid, and lactic acid are produced industrially by fermentation. Other organic acids, such as fumaric, gallic, malic, and tartaric acids, once produced by fermentation or enzyme processes, are now produced commercially, predominantly by the more economic means of chemical synthesis.

Acetic acid. Acetic acid as a chemical feedstock is manufactured by chemical synthesis. Acetic acid in the form of vinegar (at least 4% acetic acid by law) is produced largely via the oxidation of ethanol by bacteria of the *Acetobacter* genus [33].



Vinegar is one of the oldest known fermentation products, predated only by wine and possibly by certain foods from milk. First derived from the spoilage of wine, vinegar has been used as a condiment, food preservative, medicinal agent, primitive antibiotic, and even today as a household cleansing agent. Today, vinegar is produced almost entirely for use in foods. Vinegar may be defined as the product of a double fermentation: an alcoholic fermentation of a sugar source (fruits and their juices, cereals, syrups) usually by a selected strain of the yeast *Saccharomyces cerevisiae* or *ellipsoidens* and a second fermentation to oxidize the alcohol (including synthetic) to acetic acid by a suitable culture of *Acetobacter* organisms. The theoretical maximum yield of acetic acid on glucose by this route is 67% (two moles of acetic acid produced from every mole of glucose consumed). A homofermentative culture, *Clostridium thermoaceticum*, is known to be capable of fixing CO_2 and yielding three moles of acetic acid from one mole of glucose under anaerobic conditions. However, the technology for this process has not been commercialized.

Several vinegar-manufacturing processes are commercially used, including the following:

1. Trickling-bed reactor
2. Submerged cell reactor
3. Tower reactor

The circulating, trickling generator, which is still widely used, is a large tank generally of wood, such as redwood or

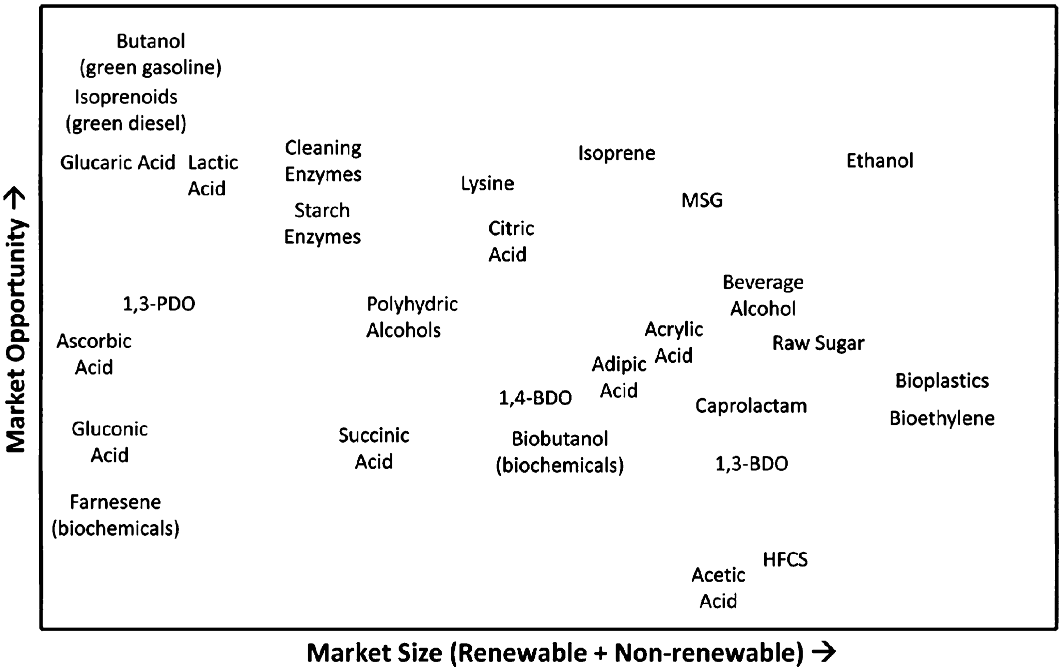


Fig. 30.14 Market opportunity (growth potential) vs. current market size of biochemicals and petrochemicals

fir but preferably cypress. Air is circulated in the generator by a number of equally spaced inlets. A pump circulates the ethanol–water–acetic acid mixture from the collection reservoir up through a cooler to the top of the tank. The liquid trickles down through the packing and returns to the bottom reservoir. The temperature of the generator is about 29°C at the top and 35°C at the bottom. A portion of the finished vinegar is periodically withdrawn from the reservoir and replaced with the ethanol-containing charge to maintain ethanol concentration in 0.2–5% range. If ethanol is depleted in the generator, the *Acetobacter* die, and the generator becomes inactive.

The Frings Acetator (produced by the Heinrich Frings Company of Bonn, Germany) is a submerged batch fermentor. It consists of a stainless steel tank with internal cooling coils, a high-speed, bottom-entering agitator, and a centrifugal foam breaker. The unique feature of this Acetator is its highly efficient method of supplying air, accomplished by high-velocity self-aspirating rotor that pulls air in from the room to the bottom of the tank. When the ethanol content falls to 0.2% by volume about 35–40% of the finished product is removed. Fresh feed is pumped in to restore the original level, and the cycle starts again. Cycle time for 12% vinegar is about 35 h, and the rate of production can be 10× that of the trickling generator. The yield of acetic acid on ethanol is higher as well, for example, 94 for the Acetator compared to 85% for the trickling generator. However, much more extensive refining equipment is necessary for filtering vinegar produced by the submerged process because the mash contains both the vinegar and the bacteria that produced it.

The tower fermentor is a relatively new aeration system applied to vinegar production. The fermentor is constructed of polypropylene reinforced with fiberglass. Aeration is accomplished through a plastic perforated plate covering the cross-section of the tower and holding up the liquid. The cost of the tower fermentor is said to be approximately half that of a Frings Acetator of equivalent productive capacity. It has been reported that the tower fermentor is satisfactory for producing all types of vinegar.

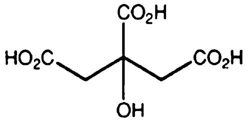
Since 1993, a two-stage submerged fermentation has been used to produce vinegar of more than 20% acetic acid. In the first fermentor, alcohol is added slowly to a total concentration of about 18.5%. After the acetic acid concentration has reached 15%, about 30% of the liquid from the first fermentor is transferred to a second fermentor. In the second fermentor, the fermentation is carried out at a reduced temperature of 18–24°C and continues until the alcohol is almost depleted.

Recombinant strains of *Acetobacter aceti*, cloned with either alcohol dehydrogenase or aldehyde dehydrogenase, have been tested for vinegar production. The bacteria with the aldehyde dehydrogenase gene produced acetic acid more rapidly than those with the alcohol dehydrogenase and were more resistant to high concentrations of acetic acid.

Vinegar clarification is accomplished by filtration, usually with the use of filter aids such as diatomaceous earth or bentonite. After clarification, vinegar is bottled, sealed tightly, pasteurized at 60–65°C for 30 min, and then cooled to 22°C. Vinegar can be concentrated by freezing or by a reverse osmosis membrane process.

The world production (excluding China and Russia) of 10% vinegar is estimated to be about two billion liters a year or about 200,000 metric tons of acetic acid. Price depends on the source (fruit, malt, grain, etc.).

Citric acid. Citric acid, whose structure is shown below, is the most important organic acid produced by means of fermentation.



Citric acid is used in soft drinks, candies, wines, desserts, jellies, jams, as an antioxidant in frozen fruits and vegetables, and as an emulsifier in cheese. As the most versatile food acidulant, citric acid accounts for about 70% of the total food acidulant market. It provides effervescence by combining the citric acid with a biocarbonate/carbonate source to form carbon dioxide. Citric acid and its salts are also used in blood anticoagulants to chelate calcium, block blood clotting, and buffer the blood. Citric acid is contained in various cosmetic products such as hair shampoos, rinses, lotions, creams, and toothpastes. More recently, citric acid has been used for metal cleaning, substituted for phosphate in detergents, for secondary oil recovery, and as a buffer/absorber in stack gas desulfurization. The use of sodium citrate in heavy-duty liquid laundry detergent formulations has resulted in a rapid increase in the use of citric acid.

Wehmet first described citric acid as a product of mold fermentation in 1893. However, it was only in 1919 that commercial fermentation processes based on sucrose were developed [34]. Although many organisms have been shown to produce citric acid from carbohydrates, *A. niger* has been the best organism for industrial production. Figure 30.15 summarizes the reactions leading to citric acid from glucose. It is worthwhile to note that one mole of glucose yields one mole of citric acid with no consumption of oxygen. The overall reaction is actually energy yielding. It yields one mole of ATP and two moles of NADH per mole of citric acid, and minimal growth is needed during production. This aspect makes fermentative production of citric acid a good candidate for the process of cellular immobilization.

Microbiological production of citric acid can be implemented by three techniques:

- Solid-state fermentation
- Liquid surface fermentation
- Submerged culture fermentation

In solid-state, or Koji, fermentation, *A. niger* is grown on moist wheat bran (70–80% water) and produces citric acid in 5–8 days. This process is practiced only in Japan and accounts for about one-fifth of Japanese citric acid production. In liquid surface, or shallow tray, fermentation, beet molasses (containing 48–52% sugar) or cane molasses of blackstrap (containing 52–57% sugar) or high test (containing 70–80% sugar) are introduced into a mixer. Dilute sulfuric acid is added to adjust the pH to about 6.0. Phosphorus, potassium, and nitrogen in the form of acids or salts are added as nutrients for proper mold growth and optimal citric acid production. The mix is then sterilized and finally diluted with water to a 15–20% sugar concentration. The medium flows by gravity into shallow aluminum pans or trays arranged in tiers in sterile chambers. Most chambers have provisions for regulation and control of temperature, relative humidity, and air circulation. Each tray holds about 400 L of solution at a depth of 76 mm. When the medium has cooled to about 30°C, it is inoculated with spores of *A. niger*. The tray fermentation requires 8–12 days. The pH drops to about 2 at the end of the fermentation, and the acid content varies from 10 to 20%. Some oxalic and gluconic acids are also formed. The temperature is maintained at 28–32°C during the fermentation. Sterile air is circulated through the chambers, and the relative humidity is controlled between 40 and 60%.

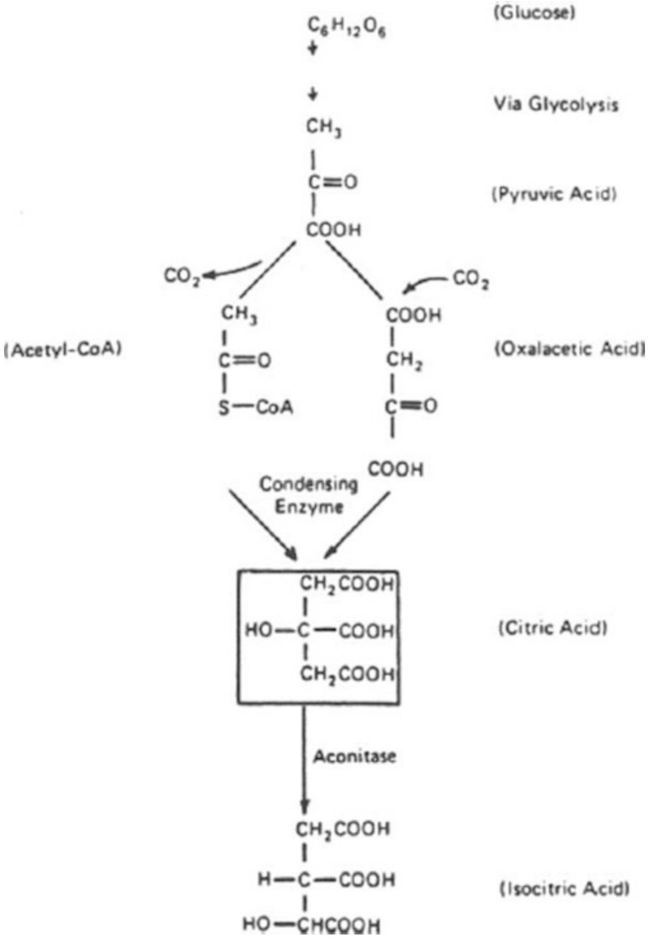


Fig. 30.15 Pathway leading to citric acid from glucose

Most newly constructed plants have adopted the submerged culture or deep fermentation process. The fermentation medium consists of sucrose (around 200 g/L) and mineral salts to provide a balanced supply of iron, zinc, copper, magnesium, manganese, and phosphate. The provision of a suitable culture medium is the most critical factor in obtaining a high yield of citric acid. The fermentation is carried out at 25–27°C. Oxygen is provided by bubbling air at a rate of 0.5–1.5 volumes of air/volume of solution/minute without mechanical agitation. It is generally accepted that the formation of pellets between 1 and 2 mm in diameter in the fermentation mash is most desirable. Pelleting reduces broth viscosity, increases oxygen transfer, and simplifies mycelium separation in the recovery scheme. The submerged fermentation has a time cycle of 6–9 days. The yield of citric acid on sugar varies from producer to producer but the theoretical maximum is 112% on sucrose. The liquid surface fermentation yield is high (90–95%) but the submerged culture fermentation is a bit lower. Reducing the formation of by-products, mainly oxalic acid, has resulted in improvements in the yield of the submerged culture process, nearly reaching those of the surface culture process.

The fermentation broth from the solid state, surface culture, or submerged culture process is treated similarly for recovery and refining of citric acid. Two recovery methods are used: precipitation and filtration, and extraction. A process flowsheet including the fermentation and refining section using the first method is shown in Fig. 30.16. The mycelium, which is initially filtered out of the fermentation liquor, may be used as fertilizer after proper weathering and processing. The clarified liquor flows to precipitating tanks fitted with stirrers where it is heated to a temperature of 80–90°C. The oxalic acid present is separated by preferential precipitation through the addition of a small amount of hydrated lime. The resulting calcium oxalate is purified separately in a manner similar to the process described for citric acid recovery.

Approximately one part of hydrated lime for every two parts of liquor is added slowly over a 1-h period while the temperature is raised to about 95°C. The precipitated calcium citrate is filtered on a vacuum filter, and the filtrate, free of citrate, is discarded as waste. The calcium citrate cake is moved to acidulation tanks, where it is acidified with dilute sulfuric acid. It is then filtered, and the citric acid mother liquor is decolorized by a charcoal treatment. The purified liquor is concentrated in a vacuum evaporator, run into a crystallizer where, upon cooling, citric acid crystallizes, generally in the form of the monohydrate. The resulting acid is of USP grade.

The extraction method treats the filtered fermentation liquor with a highly selective solvent, tri-*n*-butyl phosphate in kerosene, and then recovers free citric acid by counterextraction with water. The aqueous solution, which is further concentrated and crystallized, yields 92% citric acid with 8% soluble impurities.

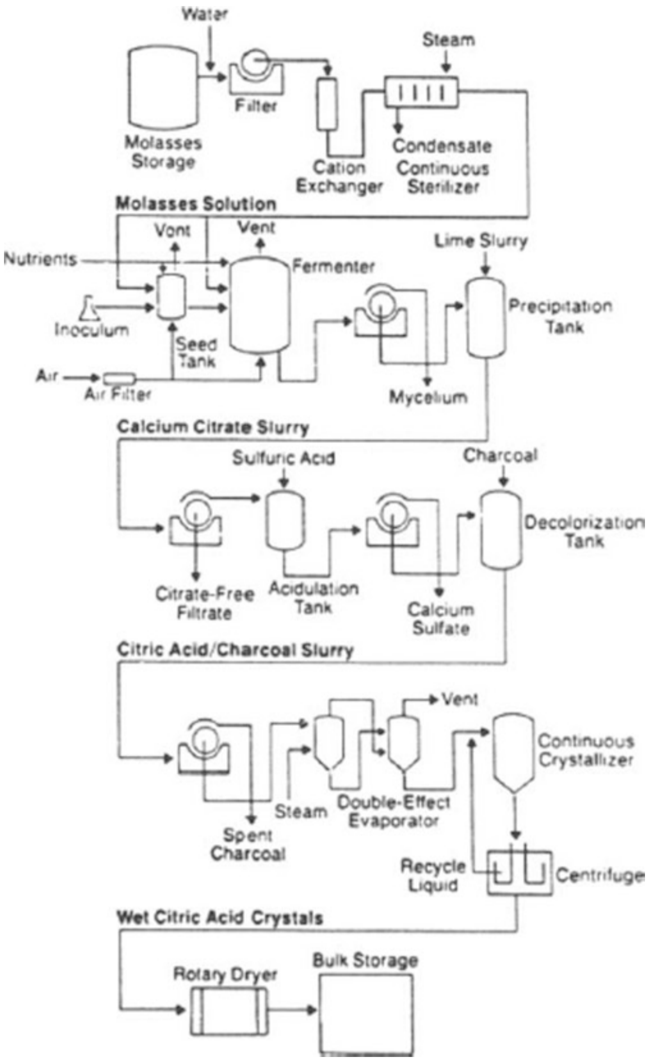
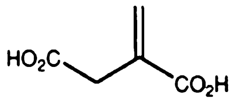


Fig. 30.16 Citric acid flow sheet

In the last two decades, the citric acid industry has seen some major changes in ownership and expansions. The worldwide production stands at about one million ton a year. The major producers include the Swiss-based Jungbunzlauer, La Citique Belge/Roche in Belgium, Bharat Starch Industries in India, and ADM, Cargill, and Tate & Lyle in the United States.

The list price of citric acid (anhydrous) has been in the range of 60–90 cents/lb, USP, 100-lb bags.

Itaconic acid



Itaconic acid (methylene succinic acid) is an unsaturated dibasic acid. It is a structurally substituted methacrylic acid, and therefore primarily used as a copolymer in acrylic or methacrylic resins. Because pure acrylic fibers are dye-resistant, it is necessary to include other components to make the fibers susceptible to dyes. An acrylic resin containing 5% itaconic acid offers superior properties in taking and holding printing inks and in bonding. In addition to its main application as a component of acrylic fibers, itaconic acid is used in detergents, food ingredients, and food shortenings.

Previously, itaconic acid was isolated from pyrolytic products of citric acid or produced by converting aconitic acid present in sugarcane juice. It is now produced on a commercial basis predominantly by direct fermentation of sugars. The biosynthesis of itaconic acid follows the metabolic sequence shown in Fig. 30.17.

Although both *Aspergillus itaconicus* and *Aspergillus terreus* are known producers of itaconic acid, the latter is superior and is used industrially with either surface (shallow-pan) or submerged (deep-tank) fermentation [35]. The medium contains a sugar source, corn steep liquor, ammonium sulfate, and mineral salts of calcium, zinc, magnesium, and copper. The fermentation, similar to that of citric acid, is very sensitive to concentrations of copper and iron. The fermentation is carried out at 39–42°C, pH of 2.0–4.0, and under vigorous agitation. Moderate, but continuous, aeration is critical. The batch cycle time is 3–6 days. The highest known product concentration is 180–200 g/L from a medium containing 30% sugar. Therefore, the yield of itaconic acid on sugar is typically 50–70%. The itaconic acid recovery scheme involves the following:

1. Acidification of itaconic precipitates, if present
2. Filtration to remove mycelium and other suspended solids
3. Activated carbon treatment (not necessary for industrial grade product)
4. Filtration to remove carbon
5. Evaporation and crystallization

If a high-purity acid is desired, further purification steps such as solvent extraction, ion exchange, and carbon decolorization are used.

The price of itaconic acid in 2002 was about $2/lb, and the world production is estimated to be about 20,000 t/year. The major producers include Cargill (using Pfizer technology) in the United States, San Yuan in Taiwan, Rhodia in France, and Merck in Germany.

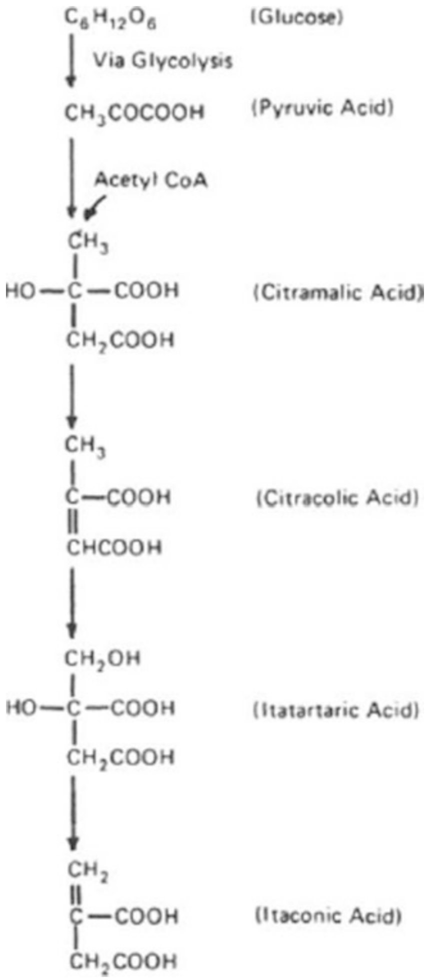
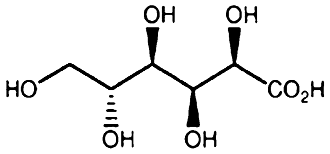
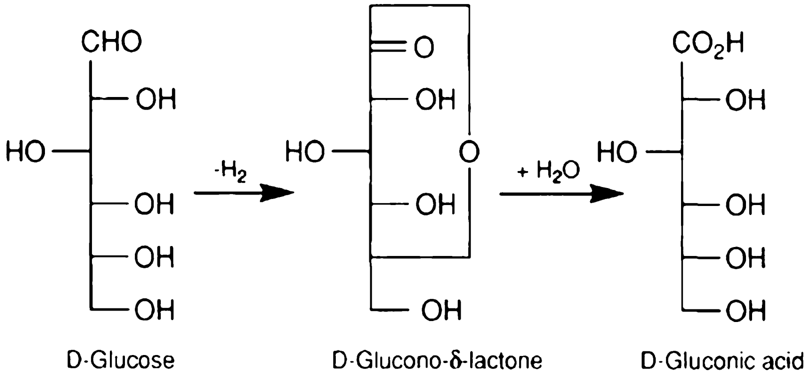


Fig. 30.17 Proposed metabolic sequence for biosynthesis of itaconic acid

Gluconic acid. Gluconic acid is produced by the oxidation of the aldehyde group of glucose to a carboxylic acid.





Gluconic acid may be prepared from glucose by oxidation with a hypochlorite solution, by electrolysis of a solution of sugar containing a measured amount of bromine, or by fermentation/enzymatic conversion (glucose oxidase, catalase) of glucose by fungi or bacteria (U.S. Patent 5,897,995). For economic reasons, the biological methods are now preferred.

Gluconic acid is marketed in the form of a 50% aqueous solution, calcium gluconate, sodium gluconate, and glucono-δ-lactone. Gluconic acid finds use in metal pickling, as an acidulant in foods, as a protein coagulant in tofu (soybean curd) manufacture, as a calcium sequestrant in detergent formulations, in the pharmaceutical area in mineral (calcium and iron) supplements, and as a cement viscosity modifier in the construction area. Calcium gluconate is widely used for oral and intravenous therapy. Sodium gluconate, a sequestering agent in neutral or alkaline solutions, finds use in the cleansing of glassware. Glucono-δ-lactone is used as a food flavor and an acidulant in making cheese and tofu, baking powders, and effervescent products.

Commonly used organisms for gluconic acid fermentation are *A. niger* and *Gluconobacter suboxydans*. The larger volume production uses the fungal process. Most of the gluconic acid produced from the *Gluconobacter* process is marketed as glucono-δ-lactone.

During gluconic acid fermentation, glucose is first oxidized (or, more correctly, dehydrogenated) to glucono-δ-lactone. This is carried out by glucose oxidase. Hydrogen peroxide is also produced in this step, but is decomposed by catalase. The fermentation can be by either surface or submerged culture, the latter more generally practiced in industry. Horizontally rotating fermentors have also been used. Calcium gluconate fermentation, in which calcium carbonate is used for neutralization of the product, is limited to an initial glucose concentration of approximately 15% because of the low solubility of calcium gluconate in water (4% at 30°C). Neutralization by sodium hydroxide instead allows the use of up to 35% glucose in the medium.

The recovery of calcium gluconate from fermentation broth involves the following:

1. Filtration to remove mycelium and other suspended solids
2. Carbon treatment for decolorization
3. Filtration to remove carbon
4. Evaporation to obtain a 15–20% calcium gluconate solution
5. Crystallization at a temperature just above 0°C

In sodium gluconate fermentation, the medium also contains corn steep liquor, urea, magnesium sulfate, and some phosphates. The pH is controlled above 6.5 by addition of sodium hydroxide. One to 1.5 volumes of air per volume of solution per minute (vvm) are supplied for efficient oxygenation and a high back-pressure, up to 30 psig, is desirable. The fermentation cycle is 2–3 days. Continuous fermentation is used in Japan to convert a 35% glucose solution to sodium gluconate with a yield higher than 95%. The continuous process doubles the productivity of the usual batch system.

Sodium gluconate can be recovered from the fermentation filtrate by concentrating it to 42–45% solids, adjusting the pH to 7.5 with sodium hydroxide, followed by drum-drying. In glucono-δ-lactone fermentation, *G. suboxydans* converts a 10% glucose solution to glucono-δ-lactone and free gluconic acid in about 3 days. Approximately 40% of the gluconic acid is in the form of glucono-δ-lactone. Aqueous solutions of gluconic acid are in equilibrium with glucono-δ-lactone. Crystals that separate out of a supersaturated solution below 30°C are predominantly of free gluconic acid; from 30 to 70°C, the crystals are principally glucono-δ-lactone; and above 70°C, they are mainly of γ-lactone.

Sodium gluconate sells for 30–70 cents/lb, depending on the liquid or dry form. Dry sodium gluconate is the main form of gluconic acid/gluconate consumed in the United States. The world market of gluconic acids and salts is about 60,000 t/year. The major producers include Glucona

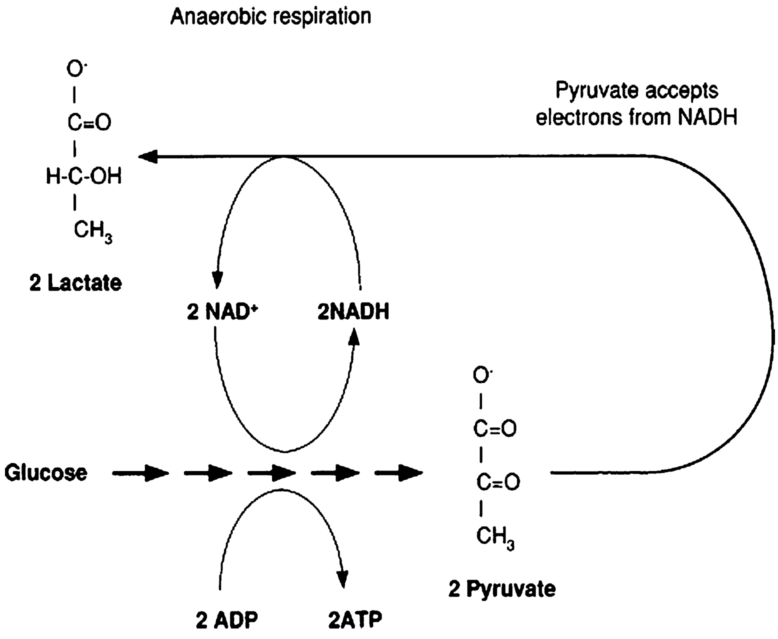
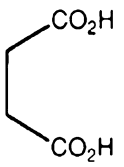


Fig. 30.18 Lactic acid from glucose fermentation

America (a subsidiary of Aveba B.A.) and PMP Fermentation Products (acquired by Fujisawa) in the United States; Benckister, Jungbunzlauer, and Roquette Freres in Europe; and Fujisawa and Kyowa Hakko in Japan. An isolated enzyme system (glucose oxidase-catalase) has recently become cost-competitive with the fermentation process and is therefore likely to expand its use.

Succinic acid

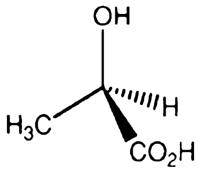


Succinic acid, as an intermediate in the chemical synthesis of 1,4-butanediol, tetrahydrofuran, and adipic acid, has a very large market potential [36]. It is a common intermediate in the metabolic pathway of several anaerobic microorganisms including *Anaerobiospirillum succiniproducens* and *Actinobacillus succinogenes.* However, succinate is produced by mixed-acid fermentations in low yields and concentrations along with several by-products. As the sole product of fermentation, it is possible to produce succinic acid at titers >100 g/L and yield of 1.2 moles per mole of glucose. In such a process, organisms use the phosphoenolpyruvate (PEP) carboxykinase pathway to make succinic acid. Other key enzymes include PEP carboxylase, malate dehydrogenase, fumarase, and fumarate dehydrogenase. Carbon dioxide concentration has been shown to regulate the levels of these enzymes for production of succinate. Carbon dioxide functions as an electron acceptor and modulates the flux of PEP.

An *E. coli* strain has also been engineered by overexpressing PEP carboxylase. Metabolic engineering of the strain was also performed in which genes encoding pyruvate:formate lyase and lactate dehydrogenase were inactivated. PEP flux was also increased through inactivation of the glucose-specific phosphotransferase system (PTS) system. Ideally, for minimizing salt formation in an acid fermentation, the pH should be lower than the pK_a. Otherwise, the conversion of the salt to the free acid adds significant costs in the final purification.

According to http://www.nrel.gov/docs/fy04osti/35523.pdf, there is a significant market opportunity for biobased C4 building block diacids. However, in order to be competitive with petrochemical-derived products, the fermentation cost needs to be below $0.25/lb. Although much progress has been made in the engineering of organisms for succinic acid production, at this point it remains commercially challenging.

Lactic acid



Lactic acid has been used as a food preservative and food-flavoring compound. Recent attention on lactic acid has been for its use in making polylactic acid (PLA), a biodegradable polymer. As a result, the market for lactic acid is rapidly growing. Under batch fermentation conditions, homolactic fermentative Lactic Acid Bacteria (LAB) make lactic acid as

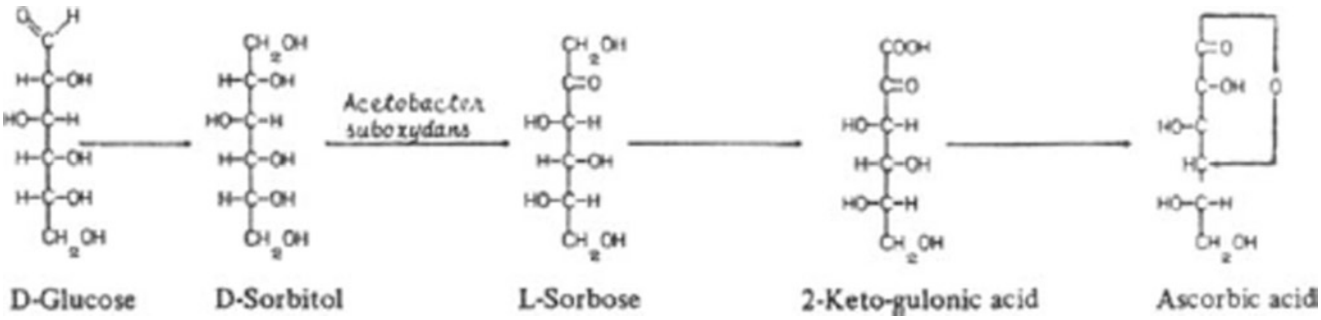


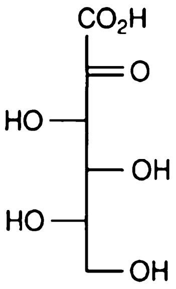
Fig. 30.19 Reactions leading to vitamin C from glucose

their endproduct. As shown in Fig. 30.18, 2 moles of lactate and 2 moles of ATP are formed per mole of glucose used. Although free lactic acid is preferred for most industrial processes, the anaerobic fermentation operates optimally at pH 5 where the salt of the acid is formed. To obtain lactic acid in its free form, the fermentation process must be carried out well below its pK_a of 3.87. This challenge has been met by cloning the bovine LDH gene into yeast, which has resulted in 1.19 mole lactate production per mole of glucose during fermentation at low pH. The fermentor titer reached is >100 g/L with a productivity of 0.8 g/L/h. A strain of *S. cerevisiae* overexpressing the lactate–proton symporter *JEN1* resulted in increased titer and yield.

Recovery of lactic acid is complicated by the high solubility of its salt. The traditional method of recovering calcium lactate is being replaced by membrane, electrodialysis, or ion-exchange purification methods. Producers such as Purac, Cargill, ADM, and Jiangxi sell lactic acid for $0.50–3.00 per kg depending on whether the product is industrial, feed, food, or pharma grade.

NatureWorks® LLC has set up a $300 million plant at Blair, NE, which is capable of producing about 140,000 t/year of polylactide polymers from corn sugar. It employs a fermentation process to produce two chiral isomers of lactic acid from glucose, which are then cracked to form three lactide isomers. The isomers are subsequently polymerized to polylactide.

2-Keto-L-gulonic acid



2-Keto-L-Gulonic acid (2-KLG) is the precursor for Vitamin C (L-Ascorbic acid). Vitamin C is used on a large scale as an antioxidant in food, animal feed, beverages, pharmaceutical formulations, and cosmetic applications. About one half of the vitamin C is used in vitamin supplements and multivitamin preparations, one quarter in food additives, 15% in beverages, and 10% in animal feed. When ascorbic acid is used in cosmetics, it is phosphorylated to prevent its oxidation.

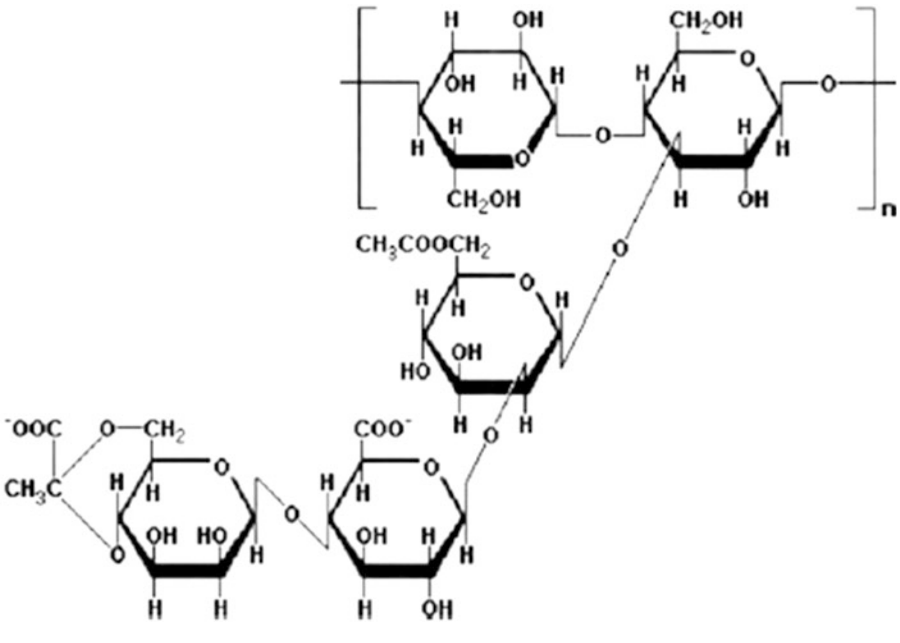
Figure 30.19 shows the reactions involved in the combined microbiological and chemical conversions of glucose, via sorbitol, to vitamin C, called the Reichstein and Grussner process. Sorbitol is made by catalytic hydrogenation of glucose and L-Sorbose is produced from sorbitol by the action of several species of bacteria, the most commonly used being *Acetobacter suboxydans*. Because this organism is very sensitive to nickel ions, it is important that the medium and fermentor be free of nickel. The medium normally consists of 100–200 g/L sorbitol, 2.5 g/L cornsteep liquor, and an antifoam agent such as soybean oil. The medium is sterilized and cooled to 30–35°C, about 2.5% inoculum is added, and the tank is aerated and sometimes stirred. This is the only step based on fermentation, and the step yield is 80–90% in 20–30 h. The chemical steps in the conversion of sorbose to ascorbic acid involve the preparation of the diacetone derivative, which is then oxidized, the acetone groups are removed, and the resultant 2-KLG is isomerized to the enediol with ring closure. It is believed that this chemical process currently operates at an overall yield of about 50% vitamin C on glucose.

Groups all around the world have pursued alternative vitamin C production technologies. A one-step fermentation process starting from glucose to produce 2-keto-L-gluconic acid was reported to be practiced in China, a microalgae process was developed in the United States and a two-stage fermentation process via the intermediate 2,5-diketo-D-gluconic acid (2,5-DKG) was developed in Japan. Genentech [37] succeeded in cloning the 2,5-DKG reductase gene from a *Corynebacter* sp. and expressing it in an *Erwina* sp. that naturally produces 2,5-DKG via the fermentative oxidation of glucose. The recombinant *Erwina* culture is thus capable of carrying out the complex oxidation and reduction steps to form 2-KLG in a single fermentation.

Genencor International, in partnership with Eastman Chemical Co. has developed such a single fermentation process using genetically engineered *Pantoea citrea*. To maximize the product yield on glucose, detailed analysis of the steps in the formation of 2-KLG was carried out. In the periplasm of this organism, glucose is first converted to gluconic acid by a membrane-bound PQQ-dependent glucose dehydrogenase. This is followed by the oxidation of gluconic acid to 2-keto-D-gluconate (2-KDG) by a cytochrome C coupled enzyme, gluconate dehydrogenase. In a third reaction, 2-KDG is further oxidized to 2,5-DKG by 2KDG dehydrogenase, another cytochrome C coupled enzyme. Purification, characterization, and determination of the enzyme structure of 2,5-DKG reductase have added significantly to the understanding and development of this production process.

The major manufacturers include DSM, BASF, and four Chinese producers, viz. Northeast Pharmaceutical, North China Pharmaceutical, Shijiazhong Pharmaceutical, and Jiangshan Pharmaceutical. The current world supply of ascorbic acid is more than 80,000 metric tons per year with annual revenues in excess of US $700 million.

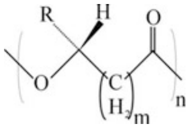
Xanthan gum



Xanthan gum is an anionic polymer of β-(1,4)-D-glucopyranose glucan with side chains of 1,3-α-D-mannopyranose and (1,2)-β-D-glucuronic acid residues [38]. It is produced naturally by bacteria to enable them to stick to plants. The negatively charged carboxyl groups on the side chains generate viscosity when the xanthan gum is mixed with water. The glucan backbone is protected by the side chains, thus making it relatively stable to acids, alkalis, and cellulase enzymes. Xanthan gum is used as a thickener, stabilizer, emulsifier, and foaming agent in products such as salad dressings, cosmetics, pharmaceuticals, paints, lubricants, and ice cream. Its key property is high viscosity at low shear and thinning character at high shear. Xanthan gum is not affected by ionic strength over a wide pH, shear or temperature range.

Xanthan gum is produced commercially via aerobic submerged fermentation of sugar by *Xanthomonas campestris*. Different strains or fermentation conditions produce differing degrees of acetylation, pyruvylation, and hence functionalities. The production strain is grown in medium containing carbohydrate, nitrogen, and salts. The production process of xanthan gum is energy intensive and costly, and product titer and productivity are relatively low due to high viscosity. In addition, the broth is difficult to filter due to cells in the broth. After fermentation, the broth is sterilized to inactivate cells; the product is precipitated or coagulated, separated by centrifuges, dried, and milled. About 20,000 t of xanthan are produced each year with a market value more than US $160 million.

Poly(3-hydroxyalkanoates)



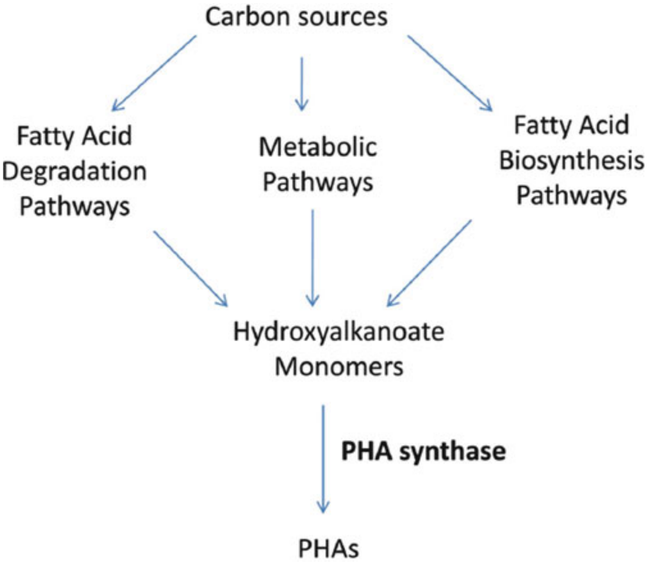


Fig. 30.20 Pathway for synthesis of poly(3-hydroxyalkanoates) using a variety of substrates

Among a wide variety of biobased polymers [39], poly(3-hydroxyalkanoates) (PHAs) also known as microbially produced polyesters are diverse in terms of constituents and biosynthetic pathways, and therefore desired as starting materials for a variety of conventional and thermo-plastics applications. Some microorganisms produce polyesters naturally, e.g., Ralstonia eutropha formally known as *Alcaligenes eutrophus*, a gram negative bacteria that produces PHAs as inclusion bodies or granules in the cytoplasm and serve as storage of carbon and energy. Natural organisms accumulate PHAs under excess carbon or some nutrient limitation, and use them later under carbon starvation. But engineered micro-organisms can accumulate up to 80–95% of the cell's dry weight, and large scale fermentors can grow >100 g/L dry cell weight. PHA is the most common polyester made in nature, and poly-3-hydroxybutyrate (PHB) was the first to be discovered. Other PHAs have since been discovered, including those containing hydroxyvalerate units, as in poly-3-hydroxyvalerate (PHV). Because the structure and monomeric composition of PHAs determine the applications for each type of polymer, a variety of polymers have been synthesized by co-feeding of various substrates or by metabolic engineering of the production organism (see Fig. 30.20).

Production organisms are fed carbon sources like glucose, sucrose, oils, propionic acid, etc. After fermentation, the cells are lysed, the cell debris is separated, and the PHA is extracted with a solvent, precipitated from water as powder, purified, and converted to pellets. Production of PHA could take up to 5 kg of raw material per kg product, and genetic improvements in microorganisms are necessary for higher PHA yield, use of cheaper feedstocks, and lower capital.

PHA commercial production has been considered by several companies like W. R. Grace, ICI, DSM, Metabolix, Monsanto, Proctor & Gamble (P&G), Kaneka, Meridian, Tianan, Tianjin, and others. According to P&G/Kaneka, PHA/PHB (Nodax^{TM}) production cost has to be <1.5 USD per kg for high volume applications, and alkaline digestibility and flushability are convenience factors of interest to the production of single-use consumer goods.

Alcohols

The organic chemicals that fall into this category and can be produced by fermentation include ethanol, 1,3-propanediol, butanol, acetone, 2,3-butanediol, and glycerol. Butanol and acetone have been produced industrially by fermentation, but currently for economic reasons, chemical synthesis is the manufacturing method of choice. However, as price and availability of ethylene and propylene as feedstocks for the synthetic processes become subjects of concern, there is a renewed interest in examining the fermentation processes. As the cost of crude oil continues to go up while the price of renewable resource such as agricultural crops remains relatively stable, there has been an increasing interest in producing chemical feedstocks from renewable resources by biological means.

1,3-Propanediol

1,3-Propanediol is a monomer with potential utility in the production of polyester fibers and in the manufacture of polyurethanes and cyclic compounds. A variety of chemical routes to 1,3-propanediol are known. For example, (1) ethylene oxide may be converted to 1,3-propanediol over a catalyst in the presence of phosphine, water, carbon monoxide, hydrogen, and an acid; (2) a catalytic solution phase hydration of acrolein followed by reduction; or (3) from glycerol, reacted in the presence of carbon monoxide and hydrogen over catalysts having atoms from group VIII of the periodic table. Although it is possible to generate 1,3-propanediol by these methods, they are capital intensive and/or generate waste streams that contain environmental pollutants.

The biological production of 1,3-propanediol from glycerol has been known for a number of years [11]. The use of natural organisms (e.g., enteric bacteria such as *K. pneumoniae*, *Citrobacter freundi*, and Clostridia, such as *C. butyricum*) to produce 1,3-propanediol from glycerol has been well studied. Continued optimization of the basic anaerobic glycerol fermentation process has produced promising results. A number of batch or fed-batch

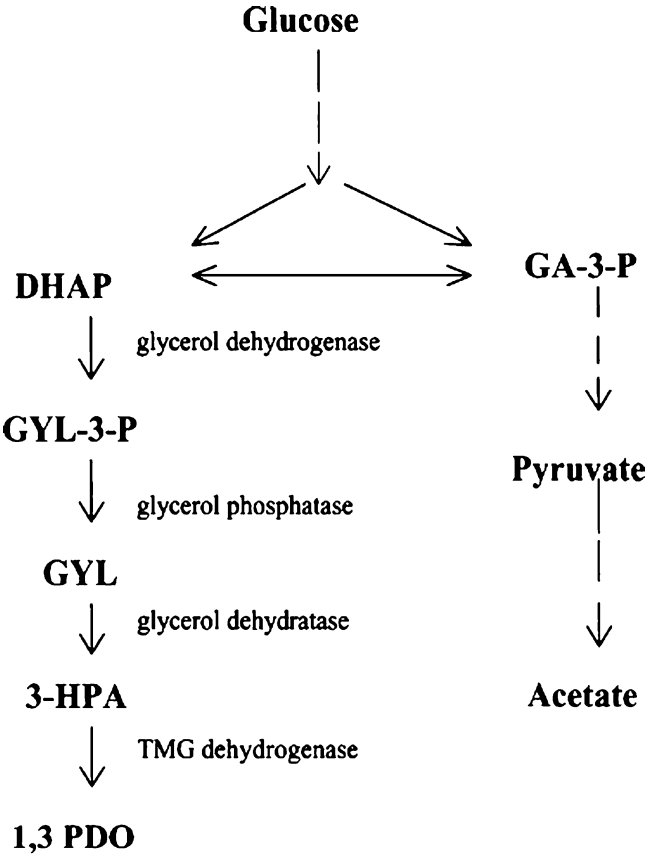


Fig. 30.21 1,3-PDO pathway from glucose

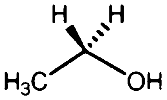
fermentations using *K. pneumoniae* or *C. butyricum* have produced titers of 50–75 g/L and a yield of 0.44–0.69 mol/mol 1,3-propanediol from glycerol. However, the complete conversion of glycerol to 1,3-propanediol is not possible due to the requirement of an additional reducing power in the form of NAD(P)H.

Neither the chemical nor the biological methods described earlier for the production of 1,3-propanediol are well suited for industrial scale. The chemical processes are energy intensive and the biological processes require the expensive starting material, glycerol. A more desirable process would be to develop a microorganism that would have the ability to convert basic carbon sources such as carbohydrates or sugars to the desired 1,3-propanediol. Such a single organism approach developed by Genencor International and DuPont overcomes this problem by generating reducing cofactors from the glucose conversion to CO_2. The ability to control both carbon and energy flow in the single organism allows for the more efficient use of the input carbon source.

As shown in Fig. 30.21, the conversion of glucose to 1,3-propanediol requires the combination of two natural pathways: glucose to glycerol and glycerol to 1,3-propanediol. The two enzymes involved in the conversion of glycerol to 1,3-propanediol have been cloned and characterized from several organisms, viz. *Klebsiella*, *Citrobacter*, and *Clostridium*. The first enzyme in the pathway is glycerol dehydratase. The dehydratase has been shown to undergo catalytic inactivation and requires the addition of a reactivation complex of two additional proteins. The second enzyme in the pathway is a NAD-linked tri-methylene glycol (TMG) dehydrogenase. Both of these enzymes have been cloned and expressed in *E. coli*. The pathways for the production of glycerol are found in the yeast, *S. cerevisiae* that produces glycerol from the glycolytic intermediate dihydroxyacetone-3-phosphate using two enzymes: dihydroxyacetone-3-phosphate dehydrogenase and glycerol-3-phosphate phosphatase. To construct a single organism to produce 1,3-propanediol from glucose, one could clone the glycerol pathway into a natural 1,3-propanediol producer, or the 1,3-propanediol pathway into a natural glycerol producer. Although either of these approaches seems simple and direct, there are problems involving natural regulation of the pathways. However, building both pathways into *E. coli* is advantageous as it is the most completely studied organism. It provides a rich set of genetic tools (sequenced genome, vectors, promoters, well-characterized metabolism, and physiology) and a large number of metabolic mutants that have been constructed and analyzed [40]. Moreover, *E. coli* has been used in large-scale fermentation for production of industrial and healthcare products, does not naturally overproduce glycerol or 1,3-propanediol, and thus has a strong likelihood of no natural regulation to overcome. Modifications of more than 70 *E. coli* genes were analyzed and the final production organism had 18 genes modified. Some of these genes were regulatory elements that indirectly affected the function of a few hundred other genes. The optimized strain of *E. coli* produced 135 g/L of 1,3-propanediol, at 3.5 g/L/h, and 51% weight yield on glucose.

The joint venture between DuPont and Tate & Lyle has demonstrated the large-scale feasibility to produce 1,3-propanediol by fermentation (bio-PDO^{TM}). The process uses the genetically engineered organism developed jointly by former Genencor and DuPont. The joint venture has announced that in 2006 they will commercially produce bio-PDO™ from corn to replace petrochemicals in the production of their polymer Sorona$^{®}$, for many end-use applications, including fibers. The bio-PDO™ will be produced at a new plant in Loudon, TN, USA. The bioprocess for making 1,3-propanediol is claimed to require 40% less capital and has a 25% lower operating cost than the chemical process for making the same product. Because of a different impurity profile, the polymer made from the biological-based 1,3-propanediol has improved properties and actually creates a superior product compared to the chemically based 1,3-propanediol (http://www.dupont.com/sorona/faqs.html).

Ethanol



Ethanol is a versatile chemical being used in industrial solvents, thinners, detergents, toiletries, cosmetics, and pharmaceuticals. It is also used as an intermediate for manufacturing other chemicals such as glycol ethers, ethyl chloride, amines, ethyl acetate, vinegar, and acetaldehyde. With the ever-increasing price and dwindling supply of crude oil, ethanol fermented from grains and other renewable organic resources is becoming an economical fuel supplement. The U.S. fermentation ethanol capacity is now about four billion gallons a year, a result of efforts by the ethanol industry to expand its use in fuels, as a replacement for the oxygenate, methyl *tert*-butyl ether.

Ethanol is produced by fermentation from sugar-containing materials such as grain products, fruits, molasses, whey, and sulfite waste liquor. Yeast, particularly strains of *S. cerevisiae*, is almost exclusively used in industrial ethanol fermentation. It tolerates ethanol concentrations up to about 20% (by volume), and under anaerobic conditions converts over 85% of the available carbohydrates to ethanol and carbon dioxide. Air or oxygen suppresses the formation of ethanol from sugar (the Pasteur effect). Under aerobic conditions, a major portion of the carbohydrates goes towards cell growth. Ethanol is formed via glycolysis (the Embden–Meyerhof–Parnas pathway). The overall reaction starting from glucose can be written as follows:

$$C_6H_{12}O_6 \rightarrow 2C_2H_5OH + 2CO_2 + 18\ \text{kcal}$$

The ethanol yield on glucose is thus 51% by weight. Because the carbohydrate is also used for cell growth and respiration, the overall yield of ethanol from total carbohydrate consumed is about 42–46 wt.% on a glucose equivalent basis.

Ethanol fermentation can be conducted on any carbohydrate-rich substrate. Molasses, which is the waste mother liquor that remains after the crystallization of sucrose in sugar mill operations, is widely used. Blackstrap molasses contains 35–40% sucrose and 15–20% invert sugars (glucose and fructose). High-test molasses contains 22–27% sucrose and 50–55% invert sugars. Most of the blackstrap molasses do not require the addition of other nutrients for ethanol fermentation. However, high-test molasses requires ammonium sulfate and other salts such as phosphates. In the molasses process, blackstrap or high-test molasses is charged into a mixing tank and diluted with warm water to give a sugar concentration of 15–20%. The pH is adjusted to 4–5, the mash pasteurized, cooled, and charged into fermentor tanks, and 1–5% yeast inoculum is added. The fermentation is carried out nonaseptically at 23–32°C. Antibiotics may be added to control possible bacterial contaminations. Because the overall reaction is exothermic, cooling is necessary. The fermentation takes 28–72 h to produce an ethanol concentration of 8–10%. Carbon dioxide normally is vented, but if it is to be recovered, then the vent gas is scrubbed with water to remove entrained ethanol and then purified using activated carbon. The bubbling due to carbon dioxide generation by fermentation could provide sufficient agitation for small tanks. Mechanical agitation may be necessary for large fermentors. The fermentation may be conducted batchwise or continuously, with or without recycling yeast. Although continuous fermentation and/or cell recycle can significantly improve productivity and thus reduce required capital investment, it may have only a limited impact on lowering product costs in as much as a significant portion of the total cost comes from raw materials.

Ethanol can also be produced by fermentation of starch, whey, sulfite waste liquor, cellulose, and hemicellulose. Potato or tapioca can also be converted to ethanol. Grain fermentations [41] require additional pretreatment because yeast cannot metabolize starch directly. The grain (corn, sorghum, barley, wheat, rice, or rye) is ground and heated in an aqueous slurry to gelatinize or solubilize the starch. Starch-liquefying enzymes (thermostable alpha amylase) are added in this step. The liquefied starch is then cooled to 60–65°C. Yeast is then added to carry out fermentation simultaneously with saccharification by amyloglucosidase (or glucoamylase). The subsequent fermentation and refining procedures are the same as those using molasses as raw materials.

The grain-based ethanol process, also called dry milling was developed and optimized for the beverage industry utilizing starch-based feedstock and energy-intensive processes. Such processes tend to produce some nonfermentable sugars during cooking of starch. Salts added for the needed pH adjustments of the enzymatic treatments. Both of these result in extra process costs. A granular starch hydrolyzing enzyme (GSHE) technology is likely to replace the traditional energy intensive process of liquefaction/saccharification (http://www.biorefineryworkshop.com/presentations/Dean.pdf). The GSHE is a new enzyme product to reduce energy usage for the production of inexpensive glucose from granular starch, which replaces the energy-intensive liquefaction/saccharification. The ethanol industry needs continued energy, cost, and conversion efficiency improvements. An ethanol production process that has the following key features is much desired.

- Energy savings—elimination of jet cooking
- Capacity increase—high solids and increased ethanol yield

- Reduction of yeast growth inhibitors—high glucose concentration, Maillard reaction products, and so on.
- Operational cost reduction—labor, time, and chemicals
- Value-added co-product, Dried Distiller's Grains and Solubles—higher protein content
- Process simplification—fewer unit operations

GSHE-like novel enzyme-processing technologies with reduced total operating cost will provide a dramatic acceleration in the timelines for the establishment of the biobased economy. This developing technology will also have a major positive impact on the life cycle costs of biobased products and address the growing concerns among policy makers and environmentalists.

The GSHE is a blend of enzymes with synergistic activity on granular starch. The key enzymes of the GSHE blend include acid stable alpha amylases and glucoamylases that "drill" holes in the starch granules. Thus the GSHE is capable of hydrolyzing insoluble granular (uncooked) starch into fermentable sugars by enabling depolymerization of starch to glucose without the need for the energy-intensive jet-cooking step, in a Simultaneous Saccharification and Fermentation (SSF) process for alcohol production. Genencor International has achieved high level of expression of these enzymes in industrial production strains resulting in an acceptable cost structure.

Research to improve ethanol fermentation has also focused on the development of a high solids, continuous feeding process as well as improved yeast strains. Other related developments include processes for the use of biomass feedstocks, such as cellulosic waste material [42–46]. The former Genencor and NREL collaboration funded by the U.S. DOE has (a) put in place the tools required for continual improvement of biomass-derived cellulases, (b) built a suite of enzymes with enhanced thermostability and improved specific performance at elevated temperatures, (c) developed an improved production strain for the enhanced cellulase components and demonstrated an enhanced production process. Therefore, plans are in place to provide a developmental product(s) in support of continued industry development.

Developments in genetic engineering have made possible the development of new organisms (yeasts and bacteria) that can survive in higher concentrations of ethanol, tolerate higher sugar concentrations, grow at higher temperatures, and utilize starch or cellulose feedstock directly (http://www.farmfoundation.org/projects/03-18_biofuels.pdf).

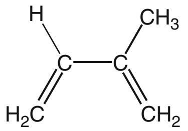
Ingram's [44] pathway engineering program has created a production organism patented by the University of Florida, with improved efficiency of ethanol production from C5 and C6 sugars (xylose, arabinose, galactose, mannose, glucose). Gelunol holds the rights to the patented engineered bacteria and is building a large-scale biomass-to-ethanol plant in Jennings, LA. The plant is expected to be operational in 2007 and will use bagasse waste from sugarcane to produce up to 30 million gallons a year of ethanol. According to the USDA and DOE, production of one billion ton of biomass per year is possible on a sustainable basis. This scale of biomass has the conversion potential of producing 100 billion gallons of ethanol per year.

In order to prevent the diversion of industrial alcohol to potable uses, it is denatured by the addition of some material that renders the treated alcohol unfit for use as a beverage. A fuel mixture such as 10% alcohol and 90% gasoline has become the most important use of fermentation ethanol. Ethanol from grain fermentations has been made competitive as a liquid fuel in the United States because of improvements in technology, especially in the area of efficient enzymes and energy usage in production plants, and various government subsidies designed to reduce the nation's dependence on imported foreign oil. Other countries in Europe as well as Brazil, India, and China also have substantial production of fermentation ethanol for use as fuel. In the United States, ADM, Tate & Lyle, and Cargill are the biggest among a large number of producers (for ethanol recovery, see "Recovery section").

Hydrocarbons

Hydrocarbons that can be produced by fermentation are many, from hydrogen and methane typically produced in anaerobic waste treatment, to ethylene and isoprene produced naturally by a variety of micro- and macro-organisms (plants, trees, animals). Very recent advancement of pathway engineering of isoprene and farnesene production systems has drawn significant attention.

Isoprene



Isoprene is an important commodity chemical used in a wide range of industrial applications ranging from the production of synthetic rubber for tires and coatings to use in adhesives and development of specialty elastomers. Styrene-isoprene-styrene block copolymers form an essential part of most hot melt pressure sensitive adhesive formulations and *cis*-polyisoprene or synthetic rubber is used in tire manufacture. All of the world's isoprene is produced from petroleum-derived feedstocks and is subject to volatility in pricing and supply linked to oil.

A first of its kind biobased process to produce isoprene (BioIsoprene™) from renewable raw materials is in

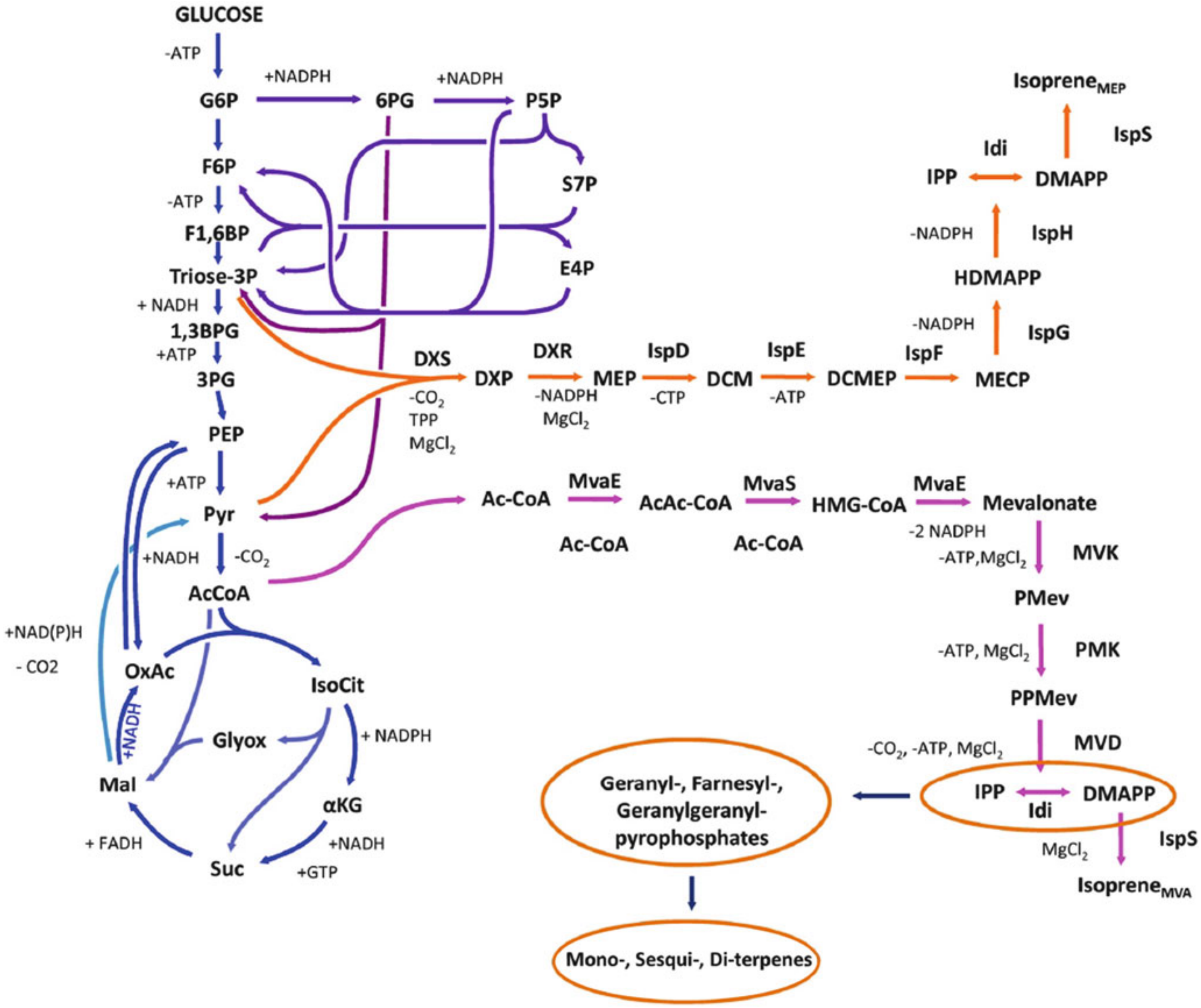


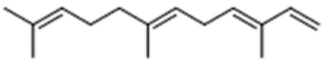
Fig. 30.22 Isoprene production pathways: mevalonic acid (MVA) and methyl erythritol phosphate (MEP)

development by former Genencor and The Goodyear Tire & Rubber Company. Production system for BioIsoprene™ is based on microbial fermentation of renewable sugars converted to dimethylallyl pyrophosphate (DMAPP), the key substrate of isoprene synthase, the key plant enzyme engineered into the production organism. DMAPP and its isoform, isoprenyl pyrophosphate (IPP), are essential compounds in all living organisms and are produced via one of two known pathways, the mevalonic acid (MVA) pathway present in eukaryotes and some prokaryotes, and the 5-methyl erythritol phosphate pathway (MEP, also known as deoxyxylulose phosphate [DXP] pathway) present in prokaryotes and plants [47] (Fig. 30.22). Carbon flux into both pathways stems from intermediates derived from sugar metabolism via central metabolism, acetyl-CoA for MVA; pyruvate and glyceraldehyde 3-phosphate for MEP. These pathways are amenable to molecular biology manipulations and metabolic pathway engineering in both prokaryotes (*E. coli*) and eukaryotes (*S. cerevisiae*). Both biosynthetic isoprenoid pathways, MEP and MVA, are being optimized towards this end. The key to successful process is engineering a microbial cell factory that is streamlined in physiology to drive sugar conversion through engineered biosynthetic pathways to produce isoprene at near theoretical yields. Unlike other biobased systems to produce biochemicals, BioIsoprene™ is produced as a gas-phase product and released into the off gas of the fermentor without any noticeable negative physical or inhibitory impacts on the production organism. The product comes out of the fermentor continuously in the gas phase and is recovered from the fermentation off gas in a continuous process. Current state-of-the-art technology has resulted in production, recovery, polymerization, and manufacture of tires with the isoprene component produced via fermentation. Polymer-grade BioIsoprene™ is recovered from the integrated process. Potential benefits of the gas-phase nature of the product include: (a) reduction and/or elimination of feedback inhibition by the isoprene product on further synthesis; (b) efficient recovery and purification of polymer-grade BioIsoprene™ from the fermentation; and (c) possibility to use crude, low cost feedstocks.

The development of BioIsoprene™ represents a major achievement for industrial biotechnology because it has the potential to enable production of isoprene from renewable raw materials to deliver commercial quantities of a basic C5 hydrocarbon that, in principle, can be used as a feedstock for a large number of value-added products. One of these products is a synthetic *cis*-polyisoprene (synthetic rubber), which for decades has been recognized as a suitable replacement for natural rubber in many applications. Additionally, there is potential for creating polyisoprenes with improved material properties. The BioIsoprene™ process offers a real possibility for obtaining the quantities of low-cost isoprene needed to produce a meaningful large-volume alternative to Hevea natural rubber and petroleum-derived isoprene.

BioIsoprene™ will have broader commercial applications beyond the biochemical uses of isoprene in synthetic rubber, adhesives, and specialty elastomers. As a C5 hydrocarbon, BioIsoprene™ has inherent fuel properties and represents a key biobased intermediate that can be converted to a drop-in transportation fuel additive using chemical catalysis to C10 and C15 biobased hydrocarbon fuels, thus addressing performance gasoline, jet fuel, and biodiesel markets.

Farnesene



Amyris has engineered yeasts that convert sugars into a hydrocarbon molecule called farnesene [48], a sesquiterpene which can be processed into cosmetics, lubricants, and diesel and jet fuel. If hydrogenated, the hydrocarbon can be turned into a diesel fuel that burns cleaner than conventional diesel. Amyris' process uses engineered yeast to convert sugar into beta-farnesene, in oil form, easy to separate from aqueous fermentation broth.

As shown in Isoprene production pathway (Fig. 30.22), farnesene, like many isoprenoids, is assembled using the five-carbon monomer IPP and its isomer DMAPP. IPP and DMAPP could also be condensed by IPP synthases into geranyl pyrophosphate (GPP, C10), farnesyl pyrophosphate (FPP, C15), or geranylgeranyl pyrophosphate (GGPP, C20). Terpene synthases convert these long chain pyrophosphate molecules into branched chain or cyclic alkenes, e.g., farnesene. Isoprenoid tailoring enzymes can then oxidize the alkenes into alcohols or reduce them to alkanes. Alternatively, isoprenyl phosphatases or IPP synthase mutants can generate long-chain alcohols such as farnesol, geraniol, and geranylgeraniol from GPP, FPP, or GGPP, respectively.

While the Amyris industrial synthetic biology platform can use a wide variety of feed stocks, their initial focus is on sugarcane due to its abundance, generally low cost, and favorable carbon footprint. That's why Amyris' initial production efforts are in Brazil, leveraging the infrastructure of existing sugar and ethanol mills. However, their success depends on increasing the sugar conversion efficiency (15–16% kg farnesene per kg sugar reported). Other success factors include productivity (less than 1 g/L/h reported), farnesene concentration in broth (about 100 g/L reported), and large-scale operation (200,000 L size fermentors reported).

Advanced biofuels

Ethanol has become a valuable additive to liquid transportation fuels. During 2010, 13 billion gallons were produced in the United States with most of this being fermented using grains as starting material. Ethanol provides high octane, but the energy density is only about 70% that of traditional gasoline. The hydroscopic property of ethanol also renders it difficult to transport in existing pipelines without resulting in corrosion. For these and other reasons, there has been increased interest in producing fuels besides ethanol from renewable resources [49, 50].

Higher Alcohols

Production of higher alcohols can be further broken out into two key groups: (1) fermentative pathways to isopropanol and *n*-butanol and (2) nonfermentative pathways to a larger diversity of branched chain alcohols.

Fermentative pathways to alcohols are typically based upon the acetone-butanol-ethanol (ABE) fermentation classically found in Clostridia. Acetone can be reduced in a single step to form isopropanol. For economic reasons, native production of these compounds in Clostridia is not currently practiced. But through synthetic biology and metabolic engineering, it has been possible to optimize heterologous pathways to maximize yield and productivity in industrial microorganisms. Considerable work has focused especially on production of butanol especially in *E. coli* and *S. cerevisiae*. Pathways for three different butanol species are shown in Fig. 30.23a. For *n*-butanol, two molecules of acetyl-CoA are condensed to form acetoacetyl-CoA, which then undergoes reduction and dehydration to generate *n*-butanol. 2-butanol is formed through acetolactate by decarboxylation, dehydration, and reduction. Isobutanol is formed by a nonfermentative pathway through 2-keto-isovalerate.

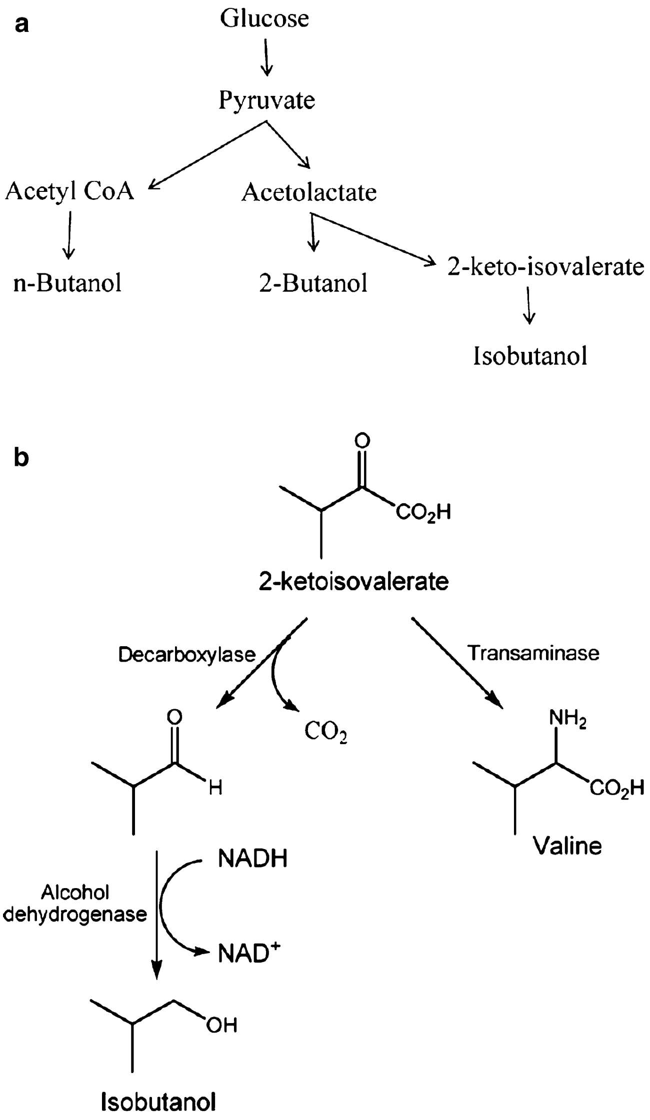


Fig. 30.23 (**a**) Pathways for three different butanol species, (**b**) pathway for the production of isobutanol

Nonfermentative pathways can result in a wider range of higher alcohols. The key steps in all these pathways are the decarboxylation of 2-keto acids and reduction of the resulting aldehyde to the corresponding alcohol. Many amino acid biosynthesis pathways have these 2-keto-acids as intermediates. As an example, 2-keto-isovalerate is the product of the penultimate step in valine biosynthesis. But instead of being transaminated to valine, it can also be decarboxylated and reduced to make isobutanol. This is shown in Fig. 30.23b. There is a wide range of alcohols that can be formed from these naturally occurring 2-keto acids. However, the range of alcohols that can be formed has been increased even further by engineering completely new pathways to make nonnatural 2-keto acids [51]. These then undergo the same decarboxylation and reduction to form alcohols. Once again, enzymes from the branched chain amino acid pathways were used. The leuABCD operon of *E. coli* acts upon 2-keto-isovalerate to add a carbon leading to 2-keto-isocaproate and then to leucine. Two enzymes were protein engineered to allow conversion of nonnative substrates. Making changes to leuA and 2-keto-isovalerate decarboxylase resulted in the formation of a large number of nonnative alcohols with carbon range of C5–C8. This work indicates that the types of molecules that can be produced are not limited by what is observed by native biosynthetic reactions.

Biodiesel Production

Biodiesel production from plant oils has been practiced. Natural triacylglycerols in oils undergo transesterification with methanol or ethanol to make fatty acid methyl and ethyl esters (FAME and FAEE). Think of modified diesel vehicles whose exhaust smells like French fries. However, the availability of plant oils limits the potential use as fuels. And the transesterification reaction can be costly. Acyltransferases can be used to perform the reaction in vivo by feeding fatty acids in strains making ethanol [52]. Conversely, deregulation of fatty acid biosynthesis has allowed for FAEE formation with alcohol feeding, but without fatty acid feeding [53]. It is only a matter of time before complete FAEE synthesis is made directly from carbohydrates.

Isoprenoid-based biofuels

Isoprenoids are naturally occurring hydrocarbons that carry out a number of biological functions. For instance, electron transport is mediated by ubiquinones and membrane structure and function are controlled by sterols [54]. Much of the recent work on the production of isoprenoids was initiated by the production of the antimalarial drug, artemisinin [55]. The key intermediates leading to all isoprenoids are the C5 isomers IPP and DMAPP. These are formed by two different pathways, the MVA pathway from acetyl-CoA and the MEP(DXP) pathway from pyruvate and glyceraldehye-3-phosphate. These can be joined together in many ways by terpene synthases to generate C10, C15, and higher hydrocarbons. Phosphatases can directly convert many of these compounds to fuel-like molecules.

Carbon sources for biofuels production

Simple sugars have been the starting point for first-generation ethanol production. Sucrose from sugarcane has driven Brazilian production while starch from grains has been enzymatically converted to glucose in the rest of the world. However, this typically uses only a portion of the plant material for incorporation into product. Agricultural residues or dedicated energy crops could greatly increase the amount of available substrate for conversion to biofuels. But the use of these lignocellulosic sources is not without issues [56]. The presence of lignin, a heterogeneous polymer of aromatic alcohols, can interfere with the hydrolysis of the desired carbohydrates. A large number of pretreatment processes have been developed to separate or modify the lignin to make the carbohydrates more accessible. Pretreatments are also required to modify the structure of recalcitrant cellulose to make it more amenable to enzymatic digestion. These processes include physical, thermal, and chemical treatments. Cellulose is then converted to fermentable glucose, just as in the starch-based processes. The final component of lignocelluloses is hemicelluloses. Hemicellulose is another heterogeneous polymer made up of a number of different sugar monomers. The most abundant of these is xylose. Different from six-carbon glucose, xylose is a five-carbon sugar that is metabolized by different pathways than glucose. Metabolic engineering has been employed extensively to combine simultaneous fermentation of both hexoses and pentoses to ethanol in a variety of organisms including *Zymomonas mobilis* [57], *S. cerevisiae* [58], and *E. coli* [59]. These strains are still being improved and in addition to ethanol, can be the starting point for the generation of advanced biofuels. Adding new pathway elements, modifying transport mechanisms, and alleviating native control hierarchy have all been utilized in the building of these strains.

Next generation processes: fuels from carbon dioxide

The use of sucrose from sugarcane, starch from corn, and even lignocellulosic biomass are all examples of utilizing chemical energy produced by photosynthetic organisms from solar energy and CO_2. Producing advanced biofuels directly from these photosynthetic organisms could ultimately result in the largest reduction of atmospheric CO_2 [60]. The most advanced systems are those producing biodiesel directly from photosynthetic microalgae. Some strains can generate lipids, used as energy storage compounds, at nearly 40% of the total dry biomass. These lipids can be extracted and undergo the same transesterification process as from plant-derived oils.

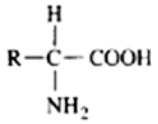
Photosynthetic organisms such as microalgae generate energy and reducing power from light. Photons are captured by specialized light harvesting complex proteins and the energy is used to split water into protons, electrons, and oxygen. Electrons are passed along an electron transport chain and eventually to NADPH. Protons are released across an organelle membrane creating a proton gradient used to generate ATP. These biological energy compounds are then used in the Calvin cycle to fix CO_2. CO_2 is added to ribulose-1,5-bis-phosphate, which is then split into two molecules of 3-phosphoglycerate, a key 3-carbon metabolic intermediate.

Metabolic engineering approaches are being applied to microalgae just as they have been applied to other organisms to generate biofuels. Here, focus has been in two main areas: (1) increasing oil production and (2) increasing photosynthetic efficiency. Most probably nonintuitive, reducing the amount of light harvesting complex proteins has led to increases in photosynthetic efficiency.

Photosynthetic organisms lead to a new problem in large-scale production. In order to function, the culture needs both abundant light and CO_2. Open ponds are one obvious choice for cultivation. However, closed bioreactor designs have numerous advantages over open ponds. These include, among others, high surface area per unit volume, ability to keep out contaminating organisms, and the more efficient use of water.

Amino Acids

Amino acids, in general, can be represented by the formula:



$$R-\overset{\displaystyle H}{\underset{\displaystyle NH_2}{\overset{|}{\underset{|}{C}}}}-COOH$$

Because the amino group is on the α-carbon, the amino acids with this general formula are known as α-amino acids. The α-carbon atom becomes asymmetric when R is not an H atom. Naturally occurring amino acids have an L-configuration. Amino acids are the building blocks of proteins, and the elementary composition of most proteins is similar; the approximate percentages are:

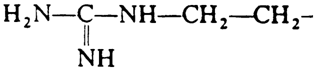
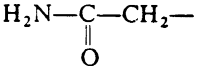
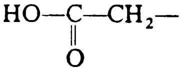
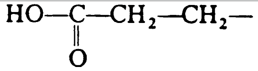
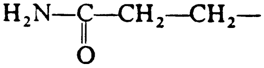
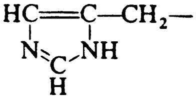
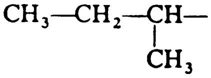
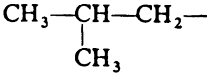
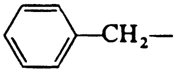
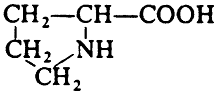
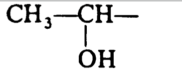
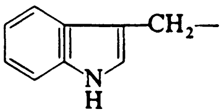
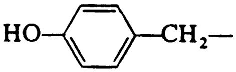
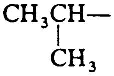
$$C = 50 - 55, \quad N = 15 - 18,$$

$$H = 6 - 8, \quad S = 0 - 4, \quad O = 20 - 23$$

Table 30.4 gives the structure of R, molecular weight, and elementary composition of each of the 20 amino acids commonly found in proteins.

Of the natural amino acids, only methionine is manufactured chemically as the racemic mixture. All other natural types are produced by fermentation as the L-enantiomer. Commercially, the most common amino acids manufactured are the L-forms of glutamic acid, mostly mono-sodium glutamate (MSG), lysine, phenylalanine, aspartic acid, threonine, tryptophan, arginine, citrulline, glutamine, histidine, isoleucine, leucine, ornithine, proline, tyrosine, and cysteine [61]. The biological methods vary (e.g., fermentation, extraction from natural sources, and enzymatic synthesis). Amino acids obtained from purified proteins are derived from chemical or enzymatic hydrolysis. They can also be isolated from industrial by-products; extracted from plant or animal tissues; or synthesized by

Table 30.4 Twenty common amino acids

Amino acids	R⁻	Elemental Composition (%wt.) M.W.	C	H	O	N	S
Alanine	$CH_3\text{—}$	89	40	8	36	16	0
Arginine	 $H_2N\text{—}\underset{\underset{NH}{\|}}{C}\text{—}NH\text{—}CH_2\text{—}CH_2\text{—}$	174	41	8	18	32	0
Asparagine	 $H_2N\text{—}\underset{\underset{O}{\|}}{C}\text{—}CH_2\text{—}$	132	36	6	36	21	0
Aspartic acid	 $HO\text{—}\underset{\underset{O}{\|}}{C}\text{—}CH_2\text{—}$	133	36	5	48	11	0
Cysteine	$HS\text{—}CH_2\text{—}$	121	30	6	26	12	26
Glutamic acid	 $HO\text{—}\underset{\underset{O}{\|}}{C}\text{—}CH_2\text{—}CH_2\text{—}$	147	41	6	44	10	0
Glutamine	 $H_2N\text{—}\underset{\underset{O}{\|}}{C}\text{—}CH_2\text{—}CH_2\text{—}$	146	41	7	33	19	0
Glycine	$H\text{—}$	75	32	7	43	19	0
Histidine		155	46	6	21	27	0
Isoleucine	 $CH_3\text{—}CH_2\text{—}\underset{\underset{CH_3}{\vert}}{CH}\text{—}$	131	55	10	24	11	0
Leucine	 $CH_3\text{—}\underset{\underset{CH_3}{\vert}}{CH}\text{—}CH_2\text{—}$	131	55	10	24	11	0
Lysine	$H_2N\text{—}CH_2\text{—}CH_2\text{—}CH_2\text{—}$	146	49	10	22	19	0
Methionine	$CH_3\text{—}S\text{—}CH_2\text{—}CH_2\text{—}$	149	40	7	22	9	2
Phenylalanine		165	66	7	19	8	0
Proline		115	52	8	27	12	0
Serine	$HOCH_2\text{—}$	105	34	7	46	13	0
Threonine	 $CH_3\text{—}\underset{\underset{OH}{\vert}}{CH}\text{—}$	119	40	8	40	12	0
Tryptophan		204	65	6	16	14	0
Tyrosine		181	60	6	26	8	0
Valine	 $CH_3\underset{\underset{CH_3}{\vert}}{CH}\text{—}$	117	51	9	12	27	0

organic, enzymatic, or microbiological means. All these methods yield the L-isomer.

Alanine and aspartic acid are produced commercially utilizing enzymes. In the case of alanine, the process of decarboxylation of aspartic acid by the aspartate decarboxylase from *Pseudomonas dacunhae* is commercialized. The annual world production of alanine is about 200 t. Aspartic acid is produced commercially by condensing fumarate and ammonia using aspartase from *E. coli*. This process has been made more convenient with an enzyme immobilization technique. Aspartic acid is used primarily as a raw material with phenylalanine to produce aspartame, a noncaloric sweetener. Production and sales of aspartame have increased rapidly since its introduction in 1981. Tyrosine, valine, leucine, isoleucine, serine, threonine, arginine, glutamine, proline, histidine, citrulline, L-dopa, homoserine, ornithine, cysteine, tryptophan, and phenylalanine also can be produced by enzymatic methods.

The worldwide amino acids market amounts to about \$5 billion. MSG and the animal feed additives, methionine and lysine, account for about 75% of this sales value. The other amino acids are used as precursors in pharmaceuticals, food additives, and animal feed. The worldwide demand for glutamic acid is about 800,000 t/year, 300,000 t/year for methionine, and 500,000 t/year for lysine. Other significant amino acids such as aspartic acid, phenylalanine, threonine, and glycine, each have a worldwide market of about 10,000–20,000 t/year. Tryptophan and cysteine command a global market in the thousands of tons as well.

Glutamic acid. MSG is an important flavor enhancer for natural and processed foods. It is also good for protecting the flavor and color of preserved foods and suppressing off-flavors. Glutamic acid is not an essential amino acid. However, it has some pharmaceutical uses and also improves the growth of pigs.

A number of glutamic acid-producing bacteria are known, for example,

- *Corynebacterium glutamicum* (synonym *Micrococcus glutamicus*)
- *Brevibacterium flavum*
- *Brevibacterium divaricatum*

Among them, *C. glutamicum* is used most commonly in the industry. Kyowa Hakko has recently completed a genome map of its glutamic acid producing organism, and hopes to improve glutamic acid fermentation efficiency and create a next-generation production system through pathway engineering.

The fermentation medium contains a carbon source (glucose or molasses), a nitrogen source (urea, ammonium sulfate, corn steep liquor, or soy hydrolyzate), mineral salts to supply potassium, phosphorus, magnesium, iron, and manganese, and less than 5 μg of biotin per liter depending upon the strain. The biotin requirement is the major controlling factor in the fermentation. When too much biotin is supplied for optimal growth, the organism produces lactic acid. Under conditions of suboptimal growth, glutamic acid is excreted. The metabolic pathway involved in the biosynthesis of glutamic acid from glucose limits α-ketoglutarate dehydrogenase activity, which is a notable phenotype of glutamic acid overproducing microorganisms.

The fermentation is conducted aerobically in tanks with k_La value in 300 mmol O_2/L/h/atm range. If aeration is not adequate, lactic acid is produced, and the yield of glutamic acid is poor. If aeration occurs in excess, even more lactic acid in addition to some α-ketoglutaric acid is produced. The fermentation temperature is 28–33°C, and the optimal pH is 7–8. Continuous feeding of liquid or gaseous ammonia controls pH and supplies ammonium ions to the fermentation. The fermentation cycle is 24–48 h, and the final concentration of glutamic acid is about 120–150 g/L. The overall yield of glutamic acid on sugar is about 65% on a weight basis. A portion of sugar is used for cellular growth; otherwise glutamic acid yield on sugar is 86% according to the equation:

$$C_{12}H_{22}O_{11} + 2NH_3 + 3O_2 \rightarrow 2C_5H_9O_4N + 2CO_2 + 5H_2O$$

For recovery of glutamic acid after fermentation, the broth is clarified by adding acid to pH 3.4, heating to 87°C, holding for a sufficient time to coagulate suspended solids, and filtering the coagulated solids. The clarified broth is concentrated by evaporation followed by glutamic acid crystallization. Other recovery schemes such as membrane filtration or ion-exchange can be used.

MSG sells for about \$2/kg, and glutamic acid (99.5% pure) for about \$4/kg. The major manufacturers of these are Ajinomoto, Asahi Foods, Kyowa Hakko, Takeda (Japan), Orsan (France), Biacor (Italy), Cheil Sugar, Mi-Won (S. Korea), and Tung Hai Fermentation Industry, Ve Wong (Taiwan).

Lysine. Lysine, biologically active in its L-configuration, is an essential amino acid in human and animal nutrition. The richest sources of lysine are animal proteins such as meat and poultry, but it is also found in dairy products, eggs, and beans. On the other hand, cereal grains, such as corn, wheat, and rice, are usually low in lysine. For applications in human food, lysine in its salt forms can be added to cookies and bread, and in solution it can be used to soak rice. As most animal feed rations are based on corn and other grains, supplementing the feedstuffs with lysine (plus methionine) significantly improves their nutritional value for breeding poultry and pigs.

A fermentation process for producing lysine was made possible by using mutants of *C. glutamicum* or *B. flavum.* Both auxotrophic and regulatory mutants have been obtained for overproduction of lysine. Figure 30.24 shows the biosynthetic pathway of lysine and its metabolic controls in *C. glutamicum.*

Molasses was the most common carbon source until being replaced by glucose for reasons of easier downstream processing. Sufficient amounts (over 30 μg/L) of biotin must be included in the medium to prevent the excretion of glutamic acid. This biotin requirement was previously met by using molasses as the carbon source and must now be added exogenously. The fermentation runs at temperatures of about 28–33°C and pH 6–8. High aeration is desirable. The final product concentration is around 100–140 g/L, and the fermentation time is 48–72 h. The yield of lysine on carbohydrate is about 40–50%. The formation of lysine from sucrose can be represented as follows:

$$C_{12}H_{22}O_{11} + 2NH_3 + 5O_2 \rightarrow C_6H_{14}O_2N_2 + 6CO_2 + 7H_2O$$

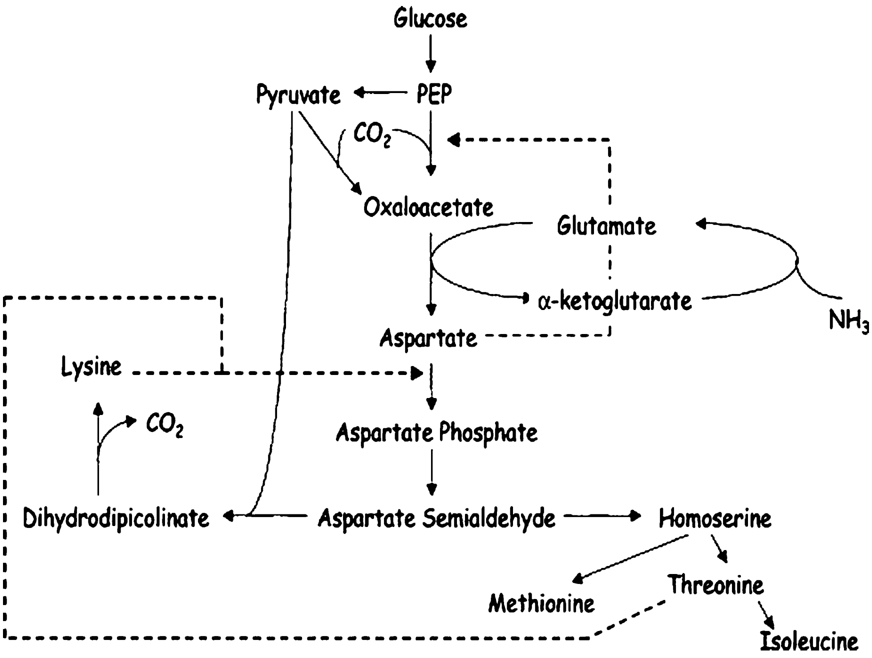


Fig. 30.24 Lysine biosynthesis from glucose in *Corynebacteria*

Ion-exchange resins are used for isolation of lysine from fermentation broths. The eluted lysine is then crystallized from water. The most common commercial form of lysine used in animal feed is 98% lysine monohydrochloride. For lower purity, the entire fermentation broth may be evaporated and syrup or the dried product is used as an animal feed supplement.

Standard commercial food-grade lysine (as 98.5% lysine monohydrochloride) is about $5/kg, and lysine as a feed supplement is about $2/kg. The major producers include Ajinomoto and Kyowa Hakko of Japan, Sewon and Chiel Sugar of South Korea, and ADM, BioKyowa (a subsidiary of Kyowa Hakko), Heartland Lysine (a subsidiary of Ajinomoto), and Degussa (formerly joint venture of Cargill and Degussa in the United States). China's Global Bio-Chem has been increasingly selling protein–lysine (protein with enhanced lysine content), as well. The world demand for lysine is expected to increase continually. In fact, the increasing ethanol fermentation capacity for bioethanol production provides an opportunity for increased lysine production. Distillers' Dried Grains (DDG), a by-product of ethanol fermentation used as a protein source in animal feeds, is deficient in lysine and other essential amino acids. Therefore, it needs to be supplemented with lysine for full-value use.

Aromatic amino acids. The aromatic amino acids, phenylalanine and tryptophan, provide some of the first examples of chemical production using microorganisms through the use of pathway engineering [62, 63]. Intermediates in the aromatic amino acid pathway can also be used as precursors to other biosyntheses with genes recruited from different organisms. Examples include catechol, adipic acid, shikimic, and quinic acids. In general, the aromatic pathway illustrates the potential of multiple product opportunities from one pathway and provides a great leveraging factor with respect to technical and commercial development costs.

Initially, work on the aromatic amino acid pathway of *E. coli* was focused on the construction of a strain for the overproduction of natural amino acids (phenylalanine, tryptophan, tyrosine). These efforts have focused on:

1. Cloning and optimization of complete primary aromatic pathway with an emphasis on removal of transcriptional and allosteric regulation, that is, removal of rate-limiting steps.
2. Enhancement of carbon flux to the aromatic pathway through modification of gene activities within central metabolism.

For the aromatic pathway (Fig. 30.25), the critical control points are the condensation of PEP and erythrose-4-phosphate to 3-deoxy-D-arabinoheptulosonate 7-phosphate, DAHP, by DAHP synthase. For tryptophan, the formation of anthranilic acid from chorismic acid by anthranilate synthase is the second critical control point. The transcriptional regulation was overcome through the use of alternative promoters and allosteric regulation was circumvented by the classical technique of selection for feedback-resistant mutants using toxic analogues of the repressing compounds.

For tryptophan production in *E. coli*, the natural regulation controlling production of tyrosine and phenylalanine was

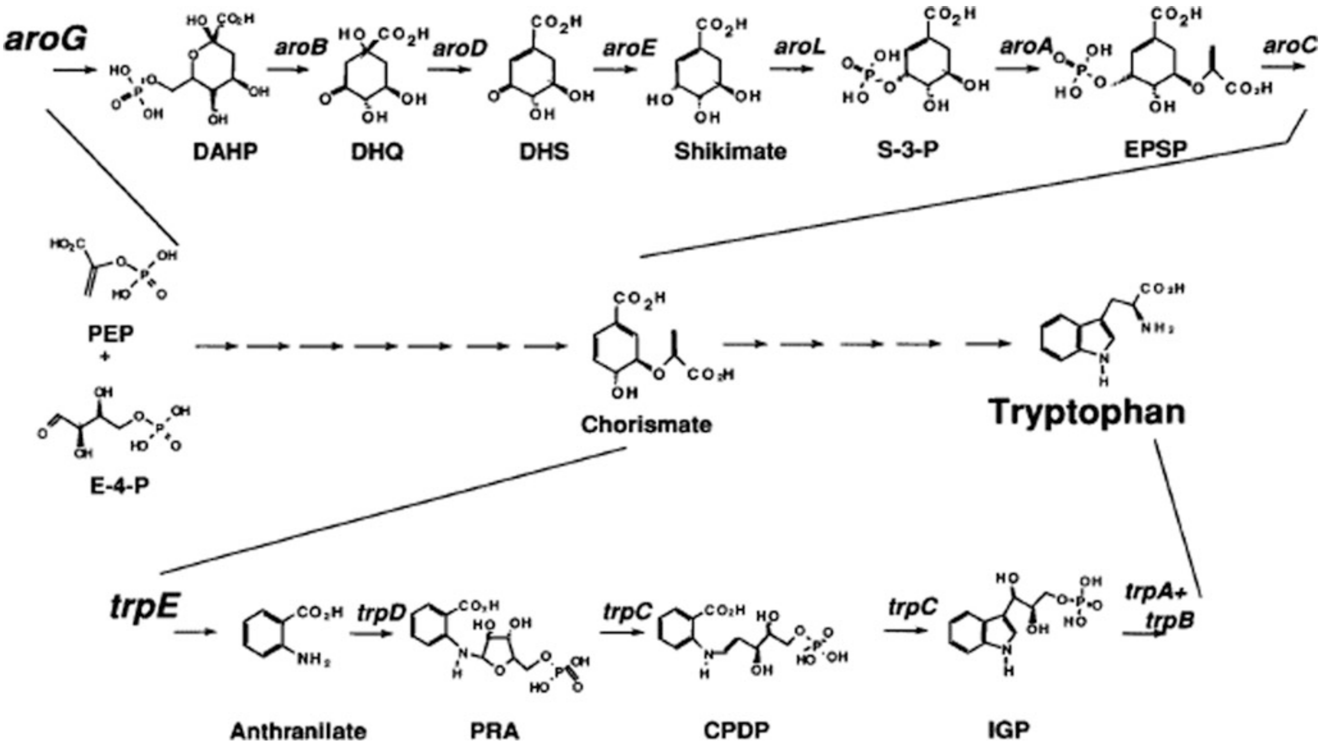


Fig. 30.25 *Escherichia coli* aromatic pathway to tryptophan

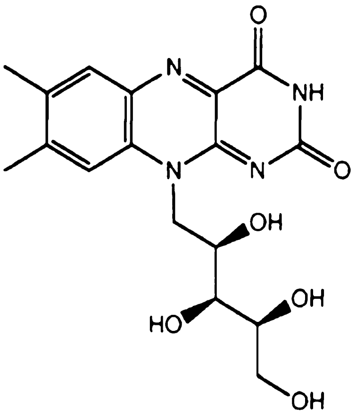
sufficient to keep carbon flowing specifically to tryptophan. This eliminated the need for addition of auxotrophic compounds to the growth medium. The major industrial producers of tryptophan are ADM, Kwoya Hakko, and Ajinomoto.

The same could not be said for the construction of a strain for the overproduction of phenylalanine. Here the control mechanisms for tyrosine were not sufficient to keep a significant amount of carbon from being diverted. However, instead of using an auxotrophic strain, a technology was developed to keep the gene for chorismate mutase and prephenate dehydrogenase present during the growth phase of the fermentation and then have it excised from the chromosome during the production phase. Due to the sweetener Aspartame, the market for phenylalanine has expanded to 10,000 metric tons per year with an average price of $10 per kg. The major producers are Nutrasweet and Ajinomoto.

Vitamins and Neutraceuticals

In the commercial market for vitamins, the most significant are ascorbic acid (Vitamin C), the D group (D2, D3, ergosterol), folic acid, and alpha-tocopherol (Vitamin E). The next important group of vitamins is biotin, cyanocobalamin (B_{12}), inositol, pantothenic acid, and riboflavin (B_2). Microorganisms can synthesize many vitamins of medical importance. Vitamin B_2 (riboflavin) and vitamin B_{12} (cyanocobalamin) are products of fermentation. Vitamin C (ascorbic acid) precursor, 2-keto-L-gulonic acid (2-KLG) is produced microbiologically as well as chemically as discussed in the “Organic Acids and Polymers” section.

Riboflavin. Riboflavin is used as a dietary supplement in both human food and animal feed. The yellow-orange riboflavin crystals are only sparingly soluble in water. To include riboflavin in water-soluble formulations, sodium riboflavin-5′-phosphate is used.



The first organism employed primarily for riboflavin production was *Clostridium acetobutylicum*, the anaerobic bacterium used for the microbial production of acetone and butanol. Riboflavin was purely a by-product and was found in the dried stillage residues in amounts ranging from 40 to 70 µg/g of dried fermentation solids. Later investigations disclosed that riboflavin could be produced by yeast such as *Candida flareri* or *C. guilliermondi*, and the yield was as high as 200 mg/L.

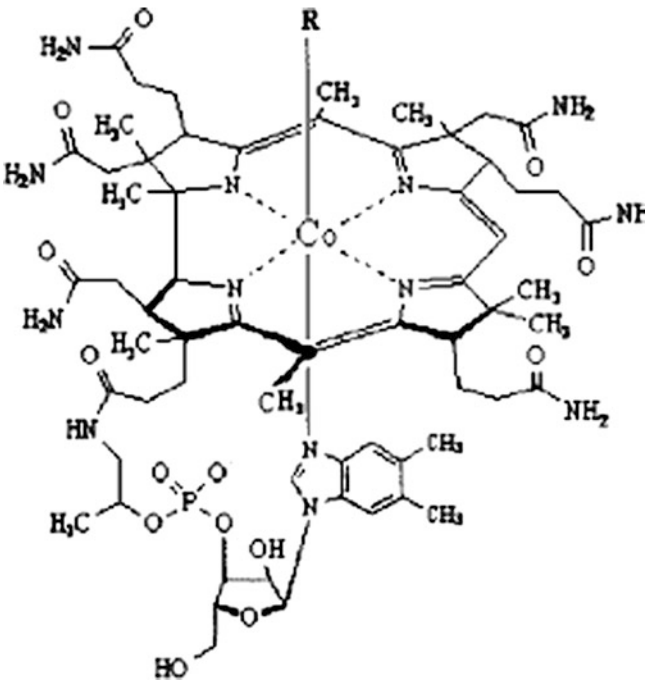
Other studies on a fungus, *Eremothecium ashbyii*, and a closely related organism known as *Ashbya gossypii* resulted in the production of much larger amounts of riboflavin. Yields as high as 10–15 g/L were possible. Then, major fermentation strain and process improvements were made with the *A. gossypii* strain. The fermentation lasts 8–10 days. Cell growth occurs in the first 2 days, and enzymes catalyzing riboflavin synthesis are formed during the growth period. Glycine and edible oil stimulate the formation of riboflavin, but they are not its precursors. The additions of carbohydrate and oil permit the overproduction of riboflavin [64]. The riboflavin carbon yield is about 50% on carbohydrate and about 100% on oil.

Upon completion of the fermentation, the solids are dried to a crude product for animal feed supplement or processed to an USP-grade product. In either case, the pH of the fermented medium is adjusted to pH 4.5. For the feed-grade product, the broth is concentrated to about 30% solids and dried on double-drum driers.

When a crystalline product is required, the fermented broth is heated for 1 h at 121°C to solubilize the riboflavin. Insoluble matter is removed by centrifugation, and riboflavin recovered by conversion to the less soluble form. The precipitated riboflavin is then dissolved in water, polar solvents, or an alkaline solution, oxidized by aeration, and recovered by recrystallization from the aqueous or polar solvent solution or by acidification of the alkaline solution.

The major producers of riboflavin include DSM and BASF. There have been significant improvements in both of the production processes. DSM opened a new riboflavin production facility in Grenzach, Germany in 2000, based on a fermentation process with a genetically modified *B. subtilis*. On the other hand, BASF, working with the University of Salamanca, Spain, engineered an improved *A. gossipii* that produces a larger amount of enzymes for riboflavin synthesis, and increased its production capacity by 20% at its Ludwigshafen facility without capital investment. The annual production of riboflavin by ADM, BASF, and DSM is about 4,000 metric ton. The feed-grade product sells for about \$30/kg and the USP-grade product \$50/kg.

Cyanocobalamin



Vitamin B_{12}, cyanocobalamin, is an important biologically active compound. It serves as a hematopoietic factor in mammals and as a growth factor for many microbial and animal species. Its markets are divided into pharmaceutical (96–98% pure) and animal feed (80% pure) applications. All vitamin B_{12} is now made commercially by fermentation [65].

Spent liquors from streptomycin and other antibiotic fermentations contain appreciable amounts of vitamin B_{12}. Bacterial strains producing high amounts have been specially selected for commercial production. Today vitamin B_{12} is obtained from fermentations using selected strains of *Propionibacterium* or *Pseudomonas* cultures. A full chemical synthesis process for vitamin B_{12} is known. However, it requires some 70 steps and for all practical purposes is of little value.

The *Pseudomonas denitrificans* strain is most often used for commercial production of vitamin B_{12} [66, 67]. It requires medium containing sucrose, yeast extract, and several metallic salts. Dimethylbenzimidazole (10–25 mg/L) and cobaltous nitrate (40–200 mg/L) must be supplemented at the start of the culture in order to enhance vitamin production. Betaine (tri-methyl glycine) stimulates the biosynthesis of vitamin B_{12}, even though the organism does not metabolize it. Similarly, choline also has favorable effects, either by activating some biosynthetic steps or altering the membrane permeability. Glutamic acid, on the other hand, stimulates cellular growth. Because of high betaine and glutamic acid contents, beet molasses (60–120 g/L) is preferentially used in industrial fermentations of vitamin B_{12}. The fermentation is conducted with aeration and agitation. Optimal temperature is about 28°C and pH near 7.0. The yield reported in the literature was 59 mg/L in 1971, using a *Pseudomonas* strain. A yield of 200 mg/L was reported for vitamin B_{12} fermentations using *Propionibacteria* in 1974. It is believed that yields of vitamin B_{12} have been significantly improved since then.

About 80% of the vitamin produced is outside the cells and 20% inside the cells. For recovery, the whole broth is heated at 80–120°C for 10–30 min at pH 6.5–8.5 and treated with cyanide or thiocyanate to obtain cyanocobalamin. B_{12} isolation can then be accomplished by adsorption on a cation-exchange resin, such as Amberlite IRC 50. Extraction can also be done by using phenol or cresol alone or in a mixture with benzene, butanol, carbon tetrachloride, or chloroform; or separation by precipitation or crystallization upon evaporation with appropriate diluents such as cresol or tannic acid. Using the extraction method, 98% pure cyanocobalamin can be obtained with a 75% yield. The total world market for cyanocobalamin is estimated to be in excess of 10,000 kg/year and sells for about $25/g. The leading producers include DSM, Wockhardt, and Merck.

Docosahexaenoic and arachidonic acids. Docosahexaenoic acid, DHA, is a long-chain polyunsaturated omega-3 (22:6 ω3) fatty acid whereas arachidonic acid, ARA, is a long-chain polyunsaturated omega-6 (20:4 ω6) fatty acid. Both DHA and ARA are found in human milk and are important nutrients in infant development, and especially for brain development. DHA is a major structural and functional fatty acid in the gray matter of the brain, the retina of the eye, and the heart cell membranes. DHA can be obtained through the diet in foods such as fatty fish (accumulated via natural algae) and organ meat. ARA is the principal omega-6 fatty acid found in the brain as well. Adults obtain ARA readily from the diet in foods such as meat, eggs, and milk. The DHA- and ARA-rich oil can also be used in supplements and functional foods for older children and adults for improvement in cardiovascular health.

Martek (http://www.martekbio.com) and DSM have developed processes to produce oils rich in DHA and ARA. The DHA is extracted from fermented microalgae (*Cryptecodinium cohnii*) and the ARA is extracted from soil fungus (*Mortierelle alpina*). In the DHA production process, the microalgae are grown in fermentors (80,000–260,000 L scale). The cell mass is then harvested by centrifugation, and the concentrated cell mass is spray-dried. The dry cell mass is then broken down into smaller particles and extracted by a solvent for the oil-containing DHA. The solvent is then evaporated from the oil, leaving crude oil. The crude oil is refined, bleached, and deodorized to cleanse it of any impurities. The final product is rich in DHA. Likewise, DSM ferments the fungus, and extracts the crude ARA oil by the DHA downstream type process. Crude oils can be extracted with the traditional hexane extraction method or with new solvents such as isopropanol or supercritical CO_2. The final cell mass, total fermentation time, and the lipid content of the organism are very important for the economic feasibility, as these factors largely determine the productivity. Media components such as the amount of sugars and nitrogen have a significant effect on growth and lipid accumulation.

Table 30.5 Examples of natural antibiotics and producing organisms

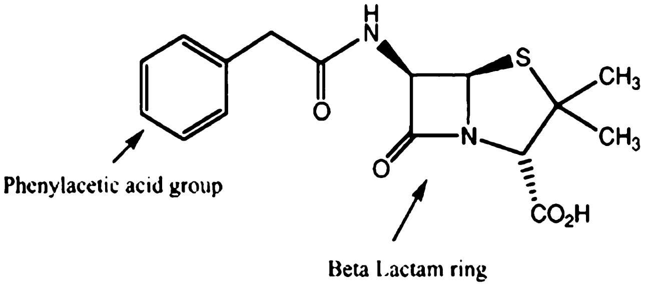
Natural antibiotic	Producing organism
Amphotericin B	*Streptomyces nodosus*
Bleomycin	*Streptomyces verticillus*
Cephalosporin	*Cephalosporium* sp.
Erythromycin A	*Saccharopolyspora erythraea*
Gentamycin	*Micromonospora purpurea*
Lovastatin	*Aspergillus terreus*
Neomycin	*Streptomyces fradiae*
Penicillin G	*Pencillium chrysogenum*
Spiramycin	*Streptomyces ambofaciens*
Tetracycline	*Streptomyces aureofaciens*
Vancomycin	*Streptomyces orientalis*

The market for DHA, ARA, and other essential neutraceuticals used in dairy, beverages, cereals, and breads, is multibillion dollars and growing as the health benefits are determined for prenatal to elderly care.

Antibiotics

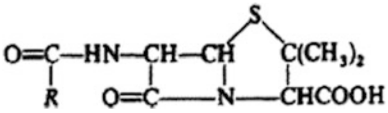
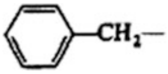
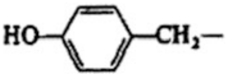
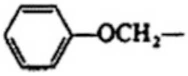
Since the early 1940s, an intensive search for new and useful antibiotics has taken place throughout the world. More than 10,000 antibiotics have been discovered from microbial sources [68–73]. *Streptomyces* spp., filamentous fungi, nonactinomycete bacteria, and non-Streptomycete strains of actinomycetales are major sources of antibiotics, examples of which are listed in Table 30.5. Of the antibiotics in clinical use, 50% are produced by microbial fermentation and others by a combination of microbial synthesis and subsequent modification.

Penicillins



Penicillin G

Table 30.6 Structural formulae of natural penicillins

Formula	
Type of penicillin	*Side chain R substitutions*
(G) Benzyl	
(X) *p*-Hydroxybenzyl	
(F) 2-Pentenyl	$CH_3—CH_2—CH{=}CH—CH_2—$
(Dihydro F) *n*-Pentyl	$CH_3—CH_2—CH_2—CH_2—CH_2—$
(K) *n*-Heptyl	$CH_3—CH_2—CH_2—CH_2—CH_2—CH_2—CH_2—$
(V) Phenoxy	

The original mold observed and preserved by Alexander Fleming was a strain of *P. notatum*, a common laboratory contaminant. Later, cultures of *P. chrysogenum* were found to be better producers of penicillin, and the present industrial strains have been derived from this species. The original strains produced the antibiotic only by surface fermentation methods and in very low yields. Improved media and productive strains under submerged aerobic fermentation conditions led to dramatic yield increases. Subsequent improvements, principally in culture selection and mutation, further improved yields, reaching 20–30 g/L.

Table 30.6 gives the structural formulae of the "natural penicillins," which are comprised of several closely related structures with aliphatic and aromatic substitutions to the common nucleus. The early impure product contained mixtures of these penicillins; however, penicillin G has become the preferred type and the crystalline product of commerce. Phenylacetic acid or its derivatives are used as precursors in the fermentation medium to enhance penicillin biosynthesis.

Nonpigmented mutants of *P. chrysogenum*, for example, Wisconsin 49-133 derivatives, are universally employed in the industrial process. The desired culture is propagated from a laboratory stock in small flasks and transferred to seed tanks. A typical production medium is as shown below.

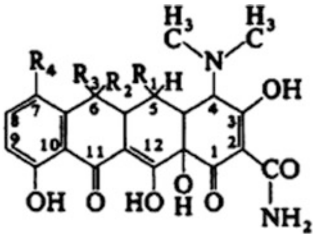
Components	Grams/liter
Corn steep liquor	30
Lactose	30
Glucose	5.0
$NaNO_3$	3.0
$MgSO_4$	0.25
$ZnSO_4$	0.044
Phenyl acetamide (precursor)	0.05
$CaCO_3$	3.0

The medium is sterilized, cooled to 24°C, and inoculated. After 24-h seed growth, larger fermentors are seeded. Sterile air is sparged through the tank, at about one volume per volume per minute. The time of the production stage varies from 60 to 200 h.

For recovery, the broth is clarified by means of rotary vacuum filters. The penicillin, being acidic, is extracted from the aqueous phase into a solvent, such as methyl isobutyl ketone or amyl acetate, at a pH of 2.5 by means of a continuous countercurrent extractor, such as a Podbielniak. The penicillin is then re-extracted with an aqueous alkaline solution or a buffer at a pH of 6.5–7.0. About 90% recovery yield is typical at this step. The aqueous solution is chilled, acidified, and extracted again with a solvent, such as ether or chloroform. Finally, penicillin is re-extracted into water at a pH of 6.5–7.0 by titration with a solution of base. The base used depends on which salt of penicillin is desired. The popular forms are sodium or potassium salts. A typical flow sheet for antibiotic recovery is shown in Fig. 30.6.

Cephalosporins. Microbiological processes for production of cephalosporin C resemble in many respects those used for penicillin production. Special strains of *Cephalosporium* have been selected that produce more cephalosporin C and less cephalosporin N than the parent culture. The growth of these strains in certain special fermentation media has resulted in higher antibiotic titers. Even with these improvements in processing, the antibiotic concentration, averaging 10–20 g/L, is much lower than those reported for the penicillins.

Table 30.7 Structure of clinically important tetracyclines



	R_1	R_2	R_3	R_4
Tetracycline	H	OH	CH_3	H
7-Chlortetracycline (Aureomycin®)	H	OH	CH_3	Cl
5-Oxytetracycline (Terramycin®)	OH	OH	CH_3	H
6-Demethyl-7-chlortetracycline (Declomycin®)	H	OH	H	Cl
6-Deoxy-5-oxytetracycline (Vibramycin®)	OH	H	CH_3	H
6-Methylene-6-deoxyl-6-demethyl-5-oxytetracycline (Rondomycin®)	OH	H	CH_2	H

Tetracyclines. In 1948, a broad-spectrum antibiotic, chlortetracycline (Aureomycin), was introduced by the Lederle Laboratories, a Division of American Cyanamid Company. This antibiotic is produced by *Streptomyces aureofaciens* when grown under submerged aerobic conditions on media composed of sugar, corn steep liquor, and mineral salts. The crystalline compound has a golden yellow color, which suggested the trade name. A related antibiotic, oxytetracycline (Terramycin), a product of *Streptomyces rimousus*, is produced by Pfizer Inc. It is chemically and biologically similar to chlortetracycline. Independent research by both companies eventually provided the structure of these two important chemotherapeutic agents. Both the compounds may be regarded as derivatives of a nucleus known as tetracycline. Their structures along with those of other clinically important tetracyclines are given in Table 30.7. Tetracycline can also be produced by fermentation of *S. aureofaciens* under special conditions, such as chloride starvation or by use of special strains of the organism that fail to halogenate efficiently. Tetracycline possesses many of the chemotherapeutic properties of chlortetracycline and oxytetracycline. Mutations of tetracycline-producing organisms have led to other tetracycline analogues, of which 6-demethyl-7-chlortetracycline (Declomycin) has clinical use. Chemical modifications of oxytetracycline have generated two other useful members of the family, known as Vibramycin and Rondomycin. Tetracyclines are active in vivo against numerous gram-positive and gram-negative organisms, and some of the pathogenic Rickettsiae and large viruses.

Dramatic price reduction has come with improved production. For instance, a million-unit vial of penicillin (1,667 units = 1 mg of potassium penicillin G) had a wholesale price of \$200 in 1943. Today, a million units of potassium penicillin G sell for as little as \$0.50 or approximately \$20/kg of free acid. Cephalosporins sell for around \$250/kg. Tetracyclines, used as animal feed supplements, sell for about \$60/kg. The antibiotic market is over \$12 billion, and volume continues to increase simultaneous with price drop. Cephalosporins, macrolides, penicillins, quinolones, and tetracyclines account for the majority of the sales of antibiotics.

In the last 20 years, it has become evident that the contribution of microorganisms does not have to be limited to the realm of infectious diseases. Metabolites of microorganisms have been found to have many other interesting therapeutic applications. Particularly as our understanding of disease processes at the biochemical and genetic levels becomes more advanced, enzymes or receptors have increasingly served as specific targets for therapeutic intervention. Recombinant DNA technology, in particular, has helped to produce reagents as tools for the development of practical and high-throughput screening methodologies based on mechanism of action. Automation and miniaturization have also allowed a rapid increase in the throughput of the screening process.

Several of these bioactive natural products have been successfully developed as therapeutics for clinical use. For example, Cyclosporin A is a fungal decapeptide principally used to suppress immune rejection in organ transplant patients. Mevinolin and compatin both control cholesterol synthesis in human. The search for enzyme- or receptor-targeted microbial products does not limit itself to medical use. Several commercially important "antibiotics" such as Nikkomycin and Avermectin have been found for agricultural applications in recent years.

Biopharmaceuticals

The term "biopharmaceuticals" has been generally accepted to distinguish the large molecule pharmaceuticals (mainly proteins) that have emerged as a result of the modern "biotechnology" from the traditional small molecule drugs.

Table 30.8 Examples of biopharmaceutical products, applications, and manufacturers

Biopharma product	Application	Companies
Insulin	Diabetes	Eli Lilly, Novo Nordisk
Erythropoietin	Anemia	Amgen, Johnson & Johnson
Human growth hormone	Growth deficiency	Genentech, Pharmacia
Interferon α	Hepatitis	Schering, Roche
Interferon β	Multiple sclerosis	Chiron, Biogen
Factor VIII	Hemophilia	Bayer
Tissue Plasminogen Activator	Blood clot	Genentech
Glucocerebrosidase	Gaucher's disease	Genzyme
Therapeutic antibodies	Cancer	Glaxo, Amgen, Genentech
GCSF	White blood cell	Amgen, Sankyo

Biopharmaceuticals that are industrially produced by microbial (bacterial or yeast) fermentation include insulin, human growth hormone, hepatitis B surface antigen vaccine, alpha interferon, beta interferon, gamma interferon, granulocyte colony stimulating factor, and interleukin-2. Table 30.8 lists some of the major biopharma products, some of which are produced by mammalian cell culture.

Insulin is the first human biopharmaceutical product that was commercialized using the recombinant DNA technology. Its fermentative expression in *E. coli*, downstream recovery, and purification serves as a good example for the large-scale production of recombinant proteins. Insulin, given its rather small size, is expressed as N-terminal extended proinsulin using the *trp* promoter. The fermentation is carried out at about 37°C and near neutral pH. It is a fast, aerobic fermentation. The *trp* operon is turned on when the fermentation runs out of its natural repressor, tryptophan. The chimeric protein as it is expressed accumulates very rapidly inside the cells as insoluble aggregates (inclusion or refractile bodies). The formation of inclusion bodies prevents proteolysis and facilitates product recovery. The recombinant *E. coli* fermentation usually runs for 18–24 h. At the end of fermentation, the inclusion bodies account for about 10–30% of total dry cell mass. Following purification from the inclusion bodies, the C-peptide is removed from proinsulin using trypsin and carboxypeptidase, and the properly folded and cleaved insulin is purified further. Fermentatively or microbially produced recombinant proteins can also be expressed as soluble proteins inside the cells or secreted outside the cells.

The biopharmaceuticals sales are about $45 billion and represent 10% of the total pharma market. Today, 25% of new drugs are biopharmaceuticals. Since the introduction of insulin over 25 years ago, 160 biopharma products, ranging from proteins, monoclonal antibodies, and nucleic acid-based products, have been approved for use. In 2004 alone, 12 new biopharmaceuticals have been approved, of which only 3 are produced by microbial fermentation, as cell culture is the preferred method for production of biopharma drugs.

Drug development, in general, is a slow, lengthy, and expensive process, requiring an average investment of $0.8 billion over 12–15 years for discovery, development, preclinical testing, Phase I, II, and III trials, and FDA approval prior to marketing. Given the complications of biopharma drug discovery and development, microbial fermentation and biocatalysis may aid in introducing rapid and innovative processes to produce potential drug candidates. It may thus be worth exploring the possibility of producing biopharmaceuticals by microbial fermentation and demonstrating that they can be made reproducibly to meet the right purity, concentration, and dose at the required market cost and volume targets.

Enzymes

Microorganisms used for production of enzymes range from prokaryotic systems including both the gram-negative and gram-positive bacteria to eukaryotic systems such as yeasts and fungi. For most of the history of enzymes, their production has occurred predominantly in strains known to make the enzyme of interest [30, 74–77]. Thus, many different types of microorganisms have been employed to make enzymes for different types of uses, as discussed in Chap. 31.

For example, an alkaline protease naturally secreted by *Bacillus licheniformis* to break down proteinaceous substrates, resulted in one of the first commercially produced enzymes, Subtilisin Carlsberg, for use in detergents. Similarly, α-amylase was produced from *B. licheniformis* because it naturally secretes a highly thermo-stable α-amylase capable of breaking down starch to more easily digestible oligosaccharides. Strains of *Bacillus* have been some of the workhorse strains for enzyme production for decades, based mainly upon their ability to overproduce subtilisin and α-amylase. Those strains and the promoters of these genes are also used to express other enzymes, wild type and engineered.

Glucoamylase, which completely breaks down oligosaccharides to glucose, is produced by strains of the fungal genus, *Aspergillus*. Strains that overproduce glucoamylase have been isolated over the years. Another carbohydrase, an acid lignocell enzyme complex, found in the fungus *Trichoderma*, is capable of breaking down cellulosic substrates to glucose, similar to the starch-degrading enzymes. This particular application has turned out to be much more challenging to commercialize but has found applications in treatment of textiles, feed, and food. New programs initiated by the U.S. Department of Energy have helped improve the expression and activity of the lignocell enzyme complex.

One of the enzymes that has been used on a large scale is glucose/xylose isomerase, which catalyzes the rearrangement of glucose to fructose, to yield a product with a sweet taste like sucrose. The enzyme is present in *Streptomyces* spp. as well as a few other organisms. Once again, the native host strain has been improved for production of glucose isomerase.

Future: Biorefineries

Biorefineries integrate biobased industries/processing facilities that utilize mostly plant materials as feedstocks to produce food, feed, fuel, chemicals, materials, and intermediates for making varieties of products. Fermentative processes discussed here and biocatalysis discussed in Chap. 31 can be applied for a wide variety of processes where renewable resource-based materials are used and produced. Bioprocesses are continually becoming more efficient through better understanding and controlling of metabolic pathways. Today, industrial biotechnology companies such as DuPont Industrial Biosciences and Novozymes sell enzymes produced via microbial fermentation processes for a wide variety of bioprocessing applications. Chemical producers are seeking new and renewable feedstocks as well as environmentally friendly and sustainable production processes, in the face of escalating nonrenewable raw material prices. In addition, there is a need for development of interesting new chemistry and products. Major chemical companies—including Dow Chemical, DuPont, BASF, Degussa, and Celanese—have invested heavily to explore these opportunities, often through alliances with small technology firms that offer specific expertise (http://pubs.acs.org/email/cen/html/060804150713.html). However, most products made today are relatively low-volume, high-value types, for example, drugs and specialty chemicals. Only a few products, such as ethanol, are produced by fermentation at a million-ton scale per year.

Improvements in fermentation organisms and processes will have to be achieved in order to raise productivity and conversion efficiency and lower production costs (http://www.nrel.gov/docs/fy04osti/35523.pdf). Enzymatic processing is also expected to play a major role in refining biofeedstocks into sugars, protein, oil, and other by-products. Such advances alone will enable the vision of biorefineries to become a reality. Analogous to conventional refineries, progress in three areas will be key for the successful development of biorefineries: low energy milling of biofeedstocks to its components, efficient bioconversion of mixed sugars to products, and the utilization of by-products. These improvements will require integration of all major areas of industrial biotechnology: novel enzymes and microorganisms, functional genomics, pathway engineering, protein engineering, biomaterial development, bioprocess design, product development, and applications.

Acknowledgment The subject of biotechnology first appeared in an earlier edition of this *Handbook* as a chapter titled “Industrial Fermentation: Principles, Processes and Product,” written by Dr. Arthur E. Humphrey of Lehigh University. In later editions he was joined first by Dr. S. Edward Lee and then by Dr. Lewis Ho, both of Pfizer. Chaps. 30 and 31 in this edition are an outgrowth of that earlier work, some of which is used in the new chapters. Grateful acknowledgment of this use is made to the authors of that work.

References

1. Industrial Biotechnology and Sustainable Chemistry (2004) Royal Belgian Academy Council, Jan 2004. www.kvab.be
2. Theil EC (ed) (1998) Principles of chemistry in biology. American Chemical Society, Washington, DC
3. Aiba S, Humphrey AE, Milis N (1973) Biochemical engineering, 2nd edn. University of Tokyo Press, Tokyo
4. Ingram JL, Maaloe O, Neidhart FC (1983) Growth of the bacterial cell. Sinauer Associates, Sunderland, MA
5. Baumberg S, Hunter I, Rhodes M (eds) (1989) Microbial products: new approaches. Cambridge University Press, Cambridge, England
6. Flickinger MC, Drew SW (eds) (1999) The encyclopedia of bioprocess technology: fermentation, biocatalysis and bioseparation. Wiley, New York
7. MolecularCloning.com. A laboratory manual on the Web. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
8. Stephanopoulos GN, Aristidou AA, Nielsen J (eds) (1998) Metabolic Engineering: Principles and Methodologies, Academic Press, San Diego, Calif, USA
9. Bailey JE (2001) Reflections on the scope and the future of metabolic engineering and its connections to functional genomics and drug discovery. Metab Eng 3(2):111–114
10. Lee L-Q, Jeff Varner, and Kwok Ko (2004) Parallel Extreme Pathway Computation for Metabolic Networks, IEEE Conference Proceedings, Computational Systems Bioinformatics Conference, 636–639
11. Chotani G, Dodge T, Hsu A, Kumar M, LaDuca R, Trimbur D, Weyler W, Sanford K (2000) Commercial production of chemicals using pathway engineering. Biochem Biophys Acta 1543 (2):434–455
12. Bader FG (1986) Sterilization: prevention of contamination. In: Manual of industrial microbiology and biotechnology. American Society for Microbiology, Washington, DC

13. Raju GK, Cooney CL (1993) Media and air sterilization. In: Stephanopoulos G (ed) Biotechnology, vol 3. VCH, New York
14. Bailey JE, Ollis DF (1977) Biochemical engineering fundamentals. McGraw-Hill, New York
15. Demain AL, Solomon NA (1985) Biology of industrial microorganisms. Biotechnology series. Benjamin/Cummings, Redwood City, CA
16. Shuler ML, Kargi F (eds) (2002) Bioprocess engineering: basic concepts, 2nd edn. Prentice Hall, Upper Saddle River, NJ
17. Roels JA (1983) Energetics and kinetics in biotechnology. Elsevier Biomedical Press, Amsterdam
18. Hughes MN, Poole RK (1989) Metals and microorganisms. Chapman and Hall, New York
19. Neway JO (ed) (1989) Fermentation process development of industrial organisms. Marcel Dekker, New York
20. Monod J (1974) in P. S. S. Dawson (Ed.): Microbiol. Growth, Halsted Press, New York, pp 88–110
21. Luedeking R, Piret EL (1959) A kinetic study of the lactic acid fermentation. Batch process at controlled pH. J Biochem Microbiol Tech Engr 1:393
22. Van't Riet K, Tramper J (eds) (1991) Basic bioreactor design. Marcel Dekker, New York
23. Arbige M, Bulthuis B, Schultz J, Crabb D (1993) Bacillus subtilis and other gram-positive bacteria: biochemistry, physiology and molecular genetics. ASM, Washington, DC, pp 871–895
24. Asenjo JA, Merchuk JC (eds) (1991) Bioreactor system design. Marcel Dekker, New York
25. Blanch HW, Clark DS (1996) Biochemical Engineering, Marcel Dekker, New York
26. Lyderson BK, D'Ella NA, Nelson KL (eds) (1994) Bioprocess engineering: systems, equipment, and facilities. Wiley, New York
27. Rushton JH et al (1959) Chem Eng Prog 46:467
28. Rothberg A, Weegar J, von Schalien R, Fagervik K, Rydström M, Lind K (1999) Optimization of an *Aspergillus niger* glucose oxidase production process. Bioprocess Engineering 21(4): 307–312
29. Lyddiatt A (2002) Process chromatography: current constraints and future options for the adsorptive recovery of bioproducts. Curr Opin Biotechnol 2:95–103
30. Chotani GK, Dodge TC, Herman A, van Scheltinga T, Gölker C, Heng MH, Kan J, Becker T, Fukui S, Tanaka A, Schmuck R (2003) General production. In: Aehle W (ed) Enzymes in industry—production and applications. Wiley-VCH, Weinheim
31. Weelright SM (1991) Protein purification: design and scale up of downstream processes. Wiley, New York
32. Becker T, Park G, Gaertner AL (1997) Formulation of detergent enzymes. In: Ee JH, Misset O, Baas EJ (eds) Enzymes in detergency. Marcel Dekker, New York, pp 299–325
33. Ebner H, Follmann H, Seller S (1995) Vinegar. In: Reed G, Nagodawithana TW (eds) Biotechnology, vol 9. VCH, New York
34. Schierholt J (1977) Fermentation processes for the production of citric acid. Process Biochem 12(9):20
35. Willke T, Vorlop KD (2001) Biotechnological production of itaconic acid. Appl Microbiol Biotechnol 56(3–4):289–295
36. Werpy T, Pedersen G (eds) Top value added chemicals from biomass. Volume 1. Results of screening for potential candidates from sugars and synthesis gas, NREL, Aug 2004. http://www.osti.gov/bridge
37. Anderson S, Marks CB, Lazarus R, Miller J, Stafford K, Seymour J, Light D, Rastetter W, Estell DA (1985) Production of 2-keto-L-gulonate, an intermediate in L-ascorbate synthesis, by a genetically modified Erwinia herbicola. RL Science 230:144−149
38. García-Ochoa F, Santosa VE, Casasb JA, Gómeza E (2000) Xanthan gum: production, recovery, and properties. Biotechnol Adv 18(7):549–579
39. Lee JW, Kim HU, Choi S, Yi J, Lee SY (2011) Microbial production of building block chemicals and polymers. Curr Opin Biotechnol 22:1–10
40. Nakamura C, Whited G (2003) Metabolic engineering for the microbial production of 1,3 propanediol. Curr Opin Biotechnol 14 (5):454–459
41. Bothast RJ, Schlicher MA (2005) Biotechnological processes for conversion of corn into ethanol. Appl Microbiol Biotechnol 67:19–25
42. Ghose TK, Bisaria VS (1979) Studies on the mechanism of enzymatic hydrolysis of cellulosic substances. Biotechnol Bioeng 21:131–146
43. Ho N (1980) Yeast alcohol tolerance and recombinant DNA for improved alcohol production. Ann Rep Ferment Process 4:235
44. Ingram LO, Gomez PF, Lai X, Moniruzzaman M, Wood BE, Yomano LP, York SW (1998) Metabolic engineering of bacteria for ethanol production. Biotechnol Bioeng 58:204–213
45. Mitchinson C (2004) Improved cellulases for the BioRefinery: a review of Genencor's progress in the DOE subcontract for ignocell cost reduction for bioethanol. In: Stanford GCEP biomass energy workshop, April 2004
46. Zaldivar J, Nielsen J, Olsson L (2001) Fuel ethanol production from lignocelluloses: a challenge for metabolic engineering and process integration. Appl Microbiol Biotechnol 56:17–34
47. Whited GW, Feher FJ, Benko DA, Cervin MA, Chotani GK, McAuliffe JC, LaDuca RJ, Ben-Shoshan EA, Sanford KJ (2010) Development of a gas-phase bioprocess for isoprene-monomer production using metabolic pathway engineering. Ind Biotechnol 6:152–163
48. Peralta-Yahya PP, Keasling JD (2010) Advanced biofuel production in microbes. Biotechnol J 5:147–162
49. Rude MA, Schirmer A (2009) New microbial fuels: a biotech perspective. Curr Opin Microbiol 12:274–281
50. Yan Y, Liao JC (2009) Engineering metabolic systems for production of advanced fuels. J Ind Microbiol Biotechnol 36:471–479
51. Zhang K, Sawaya MR, Eisenberg DS, Liao JC (2008) Expanding metabolism for biosynthesis of nonnatural alcohols. Proc Natl Acad Sci U S A 105(52):20653–20658
52. Kalscheuer R, Stölting T, Steinbüchel A (2006) Microdiesel: Escherichia coli engineered for fuel production. Microbiology 152:2529–2536
53. Keasling JD, Hu Z, Somerville C, Church G, Berry D, Friedman L, Schirmer A, Brubaker S, Del Cardayre SB (2007) Production of fatty acids and derivatives thereof. WO/2007/136762
54. Chappell J (1995) The biochemistry and molecular biology of isoprenoid metabolism. Plant Physiol 107:1–6
55. Ro D-K, Paradise EM, Ouellet M, Fisher KJ, Newman KL, Ndungu JM, Ho KA, Eachus RA, Ham TS, Kirby J, Chang MCY, Withers ST, Shiba Y, Sarpong R, Keasling JD (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440:940–943
56. Clomburg JM, Gonzalez R (2010) Biofuel production in Escherichia coli: the role of metabolic engineering and synthetic biology. Appl Microbiol Biotechnol 86:419–434
57. Zhang M, Eddy C, Deanda K, Finkelstein M, Picattagio S (1995) Metabolic engineering of a pentose metabolism pathway in ethanologenic Zymomonas mobilis. Science 267:240–243
58. Ho NWY, Chen Z, Brainard A (1998) Genetically engineered Saccharomyces yeast capable of effective cofermentation of glucose and xylose. Appl Environ Microbiol 64(5):1852–1859
59. Alterthum F, Ingram LO (1989) Efficient ethanol production from glucose, lactose, and xylose by recombinant Escherichia coli. Appl Environ Microbiol 55(8):1943–1948
60. Schenk PM, Thomas-Hall SR, Stephens E, Marx UC, Mussgnug JH, Posten C, Kruse O, Hankamer B (2008) Second generation

biofuels: high-efficiency microalgae for biodiesel production. Bioenerg Res 1:20–43

61. Aida K, Ichiba I, Nakayama K, Takinami K, Yamada H (eds) (1986) Biotechnology of amino acid production. Elsevier, Amsterdam
62. deBoer L, Dijkhuizen L (1990) Microbial and enzymatic processes for L-phenylalanine production. Adv Biochem Eng Biotechnol 41:1–27
63. Dodge T, Pepsin M, Berry A (1999) Proceedings of the 13th Forum for Applied Biotechnology, 22–23 September, Gent, Belgium. Med Fac Landbouw Univ Gent pp 281–288
64. Stahmann KP, Revuelta JL, Seulberger H (2000) Three biotechnical processes using Ashbya gossypii, Candida famata, or Bacillus subtilis compete with chemical riboflavin production. Appl Microbiol Biotechnol 53(5):509–516
65. Martens JH, Barg H, Warren MJ, Jahn D (2002) Microbial production of vitamin B_{12}. Appl Microbiol Biotechnol 58(3):275–285
66. Long RA, Parlin NJ. Production of vitamin B_{12}. US Patent 3,018,225
67. Bijl HL. Production and use of compositions comprising high concentrations of vitamin B_{12} activity. US Patent 5,955,321
68. Strohl WR (1997) Industrial antibiotics: today and the future. In: Strohl WR (ed) Biotechnology of antibiotics. Marcel Dekker, New York
69. Weinstein MJ, Wagman GH (Eds.) (1978), Antibiotics. Isolation, Separation and Purification (Journal of Chromatography Library, Vol. 15) Elsevier, Amsterdam
70. Martin JF, Demain AL (1980) Control of antibiotic biosynthesis. Microbiol Rev 44(2):230
71. Martin JF, Liras P (1989) Enzymes involved in penicillin, cephalosporin and cephamycin biosynthesis. Adv Biochem Eng Biotechnol 39:153
72. DiMarco A, Pennella P (1959) The fermentation of the tetracyclines. Prog Ind Microbiol 1:45
73. Perlman D (1975) Influence of penicillin fermentation technology to processes for production of other antibiotics. Process Biochem 10(9):23
74. Foster KA, Frackman S, Jolly JF (1995) Production of enzymes as fine chemicals. In: Reed G, Nagodawithana TW (eds) Biotechnology, vol 9. VCH, New York
75. Godfrey T, Reichelt J (1983) Industrial enzymology. Nature Press, New York
76. Aunstrup K (1979) Production, isolation and economics of extracellular enzymes. In: Wingard L, Katchalski-Katzir E, Goldstein L (eds) Applied biochemistry and bioengineering, vol 2. Academic Press, London, pp 27–69
77. Lilly MD (1979) Production of intracellular microbial enzymes L.B. Wingard Jr, E. Katchalski-Katzir, L. Goldstein (Eds.), Applied Biochemistry and Bioengineering (9th edn), vol 2. Academic Press, NewYork, p. 1

31

Industrial Enzymes and Biocatalysis

Joseph C. McAuliffe

Introduction

All life processes, whether plant, animal, or microbial, depend upon a complex network of enzyme-catalyzed chemical reactions for cellular growth and maintenance [1, 2]. As catalysts, enzymes facilitate reactions by enabling alternate reaction mechanisms with lower activation energy, but in no way modify the thermodynamic equilibrium constant or the free energy change of a chemical transformation. They generate enormous kinetic rate accelerations, often exceeding factors of 10^{12}-fold relative to the rate of the uncatalyzed reaction. Enzymes are capable of performing many different chemistries, can be produced on a large scale, and typically operate at ambient temperatures and near neutral pH [3–5]. These attributes have captured the attention of generations of scientists and engineers alike and enabled the dramatic growth of the enzyme industry over the past century.

All commercial enzymes are proteins, which are assembled in living cells from a combination of the 20 naturally occurring amino acids, according to specific sequences defined by genes within the cell's DNA. Additional structural features are often present, such as carbohydrates, metal ions, and coenzymes which influence both the physical and catalytic properties of enzymes. The biosynthesis of a given enzyme is typically subject to several layers of cellular regulation and control; however, these controls can often be genetically or environmentally manipulated when overproduction of an enzyme is desired.

Enzymes have been used quite extensively as industrial catalysts and the current trend toward green, more sustainable processes is driving strong growth of this industry [5–9]. They possess the following attributes;

1. They are specific in action, and thus minimize the occurrence of undesirable side reactions
2. They are relatively inexpensive and can perform under mild conditions
3. They are effective for chemical conversions (biocatalysis)
4. They can be produced on a large scale by fermentation from common sugars and other renewable substrates
5. They can be modified and improved to suit a particular application
6. They are safe and acceptable for applications in food processes and medicinal treatments
7. They are effective over a wide range of concentrations

Over 3,000 different enzymes have been isolated and characterized, although less than 100 have ever been used in industrial applications. Enzymes are often named according to the substrate they act upon, for example, cellulose degrading enzymes are called cellulases and those that degrade starch, a mixture of amylase and amylopectin, are referred to as amylases. Many individual enzymes also have common names such as subtilisin, a family of serine proteases, and the digestive proteases chymotrypsin, rennin, and pepsin. Another example is lysozyme, a glycosidase that hydrolyzes peptidoglycan, a key component of bacterial cell walls. Enzymes that act upon polymeric substrates are also described according to their site of action—polysaccharide-degrading enzymes that act within the polymer chain are referred to as endoglycosidases and those that remove sugar units from the termini of the substrate as exoglycosidases. A description of the system used to formally classify enzymes is given later in the chapter. Commonly used industrial enzyme classes, their substrates, and products are listed in Table 31.1.

The world market for industrial enzymes exceeded \$US3.5 billion/year in 2010 and was growing at a rate of 5–8% per annum (Fig. 31.1) [10]. Over 50% of the sales come from proteolytic enzymes for use by the detergent, dairy, and leather industries [11]. The carbohydrases, mainly the amylases, isomerases, pectinases, cellulases, and hemicellulases, used in baking, brewing, fuel ethanol, starch, and textiles industries, represent nearly 40% of the total

J.C. McAuliffe (✉)
DuPont Industrial Biosciences, 925 Page Mill Road, Palo Alto, CA 94304, USA
e-mail: joseph.mcauliffe@dupont.com

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_31,

Table 31.1 Properties of industrial enzymes

Enzyme class	Substrate	Products
Amylase	α-1,4 Glycosidic bonds in starch	Maltodextrins, maltose, glucose
Catalase	Hydrogen peroxide	Water and oxygen
Cellulase	β-1,4 Glycosidic bonds in cellulose	Cellooligosacchardies, cellobiose
Glucoamylase	Nonreducing terminal α-1,4/6 glycosidic bonds in starch	D-Glucose
Glucose isomerase	D-Glucose	D-Fructose
Glucose oxidase	D-Glucose	D-Gluconolactone
Hemicellulase	β-1,4 Glycosidic bonds in hemicellulose	Oligo- and disaccharides, xylose
Lipase	Fatty acid esters	Fatty acids, glycerol
Pectinase	Pectin	Pectin oligosaccharides
Phospholipase	Phospholipids	Fatty acids, glycerol phosphate
Phytase	Inositol phosphate esters (phytate)	Inositol, phosphate
Protease	Peptide bonds in proteins	Peptides, amino acids
Pullulanase	α-1,6 Glycosidic bonds in amylopectin	Amylose

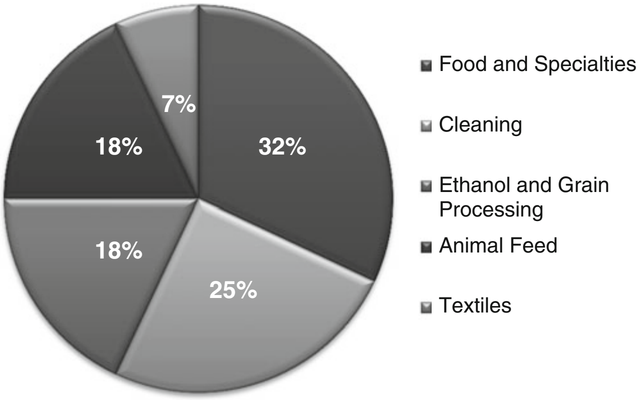


Fig. 31.1 Uses of industrial enzymes

enzyme market [12]. Lipases, phytases, oxidoreductases, and other highly specialized enzymes make up the remainder of total enzyme sales. While markets have matured in North America and Western Europe, regions including Latin America and Asia are still experiencing strong growth driven by demand from the animal feed, textiles, detergent, and grain processing industries. The major worldwide enzyme producers are Novozymes A/S and Genencor, now part of DuPont.

A multitude of non-industrial uses for enzymes have also been developed [13]. Although the total volume of enzyme protein for specialty uses is relatively small, the global market was estimated to exceed $2.5 billion by 2011 [14]. Enzymes used for analytical, clinical, and research purposes include glucose oxidase, hexokinases, pyruvate kinase, uricase, glucose-6-phosphate dehydrogenase, amino acid oxidase, aminopeptidase, and others. Restriction enzymes, endonucleases, and oligonucleotide polymerases are essential tools in Molecular biology and Forensic analysis. Glycosyltransferases and other carbohydrate-modifying enzymes are also routinely employed as tools for carbohydrate research [15]. Another growing application for enzymes is in the field of industrial biocatalysis, whereby enzymes are used for chemical and pharmaceutical manufacture, often displacing traditional chemical catalysts [5, 13, 16–18].

The first section of this chapter outlines the history of the enzyme industry, the general properties of enzymes (classification, structure, catalysis, and kinetics), and the methods used for enzyme discovery and optimization. The second section of the chapter describes the production and applications of industrial enzymes, with a focus on those of greatest commercial importance. The application of biocatalysis in industry is covered in the third section, with an emphasis on the processes employed and the resulting products. The interested reader can find more detailed information in the books and reviews cited throughout the text.

Industrial Enzymes: General Properties

History of Enzyme Use

The history of enzyme use dates back to ancient times and many cultures applied enzymes from various natural sources for the production of beverages, foods, and textiles [2, 5]. It was not until the nineteenth century that the basic nature of enzymes was investigated and their specific actions characterized. Diastase, an amylase mixture prepared from germinating barley, was described by Anselme Payen and Jean-François Persoz in 1833. Soon after, in 1835, Theodor Schwann isolated the digestive enzyme pepsin. As early as 1836, the Swedish chemist Jöns Jacob Berzelius speculated that thousands of catalytic processes took place in every organism and hypothesized that “catalytic bodies” were responsible. A mixture of protease and carbohydrases derived from solid-phase cultures of the fungus *Aspergillus oryzae* was developed in 1894 by Jokichi Takamine and is still sold under the brand name Takadiastase.

The name "enzyme" itself was coined in 1878 by Wilhelm Kühne, derived from the Greek term meaning "in yeast." Subsequent work by Emil Fischer beginning in 1894 led to a better understanding of enzymatic action, particularly the concept of substrate specificity. Further studies by Eduard Buchner demonstrated alcoholic fermentation in a cell-free system and he postulated that enzymes were responsible, although this theory was not widely accepted at first [19]. The protein nature of enzymes was demonstrated by James B. Sumner in 1926 through the crystallization of urease, a contribution for which he received the 1946 Nobel Prize [20]. The precise structure of enzymes and proteins in general remained elusive until 1965 when the 3D structure of lysozyme was finally deduced by Phillips and coworkers through X-ray crystallography [21].

Despite an incomplete understanding of enzymes, there was considerable use of crude enzyme preparations before 1950, primarily those derived from animal sources. Applications included cheese-making, leather production, brewing, cleaning, and the manufacture of malt extract. The primary limitation to the further development of enzymes was the inability to produce them on a large scale. This began to change after the mid-part of the twentieth century through the implementation of fermentation-based processes for the manufacture of bacterial amylases and proteases.

The development of large scale processes for the production of glucoamylase and proteases had significant impact within the starch processing and detergent industries, respectively. The ability to formulate enzymes and minimize sensitization of workers in enzyme production was another important development in the early 1970s, meeting a challenge which had significantly impeded the expansion of the enzyme industry. The advent of gene engineering in the late 1970s and early 1980s heralded an expansion of enzyme technologies that continues to this day. The first recombinant industrial enzyme product, a subtilisin protease variant, was marketed by Genencor International, Inc. in 1988 [22]. Additional engineered enzyme products were developed in the 1990s, as well as improved means to discover, engineer, and produce enzymes. In the last decade, the ability to rapidly sequence DNA and information technologies such as Bioinformatics have provided new tools for enzyme development. Powerful analytical methodologies including mass spectrometry, nuclear magnetic spectroscopy (NMR), and X-ray crystallography are now routinely used to solve enzyme structures and analyze the mechanisms underlying enzymatic catalysis [23]. The ability to design enzymes de novo is not yet routine, although some examples have been reported and further developments are highly likely.

The number of applications for enzymes has steadily expanded over the past 40 years and enzymes have gained wide use in the detergent, textiles, grain-processing, animal feed, pulp and paper, fuel ethanol, and chemical industries [5–10]. In addition to the economic benefits that enzymes offer, the need for clean and more sustainable industrial processes will inevitably stimulate new technologies for enzyme development and application.

Enzyme Structure and Properties

Enzymes are complex, three dimensional macromolecules that contain the chemical environment necessary to catalyze a particular reaction mechanism, in addition to possessing a template function that limits the set of possible substrates and the resulting products [3, 4]. All enzymes contain a defined region within their structure, known as the active site, where catalysis takes place. Each enzyme class catalyzes a specific reaction or a group of reactions with certain common characteristics. Many enzymes possess the ability to catalyze reactions with very high stereoselectivity in that they favor the production of only one of two possible stereoisomers. As macromolecules, enzymes can lose their catalytic properties when subjected to agents such as heat, strong acids or bases, organic solvents, or other conditions that break non-covalent bonds, a process called denaturation. Some enzymes can withstand quite extreme conditions, for example, temperatures exceeding 120°C and pH values of less than 1 [24]. The ability of an enzyme to turn over a substrate is referred to as enzyme activity. The activity of an enzyme is a function of the intrinsic nature of the enzyme itself, as well as environmental factors such as substrate concentration, temperature, and pH. All enzymes exhibit optima with regard to pH and temperature. Outside of these optimal regions, enzymes become less efficient and can undergo irreversible denaturation. Enzyme activity can be measured using an enzyme assay, where the concentrations of substrates and/or products are determined after a given period (end-point assays) or as a function of time (kinetic enzyme assays) [25].

All proteins, including enzymes, consist of a linear sequence of amino acids, referred to as the primary sequence. The N-terminal of the sequence possesses a free α-amino group, and the C-terminus contains a free α-carboxylic acid function. A typical enzyme is composed of several hundred amino acid residues with molecular weights in the 20–100 kDa range. The smallest enzyme yet described is 4-oxalocrotonate tautomerase, consisting of only 62 amino acids, whereas the large eukaryotic fatty acid synthase complex is composed of over 2,500 amino acid residues [26, 27]. Commercial enzymes such as proteases, amylases, cellulases, and lipases typically consist of 200–500 amino acids.

A key feature of enzymes and proteins in general is the ability to fold into more complex structures. It is the primary sequence of an enzyme that determines the 3-dimensional

Fig. 31.2 Ribbon diagram of a protease

structure of an enzyme, although the protein folding process is complex and not yet fully understood [28]. Secondary structure refers to features such as alpha-helices, beta-sheets, and random coils which are defined by hydrogen bonding between different amino acid residues in the primary sequence. The identity and order of amino acids defines the secondary structural features. These secondary features fold further to form the tertiary structure of an enzyme, which is stabilized by a combination of both non-covalent (hydrogen, ionic, and hydrophobic bonds) and covalent bonds (disulfide bridges). The tertiary structure of an enzyme can be depicted in several ways, the most common being a ribbon diagram where the orientation of the amino acid backbone is represented, but amino acid side chains are omitted for clarity. A ribbon diagram of a protease is depicted in Fig. 31.2. Space filling models where all atoms are depicted are also common, particularly when visualizing the active site and other binding regions of an enzyme. Quaternary structure refers to complexes formed from two or more fully folded proteins. Such complexes can take the form of dimers, trimers, and higher order assemblies. These complexes can be formed from many sub-units of the same enzyme (e.g., homodimers and trimers) or several different molecules (e.g., heterodimers, trimers, tetramers, etc.). Many enzymes are catalytically inactive in monomer form and only express activity in quaternary complexes. Protein and enzyme structural hierarchy is depicted in Fig. 31.3.

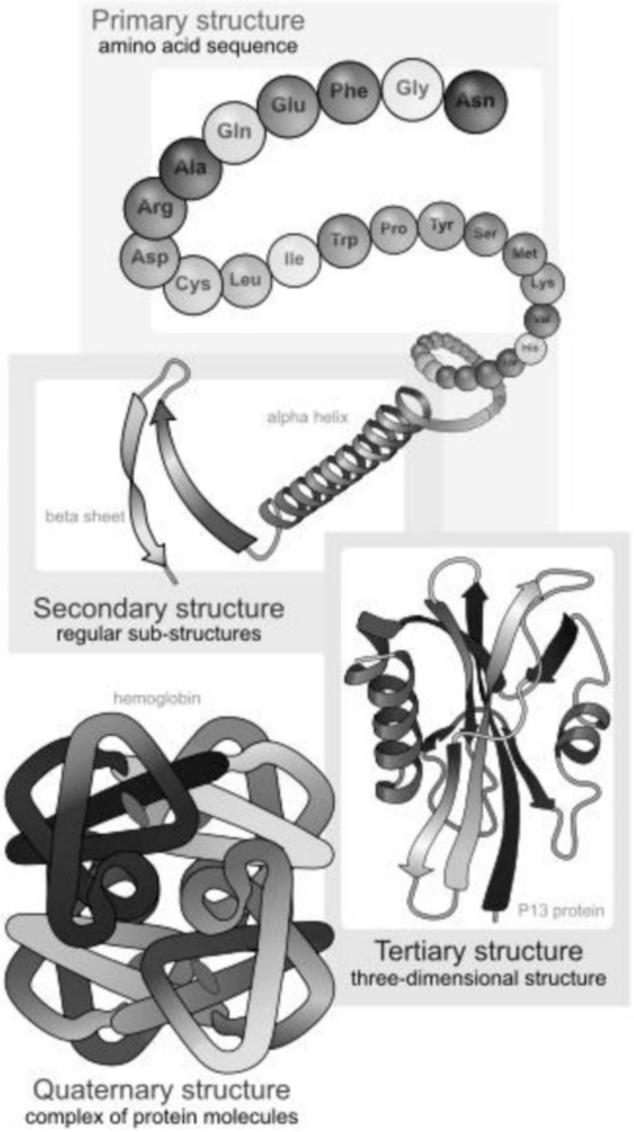


Fig. 31.3 Protein and enzyme structural hierarchy

Many enzymes contain additional structural features beyond the amino acids that define the primary structure. These elements are called posttranslational modifications resulting from subsequent processing following the biosynthesis of the primary sequence. The two most common types of posttranslational modifications are phosphorylation and glycosylation with mono- and oligosaccharides. Glycosylation is of particular interest as the carbohydrate chains can account for a large proportion of an enzyme's molecular weight. Glycosylation is described as either O-linked (serine and threonine residues) or N-linked (asparagine and glutamine residues) depending on whether the glycan chain is attached to the protein backbone through oxygen or nitrogen, respectively. Enzymes made in fungal hosts such as *A. oryzae*, *A. niger,* and *Trichoderma reesei* are often glycosylated, in contrast to those produced in bacterial cells. Enzymes often contain additional structural elements

Table 31.2 The six families of enzymes that comprise the enzyme classification system

Group	Reaction catalyzed	Examples
EC 1: Oxidoreductases	Transfer or removal of electrons or their equivalents from one substrate to another	Glucose oxidase, catalase, laccases
EC 2: Transferases	Transfer of a functional group from one substrate to another (e.g., acyl, methyl, amino, glycosyl, etc.)	Transaminases, acyltransferases, glycosyltransferases
EC 3: Hydrolases	Breaking of chemical bonds by addition of water to form two products	Proteases, amylases, cellulases
EC 4: Lyases	Non-hydrolytic scission of a substrate to form two products	Decarboxylases, aldolases, carbonic anhydrase
EC 5: Isomerases	Rearrangement of substrate molecular structure	Xylose isomerase, amino acid racemases
EC 6: Ligases	Formation of a product from two substrate molecules	DNA ligase, aminoacyl-tRNA synthetases

including metal ions, phosphate and fatty acid esters, and organic cofactors (coenzymes) like heme, cobalamin, and flavin adenine nucleotide. In many instances, these modifications have regulatory or catalytic functions.

The primary structure of an enzyme can be determined by a process known as sequencing, with Edman degradation being the most widely used technique. Mass spectrometry is now used routinely to determine the amino acid sequence of an enzyme, as well as modifications to the amino acid backbone [29]. Several analytical techniques are used to determine enzyme 3-dimensonal structure, including circular dichroism for secondary structure and X-ray crystallography for tertiary and quaternary structure. NMR is also an invaluable tool in enzyme structure determination and can be used to observe enzyme dynamics [30].

Enzyme Classification

Enzyme nomenclature is defined according to the Enzyme Commission number (EC number), first developed in 1955 at the International Congress of Biochemistry in Brussels. The system is based upon the nature of the reaction that the enzymes catalyze, as opposed to structural or other characteristics. There are six major groupings within the EC system, each of which contains many subtypes. Enzymes that perform hydrolysis reactions (EC 3) such as proteases, amylases, lipases, and cellulases dominate the current markets for industrial enzymes, although examples from each of the other classes have also been commercialized, particularly in the area of biocatalysis (oxidoreductases, lyases, isomerases) and diagnostics (oxidoreductases, ligases). Table 31.2 summarizes the EC classes and the reactions they catalyze. The EC number is also referred to as the IUBMB (International Union of Biochemistry and Molecular Biology) number. The IUBMB periodically updates enzyme nomenclature and lists the corresponding CAS number for each IUBMB number. More detailed information can be found at the International Union of Biochemistry and Molecular Biology enzyme nomenclature website (www.chem.qmul.ac.uk/iubmb/enzyme) and the ExplorEnz database (www.enzyme-database.org).

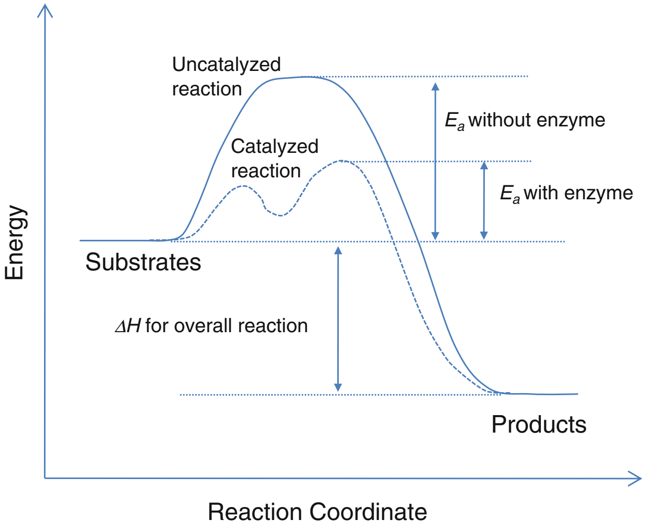


Fig. 31.4 A comparison of the reaction coordinates of a catalyzed vs. an uncatalyzed reaction. A catalyst provides an alternate reaction pathway with lower activation energy

Enzyme Catalysis and Kinetics

The basic principle by which enzymes and chemical catalysts in general operate is their ability to lower the activation energy (E_a or $\Delta G^{\ddagger}$) for a chemical reaction by altering the reaction coordinate, thus facilitating the rate at which the conversion of substrates to products proceeds [31]. Lowering of the activation energy does not change the equilibrium position of a reaction and accelerates both the forward and reverse directions of a chemical reaction (Fig. 31.4).

The relation between activation energy, temperature, and reaction rate is defined by the Arrhenius equation (31.1) where R is the universal gas constant, T is temperature in

Kelvin, k is the reaction rate coefficient, and A refers to the frequency factor. The equation demonstrates why catalysts can produce dramatic rate accelerations. For example, an enzyme that lowers the activation energy of a reaction by 10 kJ/mol at 25°C leads to a rate increase of around 50-fold, whereas a 40 kJ/mol reduction in E_a gives a rate acceleration of 10^7 fold [32].

$$k = Ae^{-(E_a/RT)} \tag{31.1}$$

Reaction rates tend to increase with temperature and enzyme catalyzed reactions are no different, exhibiting a rate increase of approximately twofold for every 10°C rise in temperature. The physical stability of an enzyme is also related to temperature, a concept known as thermostability. Thus, enzymes display temperature optima with regard to their activity. Some enzymes have broad temperature optima, while others display maximal activity over narrower ranges of temperature. Enzymes from animal sources typically have optimal temperatures around 37°C, whereas those from microbial sources have widely varying optima, often correlating to the environment from which the source organism was isolated.

The action of enzymes can be divided into three distinct events—initial substrate recognition, followed by enzymatic catalysis and subsequent product release. The binding of a substrate to an enzyme is often highly specific and the nature of this event was first explained through the lock and key model, developed by Emil Fisher in 1894, which held that the structures of the substrate and enzyme were both rigid and complementary, such that similar but different substrates would not bind with the same affinity, relative to the preferred substrate. The induced fit model was put forward by Koshland in 1958 to overcome the shortcomings of the earlier theory and proposed that the conformations of both enzyme and substrate adapt upon binding in order to produce an enzyme/substrate complex that undergoes catalytic conversion to an enzyme/product complex. This model more readily explains how an enzyme is able to produce and stabilize the transition state that lies between a substrate and the resulting product during the catalytic process.

A number of theories have been proposed to explain how an enzyme lowers the activation energy for a chemical reaction and the topic is still the subject of much debate [33]. A common view is that enzymes stabilize the transition state of a reaction, thereby reducing the free energy of this state and promoting the rate of the reaction. Enzymes accomplish this through non-covalent interactions including hydrogen-bonds, Van der Waals, and electrostatic interactions. This hypothesis is supported by the observation that compounds designed to mimic transition states are often extremely potent enzyme inhibitors with far greater affinity for the enzyme relative to either the substrate or products of the enzyme. For example, Acarbose, a nitrogen-containing pseudo-tetrasaccharide, has a K_d for glucoamylase of 10^{-12} M, far below that of maltotetraose, a substrate for this enzyme [34].

Enzymes also promote reactions through the assistance of chemical functionalities present within the active site that participate in the reaction by acting as acids, bases, nucleophiles, and counter ions. Such catalytic residues consist of amino acid side chains or elements within cofactors. Examples include carboxylic acid groups from glutamic and aspartic acid that participate in glycosidic bond cleavage by glycosidases, serine, histidine, and aspartic acid that comprise the catalytic triad of proteases and the nicotinamide groups from the soluble redox cofactors NADH and NADPH. Metal ions are also commonly employed by enzymes to perform some of the more demanding feats of chemical catalysis, for example, zinc-containing metalloproteases, iron-based oxygenases, and molybdenum nitrogenases [35].

Yet another means by which enzymes promote reactions is through conformational changes of the protein which increases strain on bonds within the substrate and bring catalytic residues into close proximity to the substrate [36]. Such conformational changes can be induced through substrate binding or from the different vibrational modes of the protein itself. The role of enzyme dynamics in catalysis is a topic of current interest, as is quantum tunneling, whereby a proton or electron transfer event occurs at rates faster than one would predict from the known activation energy by tunneling through the energy barrier instead [37]. The enzyme aromatic amine oxidase is thought to utilize quantum tunneling to oxidize tryptamine.

Enzyme Kinetics

The kinetic parameters for a free enzyme in solution are readily derived using the Michaelis–Menten approach describing pseudo-steady-state conversions [4, 38]. Consider (31.2) representing the conversion of a substrate S into a product P, catalyzed by an enzyme E. The rate of formation of an enzyme/substrate complex, ES, is denoted as k_1, the reverse reaction by k_{-1}, and the rate of subsequent conversion to the free product by k_2.

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} \mathrm{ES} \xrightarrow{k_2} E + P \tag{31.2}$$

The Michaelis constant K_m under steady-state conditions is defined as follows:

$$(k_{-1} + k_2)/k_1 = K_m \tag{31.3}$$

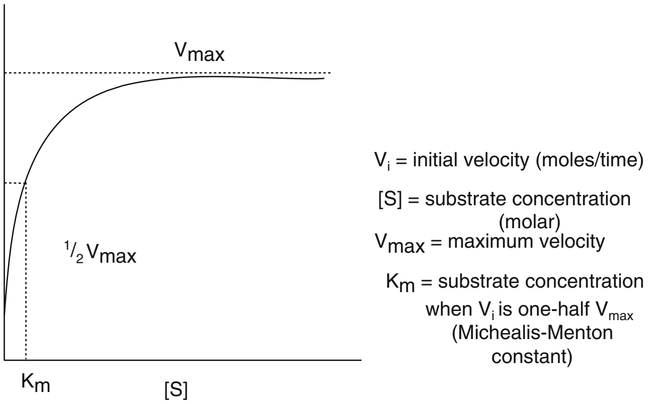


Fig. 31.5 Enzyme velocity vs. substrate concentration

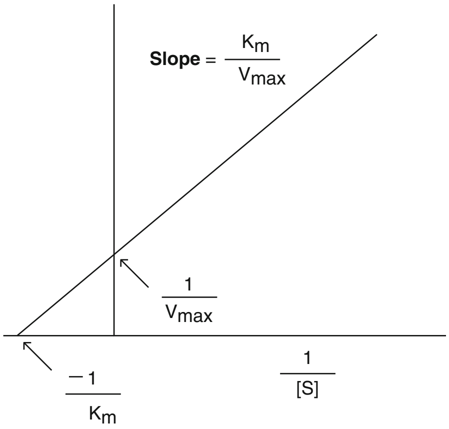


Fig. 31.6 A Lineweaver–Burke plot

The term k_{cat} can be substituted for k_2 and is referred to as the turnover number of an enzyme (units of s^{-1}). The expression $k_{\text{cat}}/k_{\text{m}}$ is widely used as a measure of the catalytic efficiency of an enzyme and is termed the specificity constant or turnover number. Where $[S] \gg K_{\text{m}}$, one can assume that all enzyme is bound to substrate (i.e., $[E_0] = [\text{ES}]$). Under these conditions, the maximal velocity of the reaction $V_{\max}$ is a function of k_2, represented in (31.4), where E_0 is the initial enzyme concentration.

$$V_{\max} = k_2[E_0] \tag{31.4}$$

From this one can derive an expression for the velocity of an enzyme catalyzed reaction as a function of the substrate concentration $[S]$, termed the Michaelis–Menten equation (31.5):

$$v = V_{\max}[S]/(K_{\text{m}} + [S]) \tag{31.5}$$

This relationship can be graphically represented by plotting reaction velocity vs. substrate concentration (Fig. 31.5) or alternately as a double reciprocal plot known as a Lineweaver–Burk plot (Fig. 31.6) where the x-intercept is $-1/K_{\text{m}}$, the y-intercept is $1/v_{\max}$, and the slope $K_{\text{m}}/V_{\max}$.

Although the Michaelis-Menten approach provides a useful approximation of enzyme kinetics, there are many situations in which the relation of reaction velocity to substrate concentration is more complex and cannot be adequately modeled using the Michaelis-Menten approach. In many instances, the substrate is a heterogeneous material, such as starch or biomass. There may be other enzymes present, resulting in synergistic effects. Many enzymes possess additional sites other than the active site where molecules can bind and alter activity, a phenomenon known as allosteric regulation. Allosteric regulation can either enhance or decrease enzyme activity and is a common form of regulation observed in metabolic networks.

Enzyme Inhibition

Enzyme inhibitors may also modulate the activity of the enzyme, altering the observed kinetic behavior of the system. An enzyme inhibitor is defined as a substance that can bind to an enzyme and reduce the catalytic rate. There are two overall types of enzyme inhibition—reversible and irreversible inhibition. Irreversible enzyme inhibitors often react with an enzyme forming a covalent enzyme-inhibitor complex. Fluorophosphonate nerve agents (e.g., sarin) are examples of irreversible inhibitors that can react with serine proteases and acetylcholine esterases. Suicide inhibitors are a type of irreversible enzyme inhibitor that is converted into the reactive form through the action of the enzyme they act upon.

In the case of reversible inhibitors, the formation of an enzyme inhibitor complex [EI] is conceptually similar to the enzyme-substrate complex [ES]. The inhibition constant K_{i} is used to describe the affinity of an inhibitor for an enzyme. Weak inhibitors have K_{i} values in the millimolar range and above, whereas potent inhibitors can inhibit enzymes on a stoichiometric (1:1) basis and have subnanomolar K_{i} values. Enzyme inhibitors can be both small and large molecules and can bind at the active site, substrate binding sites, or allosteric sites. Several types of enzyme inhibition are known and are summarized in Table 31.3. Examining the profile of Lineweaver-Burk plots at different inhibitor concentrations is a common means to determine the mode of inhibition.

Reversible enzyme inhibitors are often used to prevent unwanted enzymatic reactions, such as autoproteolysis in protease concentrates. Enzyme inhibitors can also be used to enhance the thermostability of enzyme formulations, provided the inhibitor is not so potent as to reduce the activity of the enzyme under use conditions in the presence of substrate [39].

Table 31.3 Different modes of reversible enzyme inhibition

Inhibition mode	Description	Effect
Competitive inhibition	Inhibitor and substrate compete for the active site of the enzyme	• K_m increases • V_{max} unchanged
Noncompetitive inhibition	Inhibitor binds elsewhere to the enzyme and does not compete with the substrate	• K_m unchanged • V_{max} decreases
Uncompetitive inhibition	Inhibitor only binds to enzyme-substrate complex. Relatively uncommon	• K_m decreases • V_{max} decreases
Mixed inhibition	A combination of competitive and uncompetitive inhibition	• K_m increases • V_{max} decreases

Enzyme Discovery and Engineering

When a suitable enzyme for a desired application is not available, it becomes necessary to develop a new or improved enzyme. Traditionally, this has been done by screening living microorganisms from either environmental samples or culture collections. With the recent advances in molecular biology and the explosion of genomic sequences, it is possible to isolate novel enzymes through genetic engineering of existing molecules, environmental gene screening, or by genomic database mining [40–43].

The classic method for discovering new enzymes involves screening environmental isolates and culture collections. This process was used very successfully as far back as the 1950s for the discovery of antibiotics such as penicillins and streptomycin. The technique involves examination and screening of thousands of isolates from a diverse array of environments such as soil, plant, and aquatic samples. The limiting factor in this process is the throughput capabilities. A significant investment in time is required to cultivate and screen the organisms, of which many will not have the desired activity. However, one can significantly enhance this process by subjecting an environmental sample to some selective pressure that will enrich for the desired activity. An example of this would be to take a soil sample and then provide a particular compound as the sole carbon source. This would enrich for organisms that are capable of breaking down the compound and using it as a carbon source. Another approach would be to select environmental conditions that would mimic the conditions of the biotransformation. Environments such as the hot springs, mud pots, and fumaroles found in Yellowstone National Park reach temperatures over 80°C and pH values lower than 3. Organisms such as *Sulfolobus solfataricus* that grow optimally at 80°C and pH 3 will also express extracellular enzymes that work optimally at these conditions. In addition, one the most famous research enzymes, Taq Polymerase, came from an organism, *Thermus aquaticus*, isolated from Yellowstone National Park.

One of the limitations of screening environmental isolates is the cultivation of the organisms. In a single gram of soil, more than 10^8 microbes may be present, the majority of which cannot be cultivated. It is clear from molecular ecological studies that traditional cultivation techniques only capture a fraction, usually less than 1%, of the available diversity. To overcome this, direct cloning of DNA from environmental samples enables an unbiased representation of the microbial diversity. The collective genomes of all microorganisms in a given environment is known as the "metagenome" [44]. Briefly, this involves the isolation of DNA from environmental samples and subsequent cloning into plasmids or bacterial artificial chromosome (BAC). The plasmid libraries can then be screened for the desired activity. A comparison to the traditional screening method is shown in Fig. 31.7. For example, Knietsch et al. incubated environmental samples in the presence of glycerol or propanediol to enrich for organisms and activities that would oxidize short-chain polyols or reduce the corresponding carbonyls [45]. Total DNA was isolated from the samples and approximately 100,000 *Escherichia coli* strains were screened for the desired activity. In total, 16 strains were isolated that stably expressed the desired activity and a number of novel genes were identified that encoded alcohol dehydrogenases and putative oxidoreductases.

A third approach to isolate novel enzymes is via genomic database mining. Genomic database mining is defined as the process of finding and extracting useful information from raw DNA sequence data sets. With the explosion of genomic sequencing, there is a wealth of information available to the researcher. The whole genomes of over 1,200 organisms can be found in databases such as ERGO (http://ergo.integratedgenomics.com) and NCBI (http://www.ncbi.nlm.nih.gov/genome). All three domains of life are represented, bacteria, archaea, and eukaryota, as well as viruses and various organelles (Table 31.4).

With the protein sequence of a molecule that has been characterized biochemically, one can easily identify a large number of sequence homologues in completed genomes. Genes of interest can be identified through BLAST analyses

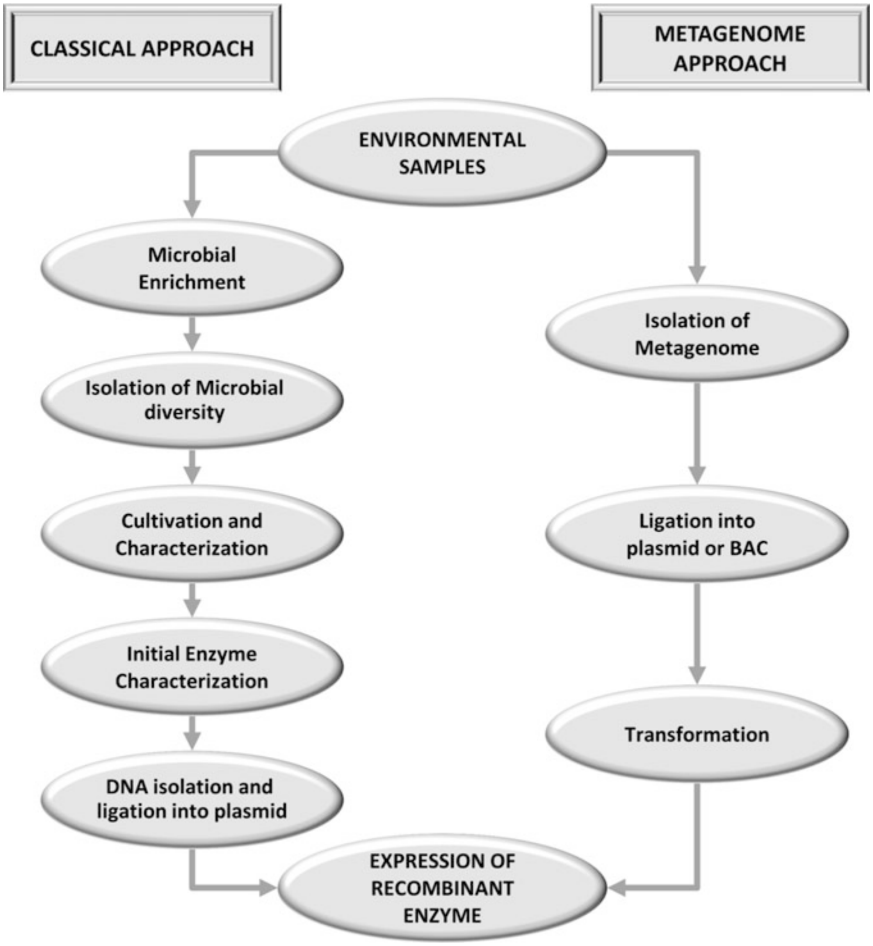


Fig. 31.7 Processes for the isolation, screening, and expression of enzymes

Table 31.4 Number of genomes available in ERGO and NCBI (2011)

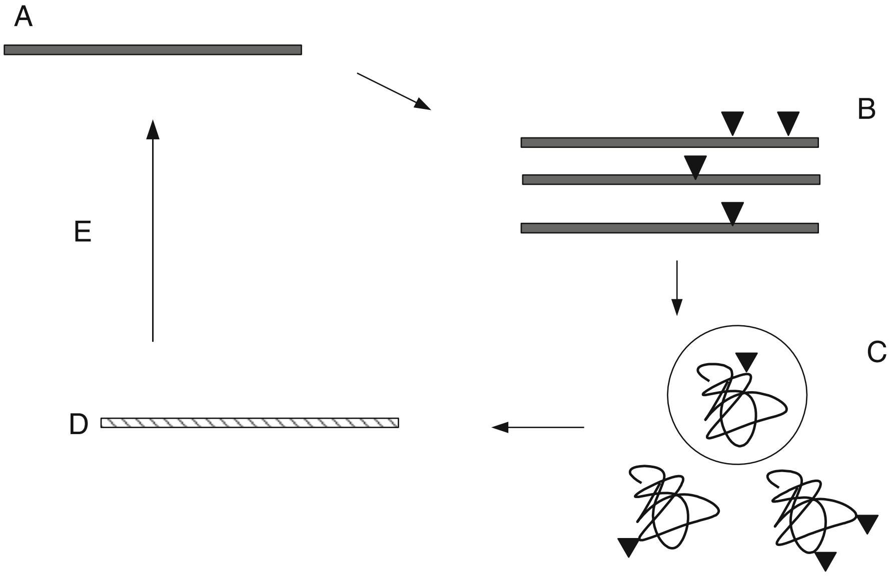
	ERGO	NCBI
All genomes	1,218	980
Bacteria	792	607
Archaea	49	44
Viral	241	213
Eukarya	136	116

of completely sequenced genomes. In addition, putative biocatalysts can be analyzed for known motifs using ScanProsite (http://ca.expasy.org/tools/scanprosite), the presence of signal peptides using SignalP (http://www.cbs.dtu.dk/services/SignalP/), and genomic organization using the STRING analysis method (http://string.embl.de), for possible enzyme–gene relationships. An example of this was done with the identification of proteases in the hyperthermophiles [46]. Initial efforts to assess the extent and variety of proteases in hyperthermophiles by biochemical methods significantly underestimated this biocatalytic feature. With the benefit of genome sequence data, it is clear that the proteolytic genotypes of these organisms are more expansive than can be inferred from biochemical analyses with over 30 proteases/peptidases identified in each genome.

The last approach involves engineering of an existing enzyme [47]. Molecular biology, high-throughput screening, and other analytical methods have made it possible to redesign and improve wild-type enzymes through the generation and screening of enzyme variant libraries. Site-directed mutagenesis is the method of choice for assessing the role of individual amino acids in the enzymes' catalytic mechanism or stability. A high-resolution crystal structure or a model based on the structure of a related protein is necessary to identify the residues to mutate. Using this technique, the amino acid of interest is either deleted or altered to another amino acid of choice. Subsequent biochemical analyses will yield information about the residues' role in the protein. This method has been an important tool in determining protein structure–function relationships and elucidation of catalytic and protein stabilization mechanisms.

Whereas site-directed mutagenesis is a technique that focuses on changing very specific residues based on the protein structure, directed evolution and gene shuffling

Fig. 31.8 Error-prone PCR for directed evolution of enzymes: (**a**) gene of interest; (**b**) error-prone PCR generating diversity; (**c**) expression of variants and characterization; (**d**) selection of variant with improved trait; (**e**) variant then serves as template for subsequent rounds of evolution



take a more global approach. An advantage of these methods is the limited knowledge of the proteins structure and mechanism that is required. The discovery of error-prone PCR was an important breakthrough in enzyme-variant generation. The mutation frequency of the error-prone PCR can be controlled at two levels, the reaction conditions and the choice of polymerase. An overview of this method is shown in Fig. 31.8.

Another method for generating diversity is gene shuffling. Using this method, one can start with either a single gene or a family of related genes. In the first case, error-prone PCR is carried out and the diversity of the library can be further increased through gene shuffling (Fig. 31.9a). The other alternative is to use multiple genes from various sources with significant levels of identity, usually >70% (Fig. 31.9b). In both cases, the genes of interest are fragmented by a DNAse treatment and then allowed to reassemble via recombination thereby generating a library of variants.

Industrial Enzymes: Properties and Applications

The global market for industrial enzymes is dominated by hydrolases from the EC 3 family, namely α-amylases, glucoamylases, proteases, cellulases, and lipases. Altogether they account for over 85% of the total sales of enzymes. In recent years, additional enzymes have shown strong sales growth, including phytase, an animal feed additive for increasing the bioavailability of phosphorous, laccase for the biobleaching of denim, phospholipases for edible oil production, and xylanases for biomass degradation. Applications of the industrially important enzymes are listed in Table 31.5, along with their microbial sources and commercial applications [6].

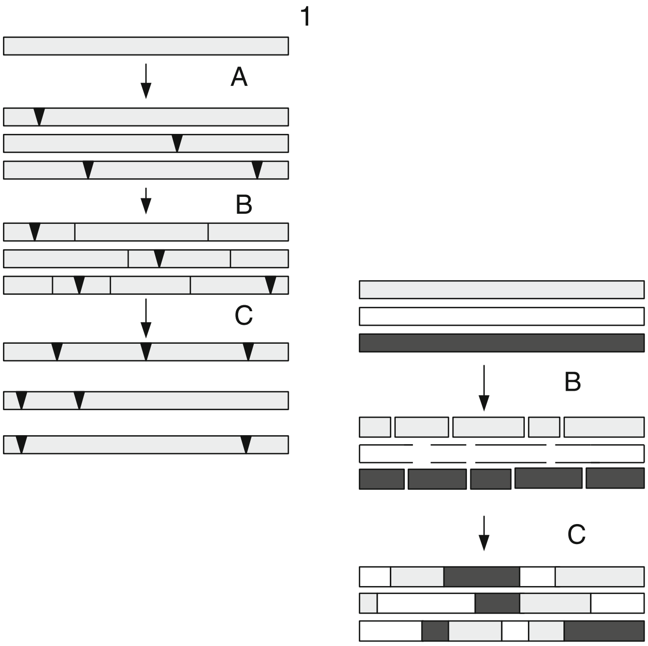


Fig. 31.9 Gene shuffling using a single gene (1) or a family of genes (2): (**a**) error-prone PCR generating diversity; (**b**) DNAse treatment; (**c**) homologous recombination of fragments and generation of diversity

Industrial enzymes are either extracted from natural sources (animal, plant, microbial) or manufactured though large-scale microbial fermentation processes. The latter method is responsible for the majority of the enzyme protein produced per year, accounting for over 85% of the total [48]. Submerged fermentation is the dominant fermentation technique, although solid state fermentation is still used for the production of some enzymes [49]. The advent of recombinant DNA technology has had a great influence on the production of enzymes for industrial purposes, both through improvements in the molecules themselves as well as the microbial hosts used in their production. The majority of technical enzymes are now produced in recombinant systems. This has offered many advantages to enzyme users. For instance, the price of enzymes has dropped significantly, while the purity of enzymes has increased overall. The improvement in purity was mainly achieved by the absence of enzymatic side activities, which has enabled much better control of enzyme activity in the industrial environment. General methods for the production, recovery, and formulation of enzymes by fermentation processes are covered in detail within the preceding chapter on Industrial Biotechnology (Chap. 30).

Despite the dominance of large-scale microbial fermentation processes for enzyme manufacture, many commercially relevant enzymes are still isolated from natural sources for both cost reasons, as well as consumer preference. The protease chymosin (rennin), used in cheese-making, is derived from rennet, a substance isolated from calf-stomach lining. While recombinant versions of chymosin expressed in microorganisms are available at lower cost, the demand for the animal-derived product remains strong, especially in Europe. Other animal-derived enzymes include lysozyme, derived from hen egg whites and the digestive aid Pancreatin, a mixture of amylase, lipase, and protease, manufactured from either cow or pig pancreatic extracts.

Plant-derived enzymes also have an important niche in the marketplace. For example, the cysteine proteases papain (papaya latex), ficin (fig latex), and bromelein (pineapple stalks) have been used as meat tenderizers, for contact lens cleaning, and as digestive aids. In the case of papain, latex from papaya trees is collected and dried, resulting in a crude extract. Further purification involves steps including extraction with water, reducing and chelating agents, salt precipitation, and solvent extraction. Highly pure material can be obtained through affinity purification using an immobilized substrate to isolate the enzyme from background components, although this form of purification is often rather costly. In recent years the FDA has restricted the use of papain for topical applications due to the risk of hypersensitivity and potential allergenicity. Malted barley is another important plant source of enzymes providing both α- and β-amylase, as well as β-glucanase, which are used on a very large scale for brewing and food applications. Lipoxygenase, an enzyme that oxidizes fatty acids, is isolated from soy beans. Production of industrial enzymes in recombinant plant systems has generated much interest, as well as controversy in recent years. While this mode of production has the potential to produce large volumes of enzymes directly from carbon dioxide and sunlight, there are concerns about contamination of non-GMO crops with enzyme-producing strains [50].

The major classes of industrial enzymes are discussed below, in addition to emerging enzymes of interest. Further information can be found in the reviews cited early in each section.

Proteases

Proteolytic enzymes (EC 3.4) are by far the most important of the commercially available industrial enzymes having found widespread use for laundry, automatic dishwashing, baking, diary, textiles manufacture, and other applications [11]. These enzymes, being essential parts of the metabolic system of most living organisms, can be isolated from innumerable sources. They can be divided into two overall categories—exoproteases (i.e., peptidases) that hydrolyze terminal peptide bonds releasing free amino acids, and endopeptidases that hydrolyze internal peptide bonds and convert proteins into smaller peptides. Proteases are also classified according to their catalytic mechanism, these subtypes are listed below;

- Serine proteases (EC 3.4.21)
- Cysteine proteases (EC 3.4.22)
- Aspartic acid proteases (EC 3.4.23)
- Metalloproteases (EC 3.4.24)

This classification is determined by amino acid sequence and through reactivity toward inhibitors that act on particular amino acid residues in the active site region of the enzyme. The serine proteases are widely distributed among microbes. The enzymes have a reactive serine residue in the active site and are generally inhibited by diisopropylfluorophosphate (DFP) or phenylmethylsulfonyl fluoride (PMSF). They are generally active at neutral and alkaline pH. The best-known serine proteases are the subtilisins, a large family of alkaline endoproteases which are the dominant enzymes used in cleaning applications owing to their wide temperature optima, broad substrate specificity, and resistance to denaturation by surfactants and other chemicals [51].

The occurrence of cysteine proteases is also widespread, occurring in both prokaryotes and eukaryotes. Examples include the plant-derived proteases papain and bromelein. These proteases are sensitive to sulfhydryl reagents, such as TLCK (tosyl lysyl chloromethyl ketone) and iodoacetic acid. Aspartic proteases are widely distributed among molds, including *Aspergillus*, *Penicillium*, and *Rhizopus*, but are less commonly found in bacteria and protozoa. They have their maximal activity at low pH, around 3–4. Many of the aspartic proteases are unstable above neutral pH and are not found in cultures growing at neutral or alkaline pH. Most

Table 31.5 Applications of industrial enzymes

Enzyme function	Origin	Applications
1. Oxidoreductases		
Glucose oxidase	*Aspergillus niger*	Increase gluten strength Improve shelf life Preserve milk Indirect enzymatic bleaching Tooth paste
Hexose oxidase	*Chondrus crispus*	Increase gluten strength
Laccase	*Trametes versicolor*, *Coprinus cinereus*	Prevent backstaining in enzymatic denim stone washing Pulp bleaching Cork treatment
Catalase	*Aspergillus niger*	Shelf life improvement of food Milk preservation Hydrogen peroxide removal
Peroxidase, lactoperoxidase	*Phanerochaete chrysosporium*	Dough improvement Milk preservation
Lipoxygenase	*Plants*	Whitening of breadcrumb
2. Transferases		
Transglutaminase	*Bacillus circulans*	Texture improvement in yoghurt and whipped cream Strength improvement of bread dough
Dextransucrase	*Leuconostoc mesenteroides*	Production of isomalto-oligosaccharide beer
3. Hydrolases		
α-Amylase	*Bacillus* sp.	Cleaning (laundry and automatic dish detergent) Baking (antistaling) Brewing (mashing, fermentation) Sweetneers Fuel ethanol Textiles processing
Cellulase	*Aspergillus* sp., *Trichoderma reesei*	Cotton finishing and depilling Biomass processing Pulp and paper Laundry detergent
Xylanase	*Bacillus* sp., *Aspergillus* sp., *Trichoderma reesei*	Baking (dough strengthening) Animal feed Lignocellulosic biomass processing Textiles processing
Protease (subtilisin)	*Bacillus* sp.	Cleaning (laundry and automatic dish detergent) Animal feed (improved digestibility) Leather production (soaking, bating, dehairing)
Protease (other)	Animal, plant and microbial sources	Cheese manufacture (Chymosin) Brewing Digestive aids
Lipase		Cleaning (removal of greasy soils) Paper and pulp (pitch removal) Biodiesel production Food applications (food oil production)
Pectin methylesterase	*Aspergillus* sp.	Apple and red berry juice pressing Citrus fruit peeling
Phytase	*Aspergillus* sp.	Mashing Releasing of phosphate from phytic acid in animal feed
Pullulanase	*Aerobacter aerogenes*, *Bacillus licheniformis*	Increase fermentability of starch worts or syrups
β-Amylase	Barley, *Bacillus cereus*	Mashing Fruit juice production
Glucoamylase, amyloglucosidase	*Aspergillus niger*	Apple juice production
		Mashing
		Hydrolysis of maltooligosaccharides
		Tooth paste

(continued)

Table 31.5 (continued)

Enzyme function	Origin	Applications
4. Lyases		
α-Acetolacetate decarboxylase	*Bacillus brevis*	Brewing (diacetyl removal)
Pectate lyase	*Aspergillus* sp.	Cotton scouring
Pectin lyase	*Aspergillus* sp.	Apple and red berry juice production Citrus fruit peeling
5. Isomerases		
Xylose isomerase	*Streptomyces* sp.	High fructose corn syrup production

aspartic proteases are sensitive to epoxy and diazoketone compounds. The pepsin-like aspartic proteases have been used commercially in processes such as soybean protein hydrolysis, and the rennin-like aspartic proteases have been used for clotting milk in a manner similar to animal rennins. The rennin-like proteases from *Mucor* and *Endothia* species have commercial applications in the manufacturing of cheese. The metalloproteases have pH optima between 5 and 9 and are sensitive to metal-chelating agents such as EDTA. Metalloproteases are widespread and most of the bacterial and fungal metalloproteases contain zinc. Calcium is required to stabilize their protein structure. The best-known metalloprotease is thermolysin produced by *Bacillus thermoproteolyticus*.

Protease Applications

Microbial proteases with widely different properties are produced commercially. *Bacillus* proteases, however, represent more than 95% of the sale of all proteases [52]. The most important use of *Bacillus* proteases is in detergents for both laundry and automatic dish (ADW) applications. Proteinaceous dirt often precipitates on clothes and it coagulates during the normal washing process. Similarly, protein soils can be difficult to remove from dishware and eating utensils through chemical-based detergents alone. The addition of proteolytic enzymes to the detergent cleaves such proteinaceous stains into soluble peptides, thus facilitating stain removal.

The alkaline serine protease of *Bacillus licheniformis*, also known as Subtilisin Carlsberg, (EC 3.4.21.62), a 274 amino acid protein, was the first protease to find widespread use in both nonionic and anionic detergents [53]. It readily dissolves proteins by cleavage of peptide bonds in a non-specific manner and is sold under the Trademark Alcalase® by Novozymes. It may be used at temperatures up to 65°C, and its pH optimum is close to 9.0, the pH normally used in washing fluids. Additional proteases have also been developed for solid and liquid detergents, including both wild-type and engineered enzymes, at both neutral and alkaline pH ranges. The serine protease Subtilisin Novo (BPN′) from *Bacillus amyloliquefaciens* has also found application as a detergent enzyme, crude versions of which contain a substantial content of α-amylase. An engineered version of BPN′ was developed by Genencor in 1984 by replacement of a methionine residue (Met 222) with amino acid residues less prone to oxidation by bleach (Ser, Ala and Leu) [22]. The subtilisins from other *Bacillus* species are also found in many commercial products, including *B. lentus*, *B. clausii*, *B. halodurans,* and *B. alcalophilus*. Commercially available products include Savinase® (Novozymes), Purafect® (Genencor), and Properase® (Genencor) [6].

Sales of proteases were small and relatively unimportant until about 1965. Since then, the use of proteases in detergents has enabled rapid expansion of the enzyme industry as a whole. However, allergic symptoms were discovered in some workers handling enzymes in detergent factories in 1971. The public, particularly in the United States, was alarmed, and proteases were taken out of most detergents. It was later found that every risk could be eliminated with proper precautions in handling by using proteases in liquid form or by encapsulating the enzymes.

Proteases are also used in the leather industry for soaking, bating, and dehairing processes [54]. Their use has enabled reductions in the amounts of hazardous chemicals including sodium sulfide used by the industry. The alkaline protease from *B. amyloliquefaciens* is used alone or in combination with sulfide for hide treatment and dewooling. A protease from an alkalophilic *Bacillus* sp. is successfully used for dehairing of ox hides in combination with lime, apparently because the enzyme is stable at pH as high as 12. Serine proteases are also used for wool scouring, a process which involves the removal of the scales naturally present on wool fibers, thus reducing the tendency of wool garments to shrink upon washing.

In the brewing industry, there is a development toward substitution of malted grain with unmalted barley and exogenous enzymes, including α-amylase, β-glucanase, and proteases of microbial and plant origin. The neutral protease from *B. amyloliquefaciens* and the thermostable neutral protease *Bacillus subtilis var. thermoproteolyticus* have been used by brewers successfully to hydrolyze barley proteins into amino acids and peptides in order to reduce foaming and haze in the finished product. The choice of enzyme and the condition applied are important as the production of certain peptides can impart undesirable flavors that should be avoided. The use of carbohydrases in brewing is covered in the "Amylases" section below.

Proteases are used in many food applications, primarily baking and cheese manufacture, as digestive aids and in the preparation of protein hydrolyzates, such as those derived from soybean protein. These proteases are derived from a range of different organisms, including *Bacillus subtilus*, *Streptomyces griseus*, *Aspergillus niger*, *Aspergillus flavusoryzae*, *Mucor miehei, and Mucor pusillus.*

Rennin, an aspartyl protease mentioned earlier, is found in the fourth stomach of nursing calves. Because of a decline in veal consumption and an increased demand for cheese, the dairy industry found it difficult to obtain sufficient quantities of this enzyme from natural sources. The increasing market for rennin (around $100 million/year) provided the necessary incentive for commercializing a recombinant DNA-derived chymosin. Prochymosin, the self-processing precursor of chymosin, was expressed in *E. coli* as an insoluble, refractile protein aggregate. The cells were lysed, and the insoluble prochymosin mass was isolated by centrifugation. The prochymosin was solubilized by 8 M urea with high pH treatment, renatured, and activated, and then chymosin was purified by ion exchange chromatography. The recombinant *E. coli* chymosin is no longer commercially available. Today recombinant chymosin is produced in the microorganisms *Kluyveromyces lactis*, *A. niger,* and *T. reesei*. These organisms secrete prochymosin. After prochymosin is separated from the cell broth, a pH drop induces the autocatalytic cleavage of prochymosin, which yields the active chymosin. This process is much cheaper than the former downstream processing from *E. coli* fermentations.

In addition to the recombinant chymosin, several coagulants for cheese-making are available from microbial origin. These are endoproteases from *Rhizomucor miehei*, *Rhizomucor pussilus*, and incidentally the plant-derived endothiapepsin from *Cryphonectria parasitica.*

Protease Production

The protease fermentation of *Bacillus* bacteria takes place under strictly contained conditions in conventional equipment for submerged fermentations. The aeration rate is about 1 vvm (volume of air per volume of medium per minute). Vigorous agitation is used to improve air distribution and oxygen transfer. The fermentation temperature is around 37°C, and the time cycle is 2–4 days. The composition of the fermentation medium is important to the yields of protease. Proteins of many different sources are used in commercial media. Carbohydrates are used as an energy source. The C/N ratio is important to the success of the process. Protein should be present in high concentration, and carbohydrate must not be in excess. A convenient way of obtaining this is to conduct fed-batch fermentation, feeding carbohydrate during the run and maintaining the carbohydrate concentration below 1%. Continuous fermentation of protease on a commercial scale is currently not practiced. The recovery and finishing of *Bacillus* protease involves the following steps.

1. Cooling to about 4°C to prevent microbial spoilage.
2. Precipitating undesirable salts using flocculants or filter aids.
3. Removing particles by centrifugation or filtration.
4. Removing pigments and odors with activated carbon treatment.
5. Concentration by ultrafiltration to product an ultrafiltered concentrate (UFC).
6. Sterile filtration to reduce microbial contamination.
7. Addition of formulation chemicals in order to produce a stabilized concentrate.
8. If a solid protease product is desired then granulation technologies are used.

Most, if not all, proteases are now produced on an industrial scale using genetically engineered organisms that overexpress either the wild-type enzymes (through homologous expression or self-cloning) or variants that have been obtained through protein engineering. The major reason for using the genetically engineered organisms is to reduce production cost. The overexpressing Bacilli can secrete proteases into the fermentation broth in very high yields. The secreted protease becomes the dominant protein in the broth. This greatly facilitates and reduces the cost of downstream processing and purification.

Amylases

The commercial importance of amylolytic enzymes is now well established and has driven significant expansion in the grain processing and fuel ethanol industries [55–58]. Amylases are also used in combination with proteases for cleaning applications. These enzymes catalyze the hydrolytic reactions of amylose (unbranched starch), amylopectin (branched starch), glycogen (animal starch), and related glucose polymers. Amylases, according to their difference in modes of action, can be divided into;

1. α-Amylase (EC 3.2.2.1), an endoglycosidase, hydrolyzes α-1,4-linkages randomly to yield a mixture of oligosaccharides, maltose, and glucose
2. β-Amylase (EC 3.2.1.2), an exoglycosidase, cleaves successive maltose units from the nonreducing end of starch

3. G4-amylase (EC 3.2.1.60), an exoglycosidase, hydrolyzes maltotetraose (G4) units from the non-reducing ends of starch
4. Glucoamylase (or amyloglucosidase) (EC 3.2.1.3) sequentially liberates glucose units from the non-reducing termini of amylose and amylopectin
5. The debranching enzymes such as pullulanase (EC 3.2.1.41) attacks the α-1,6-linkages at the branching point of amylopectin

The properties of amylases listed above have led to their use in many applications, including;

1. To produce grain syrup, glucose syrup, liquid glucose, and crystalline glucose
2. To produce high-fructose corn syrup (HFCS) in combination with glucose isomerase
3. To solubilize and saccharify starch for ethanol production in brewing, distilling, and fuel industries
4. To modify the viscosity of starch used in coating printing papers
5. To remove starch sizes applied to cotton thread before weaving in the textiles industry
6. To produce maltose-containing syrups in brewing and baking industries
7. To reduce the viscosity of sugar syrups used in various food and sugar products
8. As a component in digestive aids
9. To remove starch stains on clothes/dishes

α-Amylase is produced commercially by using both fungal and bacterial species. Fungal amylase has relatively low heat stability and its major application is in the baking industry to supplement the variable activity of amylase present in wheat flour. Bacterial amylases are much more heat stable and it is used in brewing, starch degradation, alcohol, textiles, and detergent industries. The organisms commonly used for the commercial production of α-amylase include:

Fungi
  Aspergillus oryzae
  Trichoderma reesei
Bacteria
  Bacillus subtilis
  Bacillus licheniformis
  Bacillus amyloliquefaciens

A. oryzae can be grown in either semisolid or submerged culture. In semisolid culture, it produces several enzymes, primarily α-amylase, glucoamylase, lactase, and protease. In submerged culture, the production of α-amylase is increased and the formation of other enzymes becomes minimal. The use of this fungal amylase in the baking industry speeds up the yeast (*Saccharomyces carlsbergenis*) fermentation; produces stiffer, more stable dough; and improves the texture, porosity, digestibility, and shelf life of bread. The fungal α-amylase delivers its optimal activity at pH 5–7 and at 50–55°C.

In the last decade, both protein engineering and recombinant DNA technology have been used to improve the thermostability of amylases by increasing the negative charges at its calcium-binding site and to improve the enzyme's resistance to oxidative compounds present in bleach containing detergents by substituting the oxidation-sensitive methionine residue with leucine. Genencor and Novozymes market their improved enzymes under the trade names of PuraStar® OxAm and Duramyl®, respectively.

Amylase Production

Different amylase-producing organisms may require different fermentation conditions for optimal enzyme production. When *B. subtilis* is used, the fermentation medium may contain starch, cornsteep liquor, yeast, phosphate, and some mineral salts. The amylase treatment on starch is often short to prevent the significant accumulation of glucose, which is inhibitory to the *Bacillus* amylase fermentation. The fermentation is run at neutral pH and at around 35°C. Care must be taken to prevent contamination by unwanted microorganisms. The time cycle is about 48 h. Whole mash may be used directly for starch liquefaction and saccharification or the mash may be processed to produce liquid or crystal enzyme preparation with high purity. The processing, involving filtration or centrifugation of the bacterial fermentation broth, presents a challenge to the recovery plant. Pretreatment with coagulating or flocculating agents is often needed. The amylase produced by this *Bacillus* strain is relatively unstable, but the addition of calcium chloride improves the stability.

Fungal glucoamylase, an exoglucosidase, drives the further degradation of starch following initial solubilization by α-amylase (often of a bacterial origin). Both *A. niger* and increasingly *T. reesei* are used for the production of glucoamylase. The fungal fermentation starts with a medium containing 25–30% starch and around 10% cornsteep liquor. Incremental or continuous feeding of concentrated nutrients may be used to circumvent the problems caused by a concentrated initial medium. The fermentation pH is about 4 and temperature around 28°C. The fermentation has a high oxygen demand. High oxygen tension, however, inhibits enzyme production. Levels of dissolved oxygen (DO) that approach zero are not atypical in this fermentation. After the completion of the fermentation in 4–5 days, the fermentor mash is cooled and filtered to remove cells and insoluble matters. *Trans*-glucosidase may be removed using clay, destroyed preferentially using proteases at certain pHs and temperatures, or inactivated by magnesium oxide. Contamination of glucoamylase by the transglucosidase activity may result in the loss of 5–10% of glucose to isomaltose and panose by a reversion process. These reversion products also impede the crystallization of glucose.

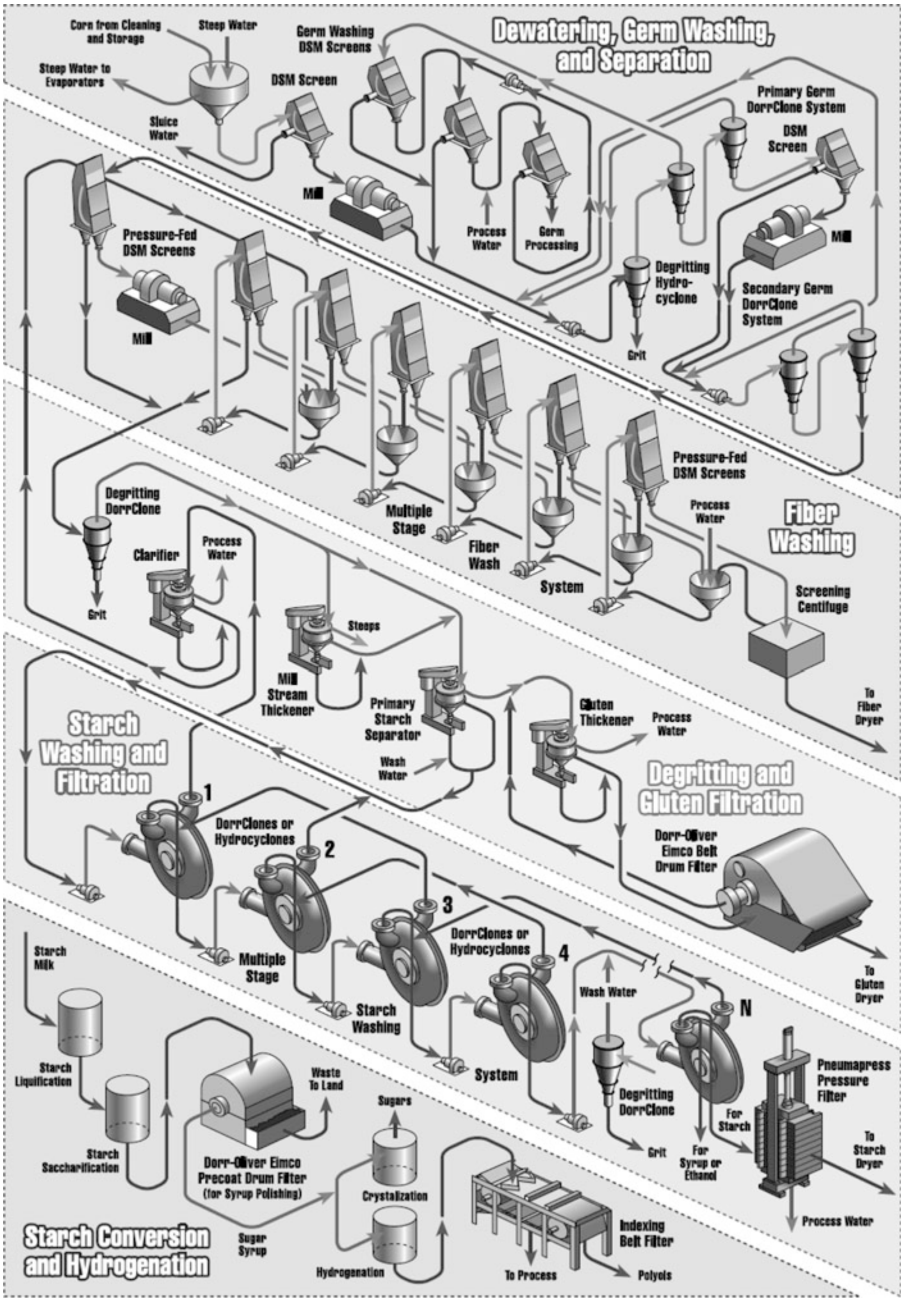


Fig. 31.10 Flowsheet of wet corn milling process (Copyright FLSmidth Co. and reproduced with permission)

Amylases in Wet and Dry Milling Processes

The enzymatic hydrolysis of starch to glucose has long since superseded the acid hydrolysis route using hydrochloric acid. The enzymatic process produces fewer side products, does not involve a corrosive acid, and allows the use of less pure starch products whose protein contaminants would, upon acid hydrolysis, give amino acids and browning reactions. Fungal glucoamylase in combination with bacterial α-amylase makes a complete enzymatic mixture for hydrolysis of starch to glucose. Corn is the dominant source of starch in the United States and is processed by either wet or dry milling processes to produce ethanol, starch, dextrins, corn syrup, glucose, and fructose syrups, in addition to byproducts of significant value such as corn oil and dry distillers grains and solubles (DDGS), a protein-rich animal feed [59].

Corn wet milling involves initial separation of the grain kernel into the germ, fiber, protein, and starch components prior to fermentation. The corn wet milling process is depicted in Fig. 31.10. Liquefaction of starch to form maltodextrins is aided by the action of bacterial α-amylase, whereas fungal glucoamylases are used in order to saccharify maltodextrins to glucose. Prior to liquefaction, granular starch is gelatinized by heat treatment at temperatures of up

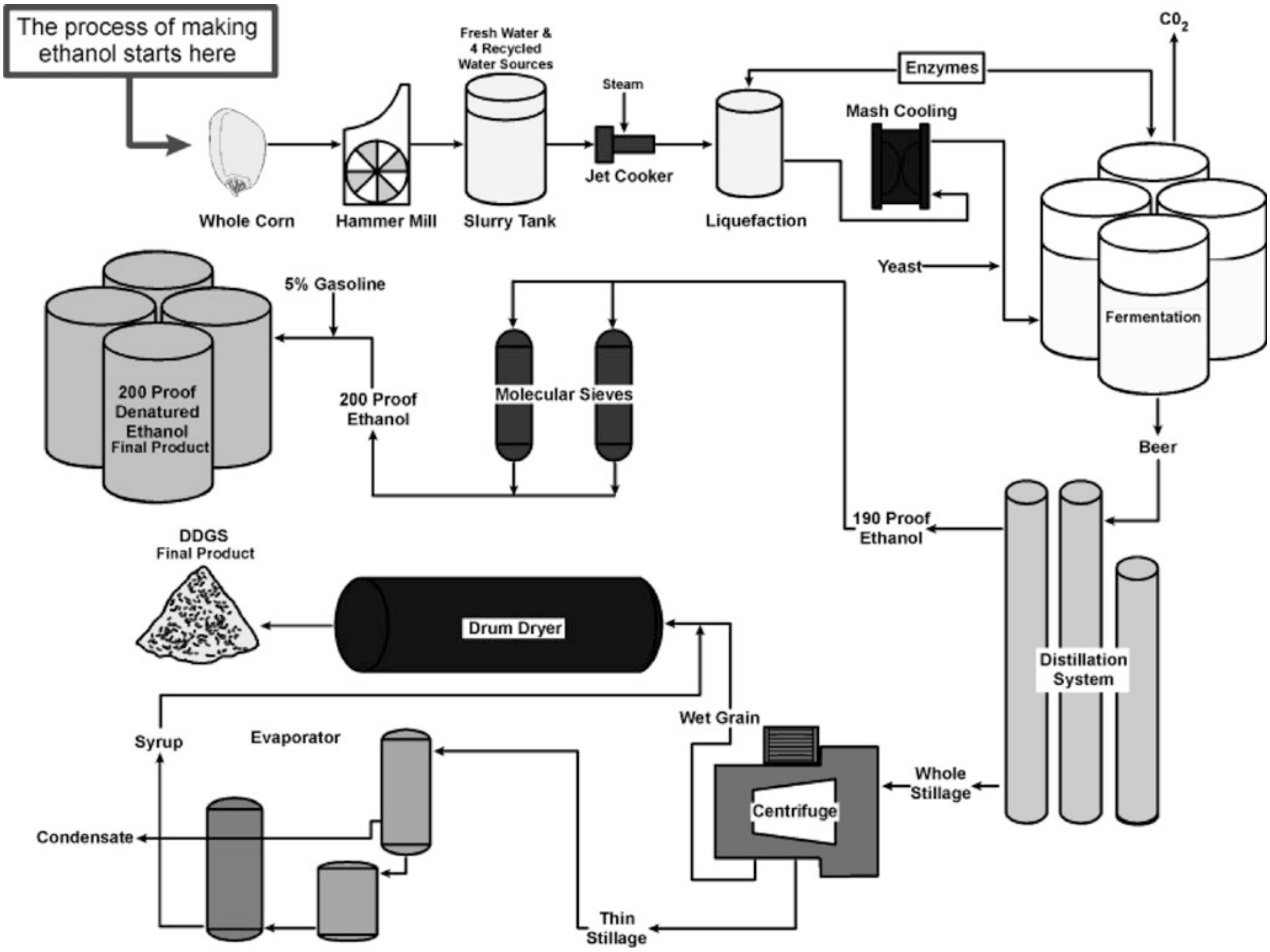


Fig. 31.11 Overview of a corn dry-milling process

to 108°C for 5 min in the presence of a thermostable amylase. The α-amylase from *B. amyloliquefaciens* functions at pH 5.5–7 and 90°C and α-amylase from *B. licheniformis* functions at pH 5.5–9 and temperatures as high as 110°C. In the subsequent saccharification process, an appropriate amount of fungal glucoamylase is added to the thinned starch (30–50% dry substance) with stirring at 55–60°C and pH 4–5 for 48–72 h. This achieves a final DE (dextrose equivalents) of about 97, with about 94% of the dry weight being glucose. The equilibrium concentrations of the saccharides formed by resynthesis limit the maximum degree of hydrolysis obtainable. Because the activity of glucoamylase toward the branching points (the α-1,6-linkages) is low, it may be advantageous to use a debranching enzyme such as pullulanase early in the hydrolysis process.

Dry milling of corn is the dominant means by which corn is converted into bioethanol (Fig. 31.11) and delivers improved process economics relative to corn wet milling. The process begins with the grinding of whole corn into flour in a hammer mill. This flour is slurried, the pH adjusted to around 5.8, and treated at 80–90°C with α-amylase for 30–45 min in order to reduce viscosity. This slurry is then gelatinized and liquefied in a similar manner to the corn wet milling process. Simultaneous saccharification and fermentation (SSF) combines the conversion of maltodextrins to glucose by glucoamylase and fermentation to ethanol. The yeast *Saccharomyces cerevisiae* is typically used, the fermentation period lasting from 48 to 72 h. Following fermentation, the spent mash (referred to as "beer") is distilled to produce 95% (190 proof) ethanol which is further dried and ultimately blended into gasoline.

Enzymes for the processing of granular starch directly to glucose at lower temperatures (<50°C) have been developed, for example, Stargen™ (Genencor) and BPX (Novozymes). The Granular starch hydrolysis (GSH) process eliminates the cooking step and combines the liquefaction, SSF steps, further streamlining the overall economics of corn drymilling.

Amylases are widely used in cleaning products, primarily laundry and automatic dish-washing detergents. The ability

Table 31.6 Glucose isomerase (GI) producing microorganisms[a]

Microorganism	Patent assignee	Yield (GIU)
Arthrobacter sp.	R. J. Reynolds	4,720
NRRL-B-3728		4,440
Streptomyces welmorensis	Standard Brands	4,640/g
Mutant-1		7,540/g
Mutant-2		6,680/g
Mutant-3		6,000/g
Streptomyces olivaceus	Miles	2,560
Mutant NRRL-3916		2,960
Streptomyces olivochromogenes	CPC	4,800–11,440
CPC-4		5,700–9,680
CPC-8		3,960–4,440
Actinoplanes missouriensis	Anheuser-Busch	2,500–35,200
Bacillus coagulans	Novo	Not known

[a] 1 GIU ~ The amount of enzyme which converts 1 mmol glucose to fructose/minute

to hydrolyze starch in food soils greatly boosts the ability of the detergent to remove stains. Many amylases are active at the high pH values and temperatures that some applications demand. In addition, detergent amylases must be resistant to proteolysis and bleaches. Typical dosing levels in the wash are in the 0.1–5 ppm range. In recent years, much effort has gone into developing amylases and other detergent enzymes that operate with improved efficiency at lower temperatures, allowing energy savings.

Glucose Isomerase

Starch degradation using α-amylase and glucoamylase produces corn syrup with glucose concentrations of up to 94% on a dry weight basis. The glucose can then be isomerized to a mixture of glucose and fructose by glucose isomerase (GI), an enzyme present in many microorganisms [60]. The product, known as HFCS, has become the dominant sweetener used in food and beverages, displacing fructose syrups derived from the action of invertase on sucrose. Industrial production of HFCS represents one of the largest biocatalytic processes in current practice and is discussed further in the third section of this chapter.

Many organisms are glucose isomerase producers. Most of them produce xylose isomerase with low specificity and glucose can be an alternative substrate for the enzyme. Table 31.6 lists microorganisms used industrially for glucose isomerase production, along with their reported yields. The desirable application conditions of glucose isomerase are pH 7.0–8.0 and 65°C. Glucose isomerase fermentation typically has a cycle time of about 2 days. The fermentation conditions vary from producer to producer. Many glucose isomerase-producing organisms require xylose and cobalt for full enzyme induction. Xylose is too rare and expensive to be used in commercial fermentation processes and cobalt ions remaining in the spent fermentation medium constitute a serious environmental hazard. However, mutants that do not require xylose and cobalt have been obtained for commercial production of glucose isomerase.

Almost every known glucose isomerase is an intracellular enzyme and can only be extracted from the cells in relatively low concentrations, so it is expensive to use this enzyme as a soluble and “once only” enzyme. Immobilized enzyme/cell technology is the key scientific advance that has made the use of glucose isomerase for HFCS production a commercial success. Improved mutants of GI are still actively sought, particularly those with altered pH optima and increased temperature resistance.

Cellulases

Cellulases (EC 3.2.1) are a large group of enzymes that find use in many industrial applications [61]. Cellulose is a homopolymer of glucose monomers, which are linked by β-1,4 bonds. The difference of the linkages distinguishes cellulose from starch and determines the difference in macroscopic properties of these two natural polymers. Starch functions as energy storage in the form of glucose and is consequently used by humans as a food ingredient, whereas cellulose is a structural polymer, which is used for production of paper and textiles. Cellulases can be separated into three distinct groups: the endo-1,4-β-glucanases (EC 3.1.2.4), which catalyze the hydrolytic cleavage within the cellulose polymers; the exo-1,4-β-glucanases (cellobiohydrolases, EC 3.2.1.91), which attack the polymer from the ends releasing cellobiose; and the β-glucanases (EC 3.2.1.21), which catalyze the hydrolysis of cellobiose into glucose monomers.

Cellulases are found in fungi and bacteria. Of commercial interest are fungal enzymes from *Aspergillus* sp. or *T. reesei* and a few bacterial enzymes. They are either used as a multicomponent, which contain all enzyme types and are found in *T. reesei* (*Hypocrea jecorina*), or as a monocomponent enzyme product, which consists of only one of the three types of enzymes. The multicomponent enzyme preparations can be produced from a selected cellulose overproducing strain of the wild-type organism, whereas the monocomponent cellulases are mainly produced in recombinant production systems.

Currently, the main use of cellulases is in the textiles, paper, and laundry detergent industries. In textiles applications, cellulases are used to change the appearance of the fabric by enzymatic removal of cellulose fibrils at the surface and in the so-called stone-washing process of denim jeans, which gives them the aged look. In the paper industry, cellulases are used as process aids, where they increase the efficiency of the papermaking process by saving energy or allowing higher machine speed. The properties of the paper can also be influenced by cellulose usage. As detergent ingredients, cellulases assist in the cleaning process and can maintain the new look of the washed textiles by removing pills or micro fibrils from the surface of cellulose-based garments.

Hemicellulases

Hemicellulases are a diverse set of enzymes that degrade mixed polysaccharides present in plant derived biomass, known collectively as hemicelluloses [62, 63]. Hemicellulose polymers comprise a variety of different sugars, principally 5-carbon sugars (xylose, arabinose), 6-carbon sugars (mannose, galactose), and uronic acids (galacturonic and glucuronic acids). They also contain ester linkages with organic acids such as acetic and ferulic acids, or with alcohols, for example, methyl esters of galacturonic acid. Accordingly, the hemicellulases span a range of different enzyme activities capable of cleaving both glycosidic and ester bonds. Hemicellulases are widely distributed in the microbial world and are found in bacteria, fungi, and yeasts from many environments. Some more common hemicellulases are listed below. A more comprehensive account on hemicellulases can be found in Chap. 33 on Biomass Conversion in this volume and the references cited within.

- α-L-Arabinofuranosidase (EC 3.2.1.55)
- Endo-α-1,5-arabinase (EC 3.2.1.99)
- α-Galactosidase (EC 3.2.1.22)
- Endo-galactanase (3.2.1.89)
- α-Glucuronidase (EC 3.2.1.139)
- Endo-β-1,4-mannanases (EC 3.2.1.78)
- Exo-β-1,4-mannanases (EC 3.2.1.25)
- Endo-β-1,4-xylanases (EC 3.2.1.8)
- Exo-β-1,4-xylanases (EC 3.2.1.37)
- Acetyl xylan esterase (EC 3.1.1.72)
- Ferulic acid esterase (EC 3.1.1.73)

Hemicelluloses, like xylan, arabinogalactan, or glucomannan, are part of many natural materials from plant origin and have a variety of uses. Consequently, xylanases, glucomannanases, and mannanases find use in many technical applications, for example, as detergent additives, for flour processing in baking, as animal feed additives, as well as in the textiles and pulp and paper industry.

The pectins are a related but distinct class of plant polysaccharide composed of poly-β-1,4-galactouronic acid chains with varying degrees of methyl ester functionality. Four different enzyme classes help degrade pectin and are referred to collectively as "pectinases" [64]. They are pectate lyase (4.2.2.2), pectin lyase (EC 4.2.2.10), polygalacturonase (EC 3.2.1.15), and pectin methylesterases (EC 3.1.1.11). The predominant use of pectinases is for the clarification of fruit juices by hydrolysis of haze-forming pectins. Pectate lyase is also used for cotton scouring by modifying pectin within the cotton structure and enhancing the subsequent washing steps. Another textiles application for which hemicellulases are used is for desizing of fabric.

Lignocellulose Degradation

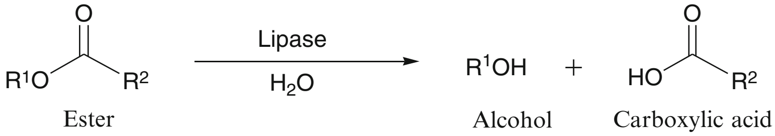
The use of cellulase/hemicellulase blends for conversion of lignocellulosic biomass into fermentable sugars for the production of ethanol and other chemicals promises to greatly expand the market for these enzymes [65]. The quest for economic processes for the utilization of cellulosic biomass has faced many challenges, for example, the cost of and amount of the enzymes required to carry out the process. As much as 100-fold more native cellulase protein (as compared to amylase protein for breakdown of starch) is required for conversion of pretreated substrate (e.g., corn stover) to fermentable sugars. In addition, chemical or thermal pretreatment of cellulosic biomass is needed in order to enable subsequent enzymatic hydrolysis of polysaccharides. Despite these hurdles, much progress has been made in recent years through intensive research programs conducted by government, industry, and academic institutions and have led to the construction of several pilot and demonstration plants, primarily for ethanol production. These efforts have led to identification of new enzymes for biomass degradation, enzymes with greater thermostability, and the optimization of the ratios of the components in enzyme blends.

One strategy to minimize ethanol production cost is to run simultaneous SSF, which utilizes ethanologens engineered to operate in high-temperature environments. Ideally, the fermentation organism has ability to utilize C5 sugars derived from the hemicellulose component, with acceptable productivities in the presence of numerous byproducts of the biomass pretreatment process, leading to lower overall production costs [66]. Several enzyme products designed specifically for biomass processing are now commercially available, for example, the Accellerase™ product line from Genencor.

Lipases and Esterases

Lipases (EC 3.1.1.3) and esterases (EC 3.1.1.1) are enzymes that catalyze the hydrolytic cleavage of ester bonds (Fig. 31.12), but differ in their substrate spectrum. Lipases have the special capability to catalyze the hydrolysis of water-insoluble substrates such as fats and lipids. Like many other enzyme-

Fig. 31.12 Lipases and esterases catalyze the hydrolysis of ester bonds



catalyzed reactions, the ester hydrolysis is a reversible process, which allows using lipases and other esterases for the synthesis of esters. The use of lipases and esterases as biocatalysts is described in the Industrial Biocatalysis section of this chapter.

True lipases show the interfacial activation phenomenon in their catalytic activity pattern. At low concentration of water-insoluble substrates, lipases are almost inactive, and the hydrolytic activity does not increase linearly. At a certain substrate concentration, however, the hydrolytic activity of lipases increases rapidly and the lipase kinetics resembles "normal" enzyme kinetics. This boost in activity is related to the formation of water-insoluble substrate aggregates such as micelles or another second phase. Only when this second phase is present, do lipases become fully active. This interfacial activation is caused by a large conformational change in the 3D structure of the lipases. In their water-soluble form, the active site is covered by a lid, which prevents the substrates from reaching it. At the lipid/water interface, the lid is opened and the active site is accessible to the substrates. In addition, the now accessible area is mainly hydrophobic, which gives the open-form lipase the shape and behavior of conventional surfactant molecules with a hydrophilic and a hydrophobic moiety in one single molecule.

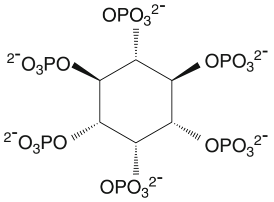
Lipases are found in a wide variety of bacteria, yeasts, and filamentous fungi. Due to their unique properties, lipases are used in all applications where lipids are modified or need to be removed [67]. In the detergent industry, lipases are added to laundry detergents for the removal of fat-containing stains. In the paper industry, lipases remove fatty pitches from paper in the manufacturing process. These pitches stem from remaining resin of the wood chips.

In the food industry, lipases are used in lipid modification processes. In these processes, the texture, digestibility, or physical properties of natural lipids are modified by lipase-catalyzed transesterification reactions with lipids other than the original fatty acids. In the baking industry, lipases are used to influence the quality of bread through modification of the wheat flour lipids. Finally, lipases are used for flavor enhancement of cheese in the dairy industry. The phospholipases (EC 3.1.1.4) related class of enzymes which act upon phospholipids such as lecithin [68]. The major use of these enzymes is for degumming of vegetable oils, for egg processing, and in the baking industry.

A new industrial application for lipases has been developed in the field of renewable energy. Lipases are used for methylation of plant oils such as rapeseed oil. The resulting methylated oils (biodiesel) are used as a replacement or supplement for fossil fuel-derived diesel. As of today, the costs of the biocatalyst, however, prevent its commercial application and chemical processes such as saponification of fats with sodium methoxide is currently preferred.

Phytase

Phytases are a class of phosphate ester hydrolase (a phosphatase) that act upon phytic acid (phytate, myo-inositol-1,2,3,4,5,6-hexakisphosphate), the dominant form of phosphate in plant feedstuffs, including cereal grains and oil seeds. The use of phytases as additives in feed for monogastric animals has expanded greatly in recent years. They increase the bioavailability of phosphate, thus enhancing animal growth while reducing the amount of phosphate that would otherwise end up in the environment [69].



Phytate (myo-inositol-1,2,3,4,5,6-hexakisphosphate)

Most commonly used in swine and poultry feed, phytases also increase the bioavailability of other minerals (calcium, magnesium, iron, and zinc) and nutrients (amino acids, proteins) by hydrolysis of complexes with phytate (phytin) that tend to sequester these nutrients.

Phytases are found in many organisms where their functions include phosphate metabolism and the modulation of inositol phosphate regulatory metabolites. Phytases are characterized according to the position at which they initially remove phosphate from phytic acid. Plant phytases (EC 3.1.3.26) remove phosphate from the 4- and 6-positions of phytic acid, whereas those from microbial sources (EC 3.1.3.8) such as *Aspergillus* species tend to remove phosphate from the 3-position. There are exceptions to these rules, for example, phytases from *E. coli*, *Buttiauxella* and *Peniophora* sp. act at the 6-position of phytate. Both acid- and alkaline-phytases have been described. Those from microbial sources usually belong in the former category with pH optima in the 4.5–6.5 range. Phytases expressed in eukaryotic organisms are typically glycosylated, a feature than has been found to confer stability by reducing the protein aggregation at elevated temperatures. There are four different types of phytase, listed in Table 31.7.

Table 31.7 The four known phytase families

Phytase family	Description	Example
Histidine acid phytases (HAP)	• Largest class of phytase • Concensus active site motif is Arg–His–Gly–Arg–X-Pro • Two-step catalytic mechanism	*Aspergillus niger* HAP *Escherichia coli* HAP
Beta-propeller phytases (BPP)	• Structure resembles a six-bladed propeller • Degrades phytate in soil and water	*Bacillus subtilis* phytase C *Pedobacter nyackensis* phytase
Purple acid phytases (PAP)	• Metalloenzyme first isolated from the cotyledons of germinating soybeans • Also present in other cereal grains	Soybean (*Glycine max*) PAP Barley (*H. vulgare*) PAP
Protein tyrosine phosphatase-like phytase (PTP)	• Isolated from microorganisms present in ruminant guts • Concensus sequence of His–Cys–(X)5–Arg	*S. rumanatium* PTP

Bacterial and fungal phytases are most readily produced at scale and possess the activity and stabilities required for an industrial product. Nonetheless, most wild-type phytases have temperature optima in the 40–60°C range and considerable activity is lost at the elevated temperatures (>90°C) attained in the steam pelleting process used to manufacture feed pellets. Ideally, a phytase product must retain at least 80% of the initial activity upon transient exposure to temperatures of 90°C or greater. For this reason, both protective enzyme granulation technologies and phytases engineered for greater thermostability have been developed to minimize loss of activity during the feed pelleting process. Commercial Phytase products include Phyzyme® XP (Danisco), Ronozyme® (DSM), and Natuphos™ (BASF).

Phytases have also found application in the starch processing industry where they are used to reduce the levels of phytate derived from corn and other feedstocks. Enzyme blends containing both α-amylase and phytase are commercially available. Reduction of phytate in the starch liquefaction process has numerous benefits including;

- Ethanol yield improvement
- Removal of phytate-mediated inhibition of α-amylases
- Viscosity reduction during liquefaction
- Release of inositol, a yeast nutrient
- Production of low-phytate, high phosphate DDGS for animal feed

Overall the market for phytase is poised for further growth, particularly in Asia and other developing regions, driven by demand for meat and need to minimize the associated effects on the environment.

Oxidoreductases

Oxidoreductases (EC family 1) are enzymes that add or remove electrons from a substrate. Cofactors are usually required to assist electron transfer and a stoichiometric electron donor or acceptor is always needed. Applications for oxidoreductases have been limited in comparison to the hydrolases, but have grown substantially in recent years [70]. Oxidases using oxygen as a terminal electron acceptor have found the greatest application, whereas those that require an electron source are still restricted in their practical application. Oxidoreductases of particular prominence are discussed below.

Glucose oxidase (EC 1.1.3.4) converts D-glucose into D-gluconolactone utilizing molecular oxygen (O_2) as the ultimate electron acceptor and producing hydrogen peroxide in the process [71]. This enzyme is used for scavenging traces of oxygen in food products, generation of hydrogen peroxide as a preservative, and for the detection of glucose in body fluids. While the later application does not consume a large amount of enzyme by weight, it forms the basis of modern glucose meters used by diabetics and others to monitor blood glucose levels. Glucose oxidase is commonly used in conjunction with catalase (EC 1.11.1.6), an oxidoreductase that converts hydrogen peroxide into oxygen and water [72]. Catalase is also used to decompose excess hydrogen peroxide in textiles bleaching applications.

Laccases (EC 1.10.3.2) are a class of copper-containing oxidases that transfer electrons from phenolic substrates to oxygen, producing water in the process [73]. The phenolic substrates are also known as mediators and can remove electrons from a wide array of substrates including lignin and dyes. Laccases occur in plants, bacteria, and fungi and have a characteristic blue color in their pure form. Commercial applications include bleaching of indigo in denim manufacture, decolorization of dyes in wastewater, and for degradation of lignocellulosic biomass [74].

Emerging Industrial Enzymes

New applications for established classes of industrial enzymes such as proteases and amylases are developed every year [6]. By contrast, products based upon new classes of enzyme are less common, although there is great potential for new enzymes that enable chemistries which cannot be performed under mild conditions. Some examples are discussed below.

Organophosphate hydrolase (OPH) (EC 3.1.8.1) and organophosphorous acid anhydrolases (OPAA) (EC

3.1.8.2) are enzymes that detoxify organophosphorous compounds including pesticides (e.g. paraxon, parathion, malathion) and chemical warfare agents (e.g., sarin, soman, tabun) [75]. These enzymes are being assessed for use by primary responders to counter chemical agents, although even larger potential lies in dealing with organophosphate pesticide contamination in agriculture.

Asparaginase (EC 3.5.1.1) is an enzyme that converts the amino acid asparagine to aspartic acid and ammonia by hydrolysis of the side-chain amide functionality. This class of enzyme is used to reduce the formation of acrylamide, a neurotoxic compound, in starchy foods such as French fries, by treatment of the foodstuff prior to cooking [76]. Products containing asparaginase are sold under the tradenames Acrylaway® and PreventASe™.

A bacterial arylesterase capable of peracid formation under aqueous conditions was developed by Genencor for biobleaching applications and is currently sold under the tradename PrimaGreen® EcoWhite. This enzyme has a unique active site configuration which disfavors hydrolysis of esters and readily performs both perhydrolysis and acyltransfer reactions [77]. The enzyme is formulated with propylene glycol diacetate and catalyzes the formation of peracetic acid in the presence of dilute aqueous hydrogen peroxide.

Section III: Industrial Biocatalysis

The use of enzymes for the synthesis of bulk and fine chemicals represents a somewhat specialized application relative to their wider industrial uses. Nonetheless, biocatalysis is becoming increasingly relevant within the chemical industry for the production of a wide range of materials (see Table 31.8) [5, 16–18]. Broadly defined, a biocatalytic process involves the acceleration of a chemical reaction by a biologically derived catalyst. In practice, the biocatalysts concerned are invariably enzymes and are used in a variety of forms. These include whole cell preparations, crude protein extracts, enzyme mixtures, and highly purified enzymes—both soluble and immobilized. The products of biocatalytic processes are as varied, ranging from the synthesis of fine chemicals and pharmaceuticals to polymer precursors, flavor, fragrances, foodstuffs, and fuels. The common theme is that the use of a biocatalyst has either lowered the cost of production or facilitated the development of a new product.

Biocatalysts often enable one to perform chemical transformations with greater selectivity or under milder conditions than might be achieved through conventional chemical means [78]. Large reductions in waste streams have been achieved in many instances. The term “bioconversion” has also been used in order to distinguish biocatalytic processes involving the transformation of a defined starting material into specific products using either isolated enzymes or resting cells. Despite the widespread recognition of the potential advantages of biocatalysis, the number of large-scale industrial processes is still relatively limited. A summary of the products of these processes is listed in Table 31.9.

The use of biocatalysis for the production of chemicals started to receive serious interest in the 1960s with the development of immobilized aminoacylases for the production of chirally pure amino acids by Tanabe Seigaku of Japan, as well as the application of penicillin acylase for the production of 6-aminopenicillanic acid (6-APA), a key intermediate in the production of semi-synthetic antibiotics. The production of HFCS began in 1968 with implementation of a batch process using whole cells originally developed by Takasaki et al. [79]. The industrial production of HFCS remains as the largest bioconversion worldwide. These applications coincided with the growing use of proteases for cleaning and significant improvements in the production

Table 31.8 Comparison of biocatalysis to chemical synthesis for chemical production

• Advantages	Disadvantages
• Mild reaction conditions	• Low volumetric productivity
• Aqueous medium	• Substrate solubility issues
• Highly stereo-, regio-, and chemoselective	• Unwanted reactions with impure preparations
• Unique and varied chemistry	• Restricted set of substrates
• Environmentally friendly	• Costly recovery of products
• High catalytic efficiency	• Poor operational stability
• Ability to fine tune catalysts	• Long development cycles

Table 31.9 Examples of products made through the application of biocatalysis

Scale (t/year)	10	$>10^2$	$>10^3$	$>10^4$	$>10^5$
Product	Glycidylbutyrate Butyl glucosides Styrene oxides Oligosaccharides	D-Pantothenic acid L-Malic acid L-Methionine L-Valine *R*-mandelic acid L-Carnitine	Aspartame L-Aspartate Cyclodextrins (*S*)-CPA	6-APA Cocoa butter Acrylamide	HFCS

(S)-CPA (*S*)-chloropropionic; *6-APA* 6-aminopenicillanic acid; *HFCS* high fructose corn syrup

Table 31.10 Enzyme classes used in biocatalysis

Enzyme	Reactions catalyzed
Lipases	Ester formation and hydrolysis
Proteases	Ester and amide hydrolysis, peptide synthesis
Esterases	Ester hydrolysis
Aminoacylases	Amide formation and hydrolysis
Acyltransferases	O- and N-acylation
Epoxide hydrolases	Hydrolysis of epoxides
Phosphoesterases	Hydrolysis of phosphate esters
Glycosidases	Glycoside hydrolysis/formation
Oxygenases	Hydroxylation, dihydroxylation, epoxidation
Glycosyltransferases	Glycosyl transfer
Nitrilases	Nitrile hydrolysis
Oxynitrilases	Cyanohydrin formation
Oxidoreductases	Reduction of aldehydes, ketones, and enolates
Aldolases	Aldol reactions (C–C bond formation)
Isomerases	Isomerization of carbohydrates, racemization, and epimerization

and isolation of industrial enzymes. The advent of recombinant DNA technology in the 1970s was followed by the development of the powerful PCR reaction for DNA amplification in 1984 by Kary Mullis while working at Cetus Corporation. Aside from revolutionizing biomedicine, these and other technologies have led to the ability to alter and improve many wild-type enzymes to the point that new bioconversions are now economically feasible.

The majority of industrial biocatalytic processes involve the use of hydrolytic enzymes including proteases, transaminases, glycosidases, aminoacylases, and lipases as well as several additional enzyme classes (Table 31.10) [80]. In some instances, hydrolytic enzymes can be used to drive the reverse reaction as in the formation of ester bonds through the use of lipases. In contrast, oxidases and other enzymes requiring movement of electrons commonly rely on the supply of a cofactor that may need to be recycled for practical application. These processes are typically limited to those involving respiring cells capable of regenerating such cofactors, or through the supply of a stoichiometric electron donor/acceptor, for example, hydrogen peroxide or oxygen. An emerging application of redox enzymes is in diagnostics and other sensitive analytical devices where direct electronic detection of a chemical reaction is possible.

One of the greatest hurdles for the application of biocatalysis is the need to operate processes under conditions that can differ dramatically from those in which the enzymes evolved. Many techniques are used in order to preserve catalytic activity and minimize the costs associated with the biocatalyst. In cases where the cost of the biocatalyst is a concern, an enzyme might be immobilized and used in a packed column or a fluidized bed reactor so as to enable reuse. Here also the enzyme must be stable for extended periods and may even be used under nonaqueous conditions and elevated temperatures. Recombinant technology has revolutionized the applications of biocatalysts through the ability to modify existing proteins. When combined with high-throughput screening and other analytical methods, it is now possible to redesign and improve wild-type enzymes such that they can be used in areas not previously possible. Coupled with these advances are those in the large-scale production and recovery of industrial enzymes, in addition to improvements in formulation and stabilization of enzymes.

The following sections focus on the elements of industrial biocatalysis, including the definition of a biocatalytic process, rendering, biocatalysts into immobilized forms, and performing bioconversions using whole cells. Key issues relating to the design of bioreactors are highlighted and several larger-scale bioconversions are discussed in further detail, followed by some emerging trends in biocatalysis.

Biocatalytic Processes

The nature of a biocatalyst often defines its utility for industrial bioconversions and determines the conditions under which it can be used. For example, the biocatalyst must be stable enough for practical application and of sufficiently low cost so as not to affect the overall economics of the process. Biocatalysts come in many forms including highly purified enzymes, enzyme mixtures, and whole cells. The form used is often related to cost and performance considerations. Although a crude enzyme preparation might be relatively cheap, the presence of additional enzyme activities might affect the purity and yield of the product. Similarly, the need for cofactor recycling might weigh against the use of a purified enzyme in favor of a respiring whole cell. Another important factor relates to the exquisite selectivity of many enzymes. This has both advantages and disadvantages in that a particular enzyme might be too narrow in its substrate preference to allow the use of unnatural substrates, thus limiting the potential application of that enzyme. Interestingly, some enzymes demonstrate broad substrate selectivity while retaining a high degree of stereoselectivity. A case in point is the lipase B isolated from *Candida antarctica* [81]. Despite the fact that this enzyme was isolated from a psychrophilic organism (cold-adapted), it demonstrates a high degree of thermostability, being functional at over 80°C, especially in an immobilized form. Another enzyme displaying broad specificity is rabbit muscle aldolase, useful for the formation of C–C bonds [82]. Selection and development of the right biocatalyst for a given transformation is often the key for successful commercialization of a biocatalytic conversion.

Immobilized Enzymes

There has been considerable effort directed toward the immobilization of both enzymes and whole cells in a wide array of formats [83, 84]. Initial attempts to immobilize enzymes on naturally derived supports such as charcoal were conducted early in the twentieth century and eventually led to the development of more robust biocatalysts immobilized on synthetic resins by the mid-1950s. Immobilization often confers a number of advantages relative to the free biocatalyst including ease of removal from the process stream, potential for reuse, improvements in stability, favorable alterations in kinetic parameters, suitability for continuous production, and in some cases the ability to operate in organic solvents. The focus of this section is on the immobilization of enzymes; however, many of the same principles apply to whole cells, the primary difference being the fact that immobilized cells are often less stable than individual enzymes and may contain additional undesired enzyme activities.

Immobilization needs to improve the performance of an enzyme enough to offset the costs associated with the procedure. Such gains can be measured in terms of improvement in the total amount of product produced per unit of enzyme, an increase in the rates and volumetric efficiency of a bioconversion, improvement in the ease of removing the biocatalyst, or through the enabling of new applications for a given enzyme. It is often the case, however, that the cheapest immobilization methods suffer from a number of drawbacks including lack of both enzymatic and mechanical stability, leaching of the enzyme, fouling of the support, and limited enzymatic activity. The many methods for producing immobilized enzymes can be divided into these subcategories:

1. Adsorption to a matrix such as carbon, chitin, diatomaceous earth, and ion exchange resins
2. Crosslinking enzyme crystals and whole cells with gluteraldehyde and other agents
3. Gel entrapment in silica sol-gels, alginate, and protein matrices
4. Covalent attachment to resins and other carriers
5. Encapsulation within a membrane or liposome

Many factors influence the catalytic efficiency and kinetics of an immobilized enzyme. The immobilization process itself can lead to loss of enzymatic activity through incomplete capture of the enzyme or through denaturation of the enzyme protein. Such denaturation may arise from the action of reagents used in the immobilization process, for example, the use of carbodiimides during covalent immobilization methods, or through partial unfolding of the protein upon contact with hydrophobic surfaces such as those encountered in the adsorption of enzymes onto polymeric resins. Immobilized enzymes often demonstrate dramatic improvements in stability over free enzymes, especially in organic solvents or at elevated temperatures. This is particularly the case where enzymes have been attached to a solid support through multipoint covalent attachment or entrapped within a rigid matrix. In such instances, the potential for loss of activity through unfolding of the enzyme is reduced.

Kinetics of immobilized enzymes. Another major factor in the performance of immobilized enzymes is the effect of the matrix on mass transport of substrates and products [85]. Hindered access to the active site of an immobilized enzyme can affect the kinetic parameters in several ways. The effective concentration of substrates and products is also affected by the chemistry of the matrix, especially with regard to the respective partition coefficients between the bulk solution and the matrix. Immobilization of an enzyme typically reduces the rate of diffusion of a substrate to the active site of an enzyme and increases the apparent K_m while reducing V_{max}. The nature of the mass transfer effect depends on the fashion in which the enzyme is immobilized. Enzymes immobilized on the surface of a carrier will experience external mass transfer limitations between the bulk solution and the surface, whereas those entrapped within a porous matrix are also affected by internal mass transfer limitations due to the reduction in the rate of diffusion of substrate and products through the matrix.

The effect of diffusional limitations is given by the Damköhler number D_a, defined as the maximal rate of reaction divided by the maximal rate of diffusion:

$$D_a = V_{max}/k_L[S_b] \tag{31.6}$$

where $D_a \gg 1$ diffusion limits the observed rate, whereas when $D_a \ll 1$ the reaction rate is limiting. The effectiveness factor η is commonly used to assess the effect of immobilization on a given enzyme and is defined as the ratio of the rate with diffusional limitation vs. the rate without diffusional limitation. The relationship is a function of bulk substrate concentration β, defined as $[S]/K_M$ in addition to the ratio of the bulk substrate concentration $[S_R]$ to that within the matrix $[S_r]$. Thus, the rate of reaction within a matrix r_s is given by the modified Michaelis–Menten equation:

$$Rs = \eta(V_{max}[S]/K_m + [S]) \tag{31.7}$$

Current methods for large-scale immobilization. Although there are numerous methods for enzyme immobilization described in the scientific literature, relatively few methods have been successfully applied on an industrial scale. By volume the largest process is the immobilization of glucose isomerase for the production of HFCS. Several methods have been used, summarized below:

- Adsorption to polyethylenimine (PEI) treated alumina, followed by gluteraldehyde cross-linking (UOP, 1981)

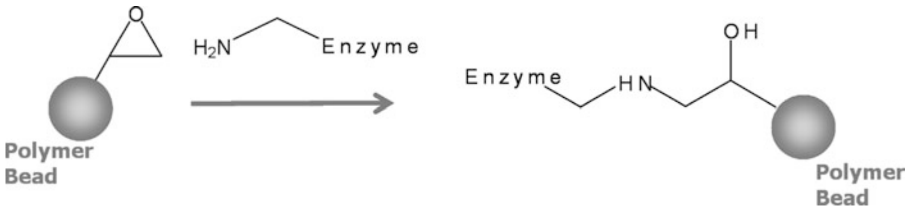


Fig. 31.13 Covalent attachment of enzymes to oxirane-functional polymeric beads (Courtesy of M. Elizabeth Miller and Rohm and Haas)

- Adsorption onto anion exchange resins (e.g., DEAE-cellulose/TiO_2 based resin)
- Cross-linking of lysed cells with gluteraldehyde followed by extrusion
- Incorporation into clay/PEI/gluteraldehyde cross-linked matrix, mechanical formation of granules

Over 500 t of immobilized GI are produced per annum, the majority through the incorporation of GI into a porous clay-based composite. The largest producer is Genencor. The process begins with a whole-cell lysate derived from GI fermentation to which is added PEI, bentonite clay, and diatomaceous earth, followed by cross-linking with gluteraldehyde. The resulting flocculant solid is filtered and pressed to remove excess water. Particles are fashioned mechanically by extrusion and dried in a fluidized bed dryer. The resulting immobilized GI is extremely stable, with a half-life of over 1 year when used in a packed bed reactor at temperatures of over 60°C. One kilogram of immobilized GI is sufficient to produce 20,000 kg of HFCS.

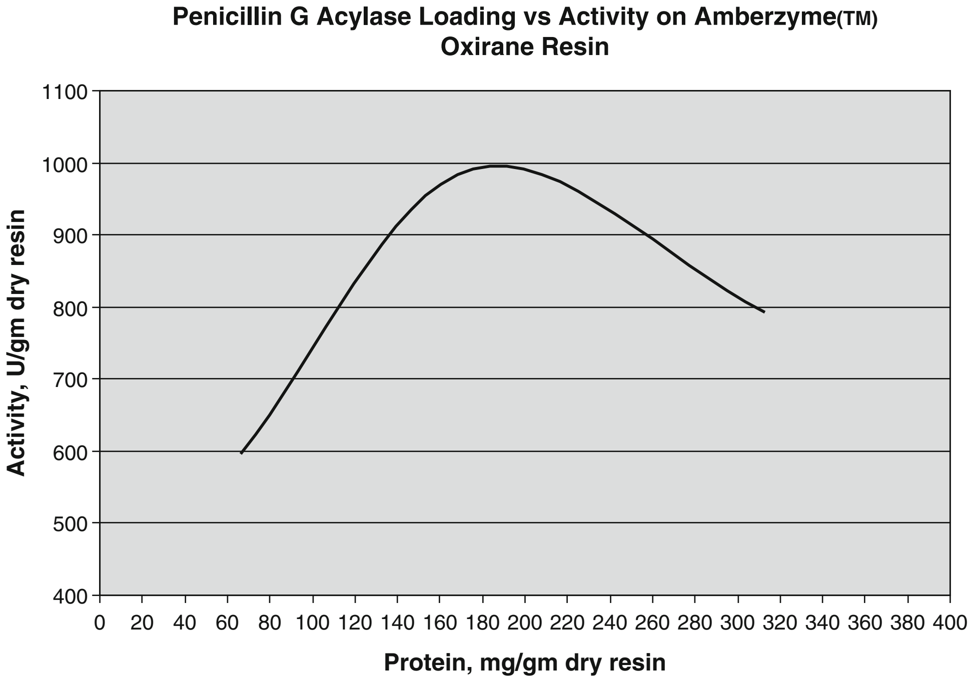
Another method for large-scale immobilization of lipases involves the formation of silica granulates, a technology developed by Novozymes A/S (Denmark) with the aim of producing a low-cost immobilized lipase for transesterification of food oils [86]. This process involves initial coating of a silicate powder with a crude lipase solution, followed by addition of a binding agent and mechanical granulation. The resulting granulates are typically 100–300 μm in diameter and are best suited to nonaqueous biocatalysis. A process for the transesterification of food oils based on a silica granulate containing *Thermomyces lanuginosa* lipase, Lipozyme TL-IM, was developed in a joint venture between Novozymes and ADM and has produced transesterified oils in the United States since 2002. The process itself uses a portable packed bed reactor operating at temperatures well below that required for chemical transesterification. Another advantage of the bioprocess is the fact that that product does not need to be purified by extraction with water as is necessary in the traditional chemical process.

A considerable number of industrial bioconversions utilize covalently immobilized biocatalysts. Examples include Penicillin acylases V and G, aminoacylases, and aspartase. In some cases the biocatalyst is immobilized through cross-linking and in others the catalyst is captured by a reactive resin [87]. Covalent immobilization often leads to extremely stable catalysts with high potential for reuse and exhibiting minimal leaching of enzyme during operating conditions. A variety of both organic and inorganic supports has been used, typically modified with a reactive functional group. The oxirane-based acrylic resins Eupergit C® (Röhm GmbH, Germany) and more recently Amberzyme™ (Rohm and Haas, USA) have proven to be versatile carriers for many enzymes. Attachment of an enzyme to an oxirane resin involves agitating a solution of the enzyme with the resin at approximately pH 8 for 12–24 h, followed by washing to remove non-covalently bound material (Fig. 31.13). The absence of amine-containing buffers (e.g., TRIS) or impurities is an important requirement for achieving high enzyme loading on the resin. The total activity of enzyme-functional beads is proportional to the amount of enzyme bound up to a certain point, past which the specific activity falls, most likely due to mass transport limitations. This effect is depicted in Fig. 31.14.

Adsorption of enzymes to various polymeric resins is a straightforward means for immobilization. Zwitterionic molecules such as proteins can bind to both anionic and cationic ion exchange resins. Hydrophobic macroporous resins are also useful for immobilizing many enzymes, particularly lipases. For example, an immobilized form of *Candida antarctica* lipase B (CAL-B) on acrylic resin has been sold for many years under the name, Novozyme® 435 (N435). The enzyme is produced in a modified *Aspergillus* organism by submerged fermentation and is subsequently adsorbed onto a macroporous acrylic resin. Once immobilized, the enzyme is very robust and ideally suited to reactions in organic media where it has been widely used by the research community. Several pilot-scale processes have been developed based upon this catalyst including the production of sugar-based surfactants through the solvent-free esterification of ethyl glucoside with long-chain fatty acids. N435 has also been used for the synthesis of chiral alcohols, amines, and carboxylic acids. The use of this biocatalyst for the synthesis of biodegradable polyesters was developed by Baxenden (United Kingdom) and scaled to 2 t per annum. The cost and limited availability of this catalyst is one of the limiting factors in the wider use of CAL-B in polymer synthesis.

The last decade has seen great progress in materials science, an example being the development of practical methods for the synthesis of a great variety of mesoporous

Fig. 31.14 Relationship between Penicillin G Acylase loading level and on-bead enzyme activity using Amberzyme™ Oxirane resin: loadings at higher challenge levels (Courtesy of M. Elizabeth Miller and Rohm and Haas)



solids and nanoparticles. A number of researchers have used these materials as substrates for enzyme immobilization [88]. Advantages of nanomaterials for enzyme immobilization include very high surface areas, narrow size distributions, and the ability to tailor surface chemistries for optimal enzyme activity. Nanomaterials have also been designed to facilitate electron transport between electrodes and redox active proteins.

Enzymes can also be immobilized using techniques that mimic natural processes for the biosynthesis of inorganic materials [89]. One strategy employs a templating molecule capable of precipitating an inorganic solid like silica (SiO_2) from a soluble precursor—sodium silicate being one example. An enzyme can be premixed with the templating molecule, or even covalently attached to a templating agent. Addition of the enzyme/template solution to an inorganic precursor results in the rapid formation of a colloidal coprecipitate consisting of the enzyme, the template, and the inorganic matrix (Fig. 31.15). A portion of the enzyme becomes entrapped within the composite matrix, often retaining high catalytic activity. The solid can be filtered and dried, or spray-dried directly. An application of this technology has been to formulate improved anti-fouling paints for ship hulls as an alternative to toxic tin-based coatings [90].

Whole-Cell Biocatalysis

Whole-cell biocatalysis is a productive and practical style of conducting biocatalytic reactions [91]. Such reactions, as implied by the term, are done with structurally intact cells. Usually, viable respiring cells are used, but not exclusively. There are many reasons why a whole-cell reaction might be preferred to a cell-free reaction using crude or purified enzymes. Often, whole-cell biotransformations use enzymes, which are not normally excreted into the growth medium. Quite often these enzymes, which are normally intracellular, are unstable outside of the cell and quickly lose activity making them unsuitable for reactions without further stabilization work. Thus, the specific productivity or total turnovers per mole of catalyst may be severely compromised compared to that seen when used in the whole cell due to inactivation of the enzyme in the reaction conditions.

Whole-cell biocatalytic reactions are most often used when the biotransformation to be conducted requires the input of energy. In biological systems this usually takes the form of reduced pyrimidine nucleotides or ATP, but can be many of a number of reduced cofactors or modified reaction components. Using the whole cell allows the technologist to take advantage of the intact, preformed cellular machinery to efficiently provide the required cofactors or components. In order to provide the energy to catalyze these reactions, a source of reducing power is usually required. The co-oxidation of an oxidizable substrate such as glucose or ethanol can provide this energy. The partitioning of the enzymes inside the cells from the bulk medium of the reaction also can be advantageous to let the cellular machinery keep the reaction medium around the enzymes constant and allow for the accumulation, removal, or in other respects processing of the reaction products (see Fig. 31.16). Many techniques exist for the in situ removal of reaction products or side products using whole cells [92]. This task is often

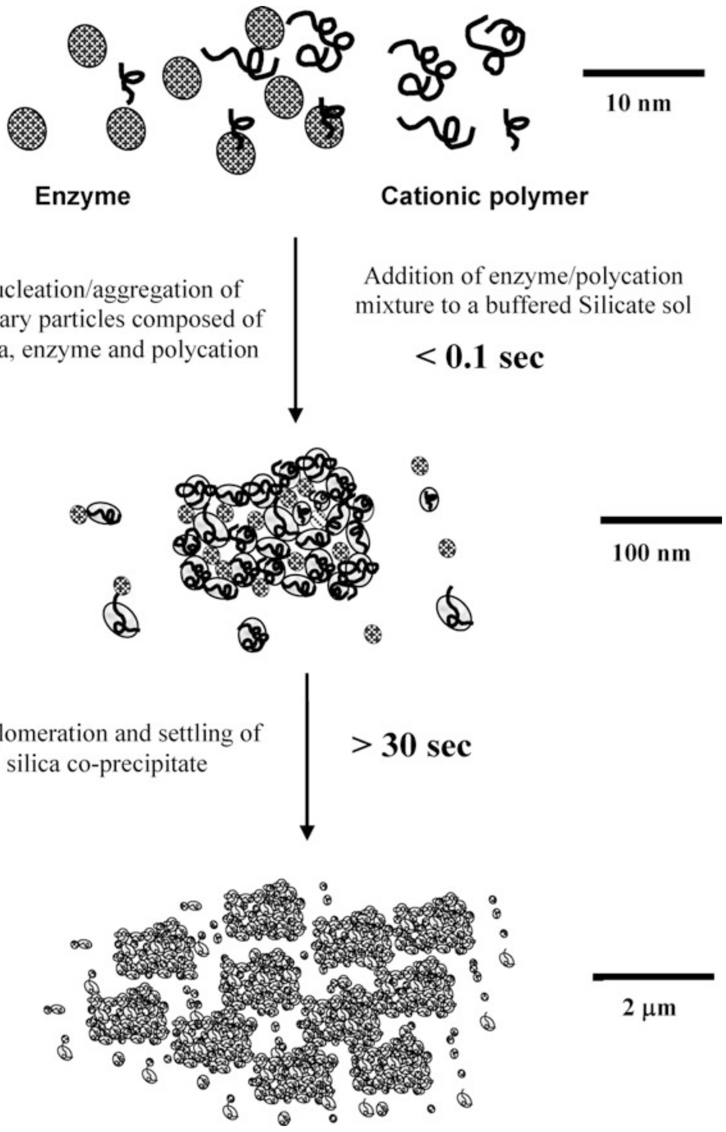


Fig. 31.15 A template-driven silicate coprecipitation process results in the formation of a active enzyme/template/silica coprecipitate

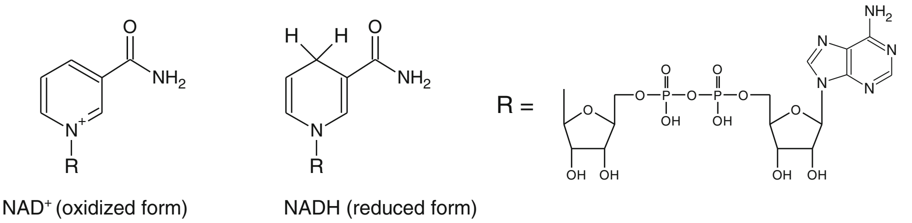


Fig. 31.16 Oxidized and reduced forms of nicotinamide adenine dinucleotide

more difficult with cell-free reactions unless the catalyst is immobilized or in some other way partitioned from the bulk reaction medium.

Another major factor when considering whole-cell vs. cell-free reactions is the overall reaction kinetics. Some enzymatic reactions utilize a complex multicomponent enzyme system. Reconstitution of the crude or purified enzyme components are not usually as effective in vitro as they are when they remain in the intracellular milieu. Whole cells have often been called “little bags of enzymes.” Although this is an oversimplification, it is a useful concept to consider. Whole cells sequester the enzyme components in a small but concentrated form, which is usually optimal for high efficiency. Whole cells also contain co-factors, including the systems that recycle them and control pH and ionic strength. Altogether these factors combine to make whole cells a very useful form for the presentation and use of sensitive enzyme catalysts.

A practical point of using whole-cell biocatalysts is their inherent ease of preparation, use, and removal. There is a variety of growth considerations, as with any fermentation, but the major ones with respect to the use in biotransformations surround the induction of the enzyme(s) of interest and the repression of enzymes which might compete with the desired process or degrade the desired catalyst. Generally, a medium can be used in which the cells can be used directly after growth for the biotransformation. Some consideration needs to be given to the extraction of the product with respect to the growth medium. Often medium components from rich media used for fermentations can interfere with recovery or extraction procedures. The effort required to reduce the complexity of the medium, to one as defined as possible, is usually worth the trouble. Having the resulting growth medium as much like water as possible simplifies downstream work. If the cells must be removed from the growth medium and resuspended in a different biotransformation medium, they can be removed by centrifugation or the medium exchanged by microfiltration. Once the biotransformation is complete, the catalyst can be removed by centrifugation or filtration. Concentrating the medium after the biotransformation by removal of water will enhance many recovery processes. The concentration of the medium salts in this process often makes extraction much more efficient.

Although there are many advantages to the use of whole cells for biotransformations, there are certain limitations that must be considered. One consideration is the transport of substrates and products across the cell membrane. In life, the cell membrane is a proton-tight barrier to the rest of the world. It is generally impermeable to charged molecules and to water, but may have permeability to hydrophobic molecules. Often cells have specific transport systems that move compounds in and out as the cell needs them. In order for a whole-cell biotransformation to proceed, the substrate of interest must be transported across the cell membrane to the active enzyme or enzyme system. The same issues exist, of course, for the reaction product. Although it is not generally an issue, for biotransformations where the substrate is not transported but enters the cell by passive diffusion, mass transport must be considered. A simple dilution of the cell density can demonstrate if this is a problem if the rate is not proportional with cell density. Another limitation, which may be observed, is the nonspecific metabolism of the substrate or product by other enzymes in the cell. It is not uncommon to find a small percentage of the substrate or product oxidized or reduced by enzymes not in the desired pathway or reaction. This can sometimes be addressed by genetic techniques or by growth conditions. The commercial-scale use of whole-cell biocatalysts is typically limited to cells produced on-site and transferred under containment to the process vessel given the biosafety and economic issues inherent to transporting live microbes in large quantities.

A variation on the whole-cell biotransformation theme is the use of permeabilized cells. Whole cells can be rendered permeable to small molecules yet remain essentially intact by contacting them for a short time with low concentrations of solvents. This process has the effect of making “holes” in the cellular membrane while leaving enough of the cell membrane and cell wall intact to still contain the enzymes and other macromolecules. The permeabilized cells can then be treated much like immobilized enzymes. This technique is especially useful when transport issues are found to be limiting the reaction.

A related topic appropriate to this discussion is the use of solvents in whole-cell or permeabilized cell biotransformations. Solvents can be added to increase the solubility of substrates in the reaction medium. The specific solvent and concentration must be empirically determined and can be highly variable. Solvents, in low concentration, can serve as a source of energy for reactions requiring reducing equivalents. In such cases the solvents may serve to both supply energy and help solubilize the substrate. Although not generally considered a solvent, the use of a second nonmiscible liquid such as vegetable oil or hexadecane can be used in many whole-cell biotransformations to “buffer” the concentration of substrates and products with respect to the aqueous phase. Two liquid phase reactions can dramatically increase the yield obtainable from a specific biotransformation.

Bioreactor Configurations

The manner in which a bioconversion is performed is dictated by the nature of the biocatalyst, the chemistry involved, and process economics [93]. The overall aims of a bioconversion are the same as for any process, to maximize the production of a given material at the lowest overall cost. In some cases, this might mean maximizing the volumetric productivity (Qp in units of mol m^3/s) of the reactor. Alternately, it might be most important to enable the more efficient recovery through maximizing the ratio of desired to undesired products. If the cost of the biocatalyst is limiting then the catalyst productivity (P_{cat}) must be maximized, a function of the intrinsic activity of the catalyst itself and the fashion in which it is used.

A variety of reactor configurations has been developed for both batch and plug-flow modes of operation. Bioconversions, involving the transformation of a defined substrate into a product, are typically less demanding than fermentations involving growing cells where many additional factors such as oxygenation, feed rate, and the supply of trace nutrients may need to be tightly monitored. Nonetheless, the implementation of bioconversions, particularly larger scale operations, requires a considerable degree of engineering. Important factors that must be considered when selecting a

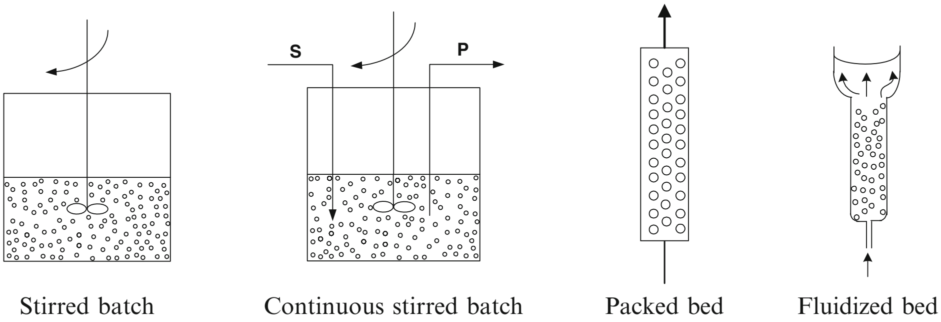


Fig. 31.17 Bioreactor configurations

reactor configuration for performing a given bioconversion include:

- Nature and solubility of substrates and products
- Form of the biocatalyst (whole cells, soluble or immobilized enzymes)
- Cost of biocatalyst relative to the overall process
- Kinetic parameters of the biocatalyst
- Substrate/product equilibrium
- Degree of product or substrate inhibition of the biocatalyst
- Operational stability of biocatalyst
- Product recovery

The bioreactor configurations in common use are illustrated in Fig. 31.17.

Batch reactors. The majority of smaller-scale bioprocesses are carried out in batch mode. A key advantage of many batch reactors is the ability to run a range of different bioconversions with the one piece of equipment. The main disadvantage is the downtime required for emptying and recycling the reactor and cases where product inhibition of the biocatalyst is significant. The simplest configuration involves the use of a stirred vessel containing the biocatalyst to which one or more substrates (S_t concentration) are added resulting in conversion to the desired product(s) over a given time. The required batch time (T_b) of such a system is a function of the kinetic parameters of the biocatalyst (K_m, V_{max}), the catalytic density (maximum enzyme activity per volume), and the initial substrate concentration (S_o), represented in the equation below.

$$(S_o - S_t) - K_m \ln(S_t/S_o) = \text{Catalytic density} \cdot T_b \quad (31.8)$$

In some cases, continuous extraction of products allows extended operation and high volumetric efficiency. Reactors run in this mode are referred to as continuous-stirred batch reactors (CSBR). In this case, the bioconversion is run under approximately steady-state conditions where the position of reaction equilibrium lies toward the products of the conversion. In this case the concentration of product (proportional to $S_i - S_o$) at a given reactor residence time becomes a function of both the flow rate (Q) into the reactor and reactor volume, in addition to the factors discussed above for batch mode reactors (i.e., catalyst parameters and density, inlet substrate concentration S_i, and outlet substrate concentration S_o).

$$\begin{aligned} Q \cdot (S_i \cdot S_o) &= V_{max} \cdot S_o/(K_m + S_o) \\ &= (\text{Catalytic density} \cdot \text{Reactor volume}) \\ &\quad \cdot S_o/(K_m + S_o) \end{aligned} \quad (31.9)$$

Another favorable aspect of stirred batch reactors is the fact that they are compatible with most forms of a biocatalyst. The biocatalyst may be soluble, immobilized, or a whole-cell preparation; in the latter case a bioconversion might be performed in the same vessel used to culture the organism. Recovery of the biocatalyst is sometimes possible, typically when the enzyme is immobilized or confined within a semi-permeable membrane. The latter configuration is often referred to as a membrane reactor. An example is the hollow fiber reactor where enzymes or whole cells are partitioned within permeable fibers that allow the passage of substrates and products but retain the catalyst. A hollow-fiber reactor can be operated in conjunction with the stirred tank and operated in batch or continuous mode. A tubular reactor configuration consisting of hollow fibers is also useful for continuous plug-flow mode operation.

Plug-flow bioreactors. Continuous bioconversions can also be carried out using so-called “plug-flow” reactors where a substrate is converted to a product by passage through a biocatalyst in immobilized form. The packed-bed bioreactor is the most common, although fluidized-bed bioreactors are also used. The principle advantages of packed-bed reactors are the greatly increased catalytic density relative to batch reactors allowing for compact designs and suitability for continuous production. There are several technical issues that need to be addressed before a given bioprocess can be performed in this fashion. The catalyst must be immobilized in a form compatible with a packed-bed format, usually immobilized on rigid particles that enable optimal fluid flow through the bed. The particles must be strong enough to support their own weight and also resist fouling by

Table 31.11 Process metrics for the glucose isomerase process

Parameter	Batch soluble	Batch immobilized	Continuous immobilized
Reactor volume (m^3)	1,100	1,100	15
Temperature (°C)	65	65	60
pH	6.8	6.8	7.6
Normalized comparison			
Reactor size	75	75	1
Enzyme consumption	90	6	1
Activity half-life	1	10	50
Active half-lives	1	3	5
Residence time	40	40	1
Co^{2+} usage	2	1	0
Mg^{2+} usage	6	6	1
Energy usage	5	5	1
Conversion cost	100	6	1
Color formation	1	0.3	<0.1

Source: Courtesy Genencor International

impurities such as particulates in the substrate. The rate of the conversion dictates the required bed thickness, or alternatively, the need for recycling in order to drive a conversion to completion. In this case, the concentration of product (proportional to $S_i - S_o$) at a given reactor residence time (T_p) becomes a function of both the flow rate (Q) into the reactor and reactor volume in addition to the factors discussed above for batch mode reactors (i.e., catalyst parameters and density, inlet substrate concentration S_i, and outlet substrate concentration S_o).

$$(S_i - S_o) - K_m \cdot \ln(S_o/S_i) = \text{Catalytic density}$$

$$T_p = \text{Catalytic density}\cdot(\text{Reactor volume/flow rate}) \qquad (31.10)$$

Continuous bioconversions using packed-bed reactors are particularly well suited for high-volume bioprocesses including HFCS and enzymatic transesterification of fats and oils. As shown in Table 31.11 for an HFCS production reactor using Glucose isomerase, PFR is advantageous in terms of short residence time critical for minimum color formation, high productivity, and low cost. An overall summary of the relative advantages and disadvantages of different bioreactor configuration is listed in Table 31.12.

Recovery of products often dictates the fashion in which a bioconversion is carried out. Liquid/liquid two-phase systems have proven useful, particularly when substrates of low water solubility are used. In these cases it is often possible to operate the reactor in continuous mode by removing and recycling the phase containing the products. The mode of product recovery depends on the nature of the product and the medium in which it is dissolved, in most cases water. Distillation can be used for product recovery in some instances, whereas extraction using a solvent might be used for nonpolar products. In many cases the removal of water from dilute solutions of product can add considerable cost to a bioprocess. For this reason there has been widespread interest in the use of solvents other than water for biocatalysis, discussed below.

Table 31.12 Comparison of bioreactor configurations

Configuration	Advantages	Disadvantages
Stirred batch	• Simple versatile format • Compatible with most biocatalyst forms	• Low catalytic density • Downtime between cycles
Continuous-stirred batch	• Higher volumetric efficiency • Continuous production	• Incomplete conversion to product
Packed bed	• High catalytic density • Extended operation	• Biocatalyst must be immobilized in compatible form
Fluidized bed	Resistance to blockage	Lower catalytic density
Membrane	Compatible with soluble enzymes	Membrane fouling

Nonaqueous Biocatalysis

Enzymes are, for the most part, soluble in water and not obviously suited for use in organic solvents. Many enzymes are denatured by exposure to solvents and still others require water as part of their catalytic action. Prior to 1980, there were several reports of biocatalysis in solvents; however, it was Klibanov and co-workers in the 1980s who first clearly demonstrated the potential of biocatalysis in organic media [94–96]. The use of such media, typically organic solvents, has greatly expanded the scope of biocatalysis for several reasons:

- Ability to use substrates with low water solubility
- Modulation of enzyme regio- and enantioselectivity
- Improved product yields in some cases
- Improved product recovery
- Enhanced biocatalyst stability
- Prevention of microbial contamination

Hydrolytic enzymes such as lipases and proteases catalyze readily reversible reactions and will often promote reverse hydrolysis at reduced water activities. Water can be removed with desiccants, as an azeotrope with a solvent or through application of a vacuum. Lipases have proven particularly useful in this regard, allowing the formation of esters from alcohols and either free carboxylic acids or esters.

An example is the synthesis of fatty acid esters in hexane from fatty acids and alcohols, a reaction catalyzed by many lipases including Porcine pancreatic lipase (PPL), *P. alcaligenes* lipase (LIPOMAX^{TM}, Genencor), CAL-B (N435, Novozymes), and *Candida rugosa* lipase (Amano

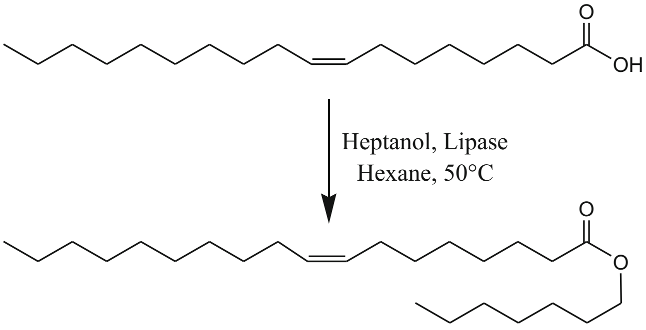


Fig. 31.18 Lipase-mediated ester formation

AYS, Amano) among others (Fig. 31.18). Many other classes of enzyme have also been used in organic solvents such as acyltransferases, glycosidases, dehydrogenases, oxidases, and dehalogenases, for the most part on the laboratory scale. Another major advantage of nonaqueous biocatalysis is the greatly expanded set of possible substrates, particularly those of low or negligible solubility in water. Examples of such compounds include long-chain fatty acids and ester derivatives, steroids, aromatic compounds, and many pharmaceutical intermediates. Several texts have been dedicated to the topic of synthetic biocatalysis, for example, the comprehensive handbook edited by Drauz and Waldmann [97].

A wide variety of organic solvents have been used to conduct bioconversions including nonpolar solvents such as isooctane, *n*-hexane, and toluene, in addition to methanol, acetone, and other water-miscible solvents. Dipolar aprotic solvents dimethylformamide (DMF) and dimethylsulfoxide (DMSO) are also compatible with many enzymes and are often used to enhance the solubility of substrates in combination with a nonpolar solvent. Tertiary alcohols such as *t*-butanol and *t*-amyl alcohol have been used for many lipase-mediated esterifications as sterically hindered tertiary alcohols are not typically good substrates for most enzymes. It should be noted that the presence of small amounts of water is essential for the effective use of most biocatalysts in organic solvents. In some cases an enzyme may only require a monolayer of water molecules on its surface in order to operate. In other cases there may need to be enough water to form reverse micelles where the biocatalyst is contained within a predominantly aqueous environment. The amount of water needed for maximal enzyme performance is usually cited in terms of water activity (A_w). Some enzymes (class “A”) operate efficiently at low water activities and become less efficient at catalyzing certain reactions such as ester formation at high A_w. Other enzymes will operate most efficiently at high water activities (class “B”). The enzyme activity as a function of water activity for these two classes of enzyme is represented graphically in Fig. 31.19.

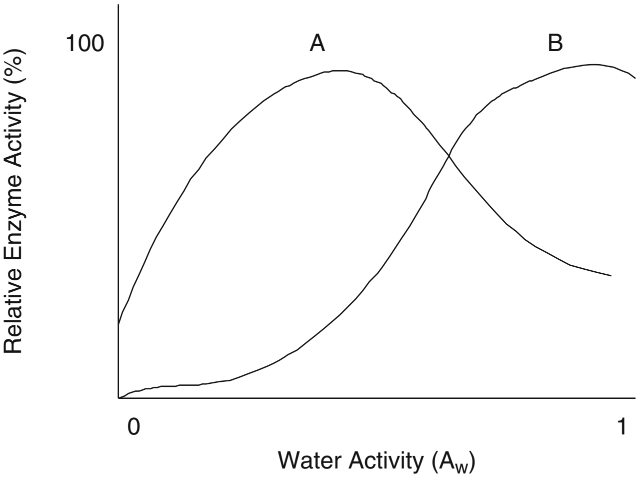


Fig. 31.19 Enzyme activity vs. water activity plot

Few industrially important bioprocesses are carried out exclusively in organic solvents. Two-phase systems are more common, however, consisting of a water-immiscible solvent that allows adequate dissolution of substrates and products and an aqueous phase containing the biocatalyst [98].

Products of Biocatalysis

Although numerous chemicals have been produced at a pilot scale using biocatalysis, there are only a modest number of materials produced at the ton scale or greater [99]. Some materials such as fructose syrup, acrylamide, and aspartame are produced on a large scale (>1,000 t/year), whereas others, including most pharmaceutical intermediates, are manufactured at considerably smaller volumes. A summary of some current products produced through biocatalytic processes is given in Table 31.13.

High-fructose corn syrup. Glucose isomerase (GI) is used for the large-scale isomerization of glucose syrups produced from the enzymatic hydrolysis of corn starch (Fig. 31.20). The product, commonly known as HFCS, typically contains 42% fructose, 50% glucose, 6% maltose, and 2% maltotriose. Fructose, the monosaccharide commonly called fruit sugar, is about 50% sweeter than sucrose, the disaccharide familiarly known as table sugar.

Annual production of HFCS amounts to over ten million tons, selling at around 30 cents/lb for the 42% grade (HFCS-42). HFCS is also available in the form of 55% syrup (HFCS-55), 90% syrup (HFCS-90), or as 99% pure crystalline fructose. The lowest grade, HFCS-42, is 15–20% cheaper than liquid invert sugar on a dry weight basis and is predominantly used in soft drinks. The baking industry ranks as the second largest user. Combining HFCS-42 and

Table 31.13 Selected products of industrial biocatalysis

Product	Biocatalyst	Substrate	Application
Food stuffs			
High-fructose corn syrup	Glucose isomerase	Corn syrup (glucose)	Sweetener
Aspartame	Thermolysin	L-Aspartic acid + D, L-phenylalanine	Artificial sweetener
L-Aspartic acid	Ammonia lyase	Fumaric acid + NH_3	Amino acid, aspartame precursor
L-Carnitine	Agrobacterium	γ-Butyrobetaine	Dietary supplement
L-Methionine	L-Aminoacylase	*N*-Acetyl-DL-methionine	Supplement
Pharmaceuticals			
7-ACA	D-amino acid oxidase and glutaric acid acylase	Cephalosporin C salt	Intermediate for semisynthetic antibiotics
Cephalexin	Penicillin acylase	7-Amino-deacetyoxy-cephalosporanic acid	Antibiotics
6-APA	Penicillin G acylase	Penicillin G	Intermediate for semisynthetic antibiotics
(*S*)-Naproxen	Lipase	(*RS*)-Naproxen ethyl ester	Drug
Isosorbide 2-acetate	Lipase	Isosorbide diacetate	Intermediate
(2*R*, 3*S*)-4-methoxy phenylmethyl glycidate	Lipase	Racemic 4-methoxy-phenylmethyl glycidate	Intermediate
L-Valine	Aminopeptidase	DL-Valinamide	Intermediate
Ethyl (*S*)-4-chloro-3-hydroxybutyrate	Carbonyl reductase	Ethyl acetoacetate	Statin intermediate
Chemicals			
Acrylamide	Nitrile hydratase	Acrylonitrile	Monomer
Cyclodextrins	Glucosyltransferase	Amylose	Formulation aid
Nicotinamide	Nitrile hydratase	3-Cyanopyridine	Vitamin precursor
(*S*)-2-Chloro-propionic acid	Dehalogenase	(RS)-2-Chloro-propionic acid	Intermediate for herbicides
(*S*)-2-Glycidyl-butyrate	Lipase	(RS)-Glycidyl-butyrate	Chemical intermediate
(*S*)-2-Methoxyisopropyl amine	Acylase	(RS)-2-Methoxyiso-propylamine	Herbicide intermediate

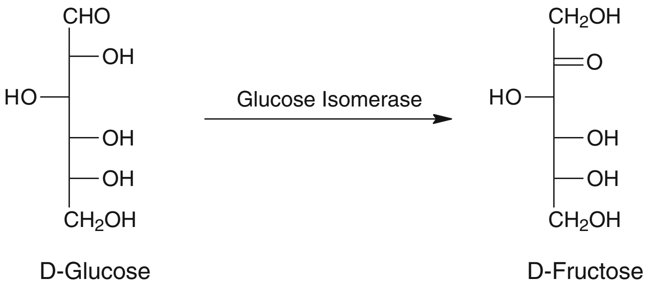


Fig. 31.20 Isomerization of D-glucose to D-fructose catalyzed by glucose isomerase

HFCS-90 gives HFCS-55, which has about the same degree of sweetness as sucrose. It is used as a sweetener and flavor enhancer in fruit-flavored soft drinks. Fructose enhances flavors, whereas sucrose masks them. Because the molecular weight of fructose is approximately half that of sucrose, a smaller amount of it is needed to sweeten a product to desired levels, and sweetener calories in the product can be reduced by about one third. The major use of HFCS-90 is in dietetic foods and drinks. Crystalline fructose of over 99% purity is obtained by drying the 90% pure HFCS. Pure fructose is about 70% sweeter than sucrose, and as an essentially pure sweetener, fructose allows the full taste of product flavors to develop. It has also found uses in low calorie foods and drinks. The amount of HFCS consumed per capita has risen dramatically in the last 30 years and there is some controversy about the possible health effects of high level consumption [100].

In many industrial cases, whole microbial cells containing GI are immobilized by physical means such as entrapment or encapsulation in polymeric materials or by chemical methods such as intermolecular cross-linking with gluteraldehyde or covalent binding with diazotized diamino compounds. Commercially, soluble glucose isomerase is also immobilized on DEAE-cellulose. The immobilized glucose isomerase can usually be used for over 1,000 h at a temperature of around 65°C. When the column enzyme activity decreases, the flow rate of the incoming glucose syrup can be adjusted so that the conversion of glucose to fructose is maintained constant.

Acrylamide. One of the largest scale industrial bioconversions is the process for the bioconversion of acrylonitrile to acrylamide, which is used in coagulants for water treatment, soil conditioners, paper treatment, adhesives, paints, and secondary oil recovery [101]. The worldwide production volume for the bioconversion process is in the tens of thousands of tons per year range. The reaction is a hydration of a nitrile to an amide and has traditionally been done by either a sulfuric acid or a copper-catalyzed

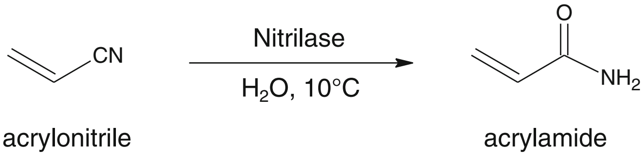


Fig. 31.21 Enzymatic conversion of acrylonitrile to acrylamide

hydration process. These processes are rapidly being phased out and replaced with a bioconversion process that has increased volumetric productivity, higher energy efficiency, decreased costs, and lower environmental impact.

The commercial bioconversion process employs the enzyme nitrile hydratase, which catalyzes the same reaction as the chemical process (Fig. 31.21). The bioconversion process was introduced using wild-type cells of *Rhodococcus* or *Pseudomonas*, which were grown under selective conditions for optimal enzyme induction and repression of unwanted side activities. These biocatalysts have since been replaced with recombinant cells expressing nitrile hydratase. The process consists of growing and immobilizing the whole-cell biocatalyst and then reacting them with aqueous acrylonitrile, which is fed incrementally. When the reaction is complete, the biocatalyst is recovered and the acrylamide solution is used as is. The bioconversion process runs at 10°C compared to 70°C for the copper-catalyzed process, is able to convert 100% of the acrylonitrile fed compared to 80%, and achieves 50% concentration of product compared to 30% concentration. Thus, the bioconversion process does not have to recycle the reaction to complete the hydration and the final liquor does not have to be concentrated as most of the commercial acrylamide is sold as a 50% solution.

It should be noted that this elegant example of replacing a chemical catalyst with a biocatalyst that has higher efficiency, lower costs, and is environmentally friendly did not happen overnight. It was first introduced in 1985 and has become progressively improved to the currently described state. That caveat being said, the current availability of biotechnological tools would today probably be fivefold shorter. The ability to identify and manipulate the enzyme, select the improved catalyst, and express it in a variety of host organisms has improved significantly in the last 20 years.

Semi-synthetic antibiotics. In 1959, Batchelor and coworkers in the Beecham Research Laboratories in England discovered that the penicillin nucleus, 6-APA, accumulated during fermentation when side chain precursors were omitted. This 6-APA could be used for the chemical synthesis of entirely new types of penicillin by coupling with new side chains. Shortly thereafter, several sources of penicillin amidase were found that would cleave the phenylacetyl side chain from penicillin G, thus producing a more economical source of 6-APA. A vast number of "synthetic penicillins" have been generated, and a few have achieved clinical importance [102]. Several objectives were sought:

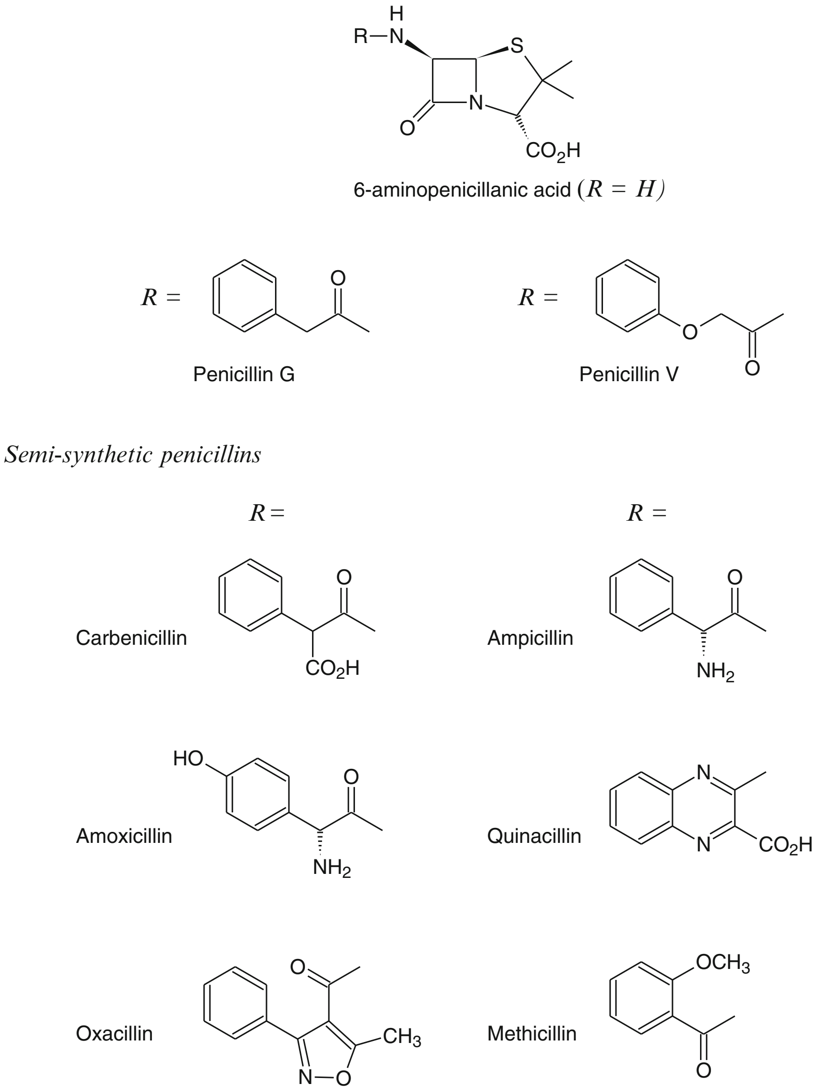
1. To broaden the inherent utility of penicillin to include gram-negative pathogens not inhibited by the natural penicillins
2. To improve its stability and absorption
3. To increase resistance to penicillinase producing pathogens
4. To decrease allergenicity
5. To improve other factors pertinent to clinical use

The broad objectives have been achieved with varying degrees of success. Table 31.14 shows the structures of some of the semisynthetic penicillin that have become important chemotherapeutics.

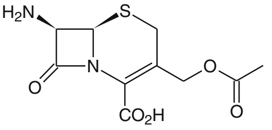
The natural penicillins, primarily G and V, have a relatively narrow spectrum. They act mostly on gram-positive organisms. The fact that proper selection of precursors could lead to new variations in the penicillin side chain offered the first source of synthetic penicillins. Penicillin V, derived from a phenoxyacetic acid precursor, attracted clinical use because of its greater acid tolerance, which made it more useful in oral administration. Also, the widespread use of penicillin eventually led to a clinical problem of penicillin-resistant staphylococci and streptococci. Resistance for the most part involved the penicillin-destroying enzyme, penicillinase, which attacked the beta-lactam structure of the 6-APA nucleus. Semisynthetic penicillins such as ampicillin and carbenicillin have a broader spectrum. Some, such as methicillin, oraficillin, and oxacillin, are resistant to penicillinase. In 1984, Beecham introduced Augmentin, which was the first combination formulation of a penicillin (amoxicillin) and a penicillinase inhibitor (clavulanic acid). Worldwide production of semisynthetic penicillins is currently around 10,000 t/year, the major producers are Smith Kline Beecham, DSM, Pfizer, and Toyo Jozo.

As in the penicillin studies, the chemotherapeutic properties of cephalosporin C can also be improved through synthetic modification of the 7-aminocephalosporanic acid (7-ACA) nucleus. Several semi-synthetic cephalosporins have been produced and are used clinically. The leading agents in this category include cefaclor (first-generation cephalosporin, developed by Lilly), cefoxitin (second-generation, by Merck), cefuroxime axetil (oral form, second-generation, by Glaxo), and ceftriaxone (third-generation, by Roche). Large-scale enzymatic removal of the 6-acyl side chain of penicillins is widely used in the commercial production of both 6-APA and the ring-expanded analogue, 7-aminodesacetoxycephalosporanic acid (7-ADCA). Two enzymes are used commercially, Penicillin acylases G and V. Penicillin G acylase is used in most cases, originally isolated from various sources

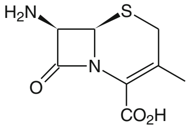
Table 31.14 Structural formulae of natural and semisynthetic penicillins



including *E. coli*, *Bacillus megaterium*, and *Streptomyces lavendulae*. Penicillium acylase V, used to a lesser extent, is derived from organisms including *Beijerinckia indica var. Penicillium*, *Fusarium* sp., and *Pseudomonas acidovorans*. Bacterial acylases have also been found that cleave the 7-acyl side chain of cephalosporin C to form 7-ACA.



7-aminocephalosporanic acid
(7-ACA)



7-aminodesacetoxycephalosporanic acid
(7-ADCA)

Immobilized forms of penicillin amidases and acylases have replaced whole-cell biocatalysts for the production of 6-APA and 7-ACA as they can be reused many times, in some cases for over 1,000 cycles. Another major advantage is the purity of the enzyme, lacking the β-lactamase contaminants often present in whole cells. The productivity of these biocatalysts exceeds 2,000 kg product/kg catalyst. A typical process for the production of 6-APA employs immobilized penicillin G acylase covalently attached to a macroporous resin. The process can be run in either batch or continuous modes. The pH of the reaction must be maintained at a value between 7.5 and 8 and requires continuous adjustment to compensate for the drop caused by the phenylacetic acid generated during the course of the reaction. Recycle reactors have been used, as they allow both pH control and the use of packed bed reactors containing the immobilized catalyst. The enzymatic process is cheaper, although not greatly so, than the older chemical route to

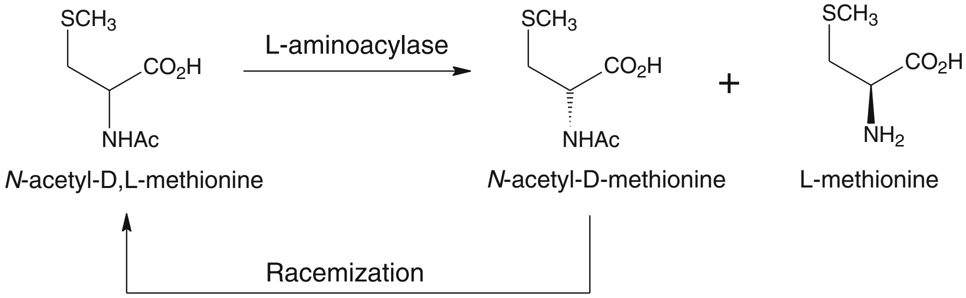


Fig. 31.22 Chiral resolution of racemic *N*-acetyl-DL-methionine

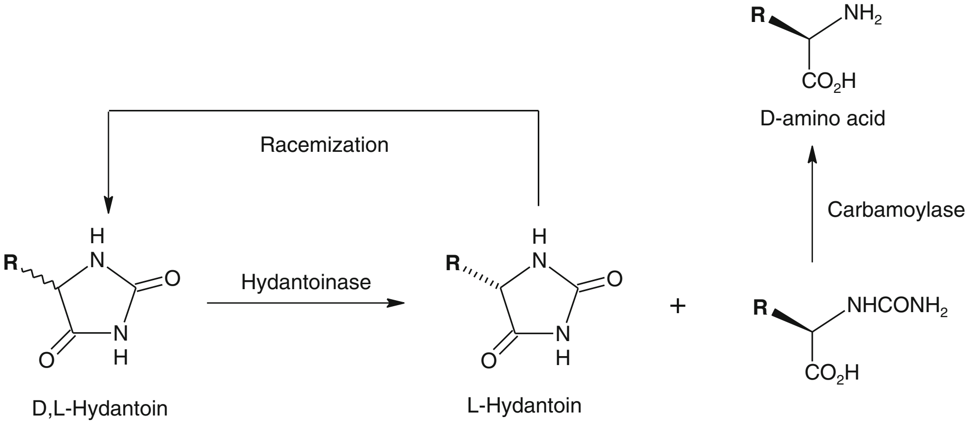


Fig. 31.23 Hydantoinase/carbamoylase process for the production of D-amino acids

6-APA, although it produces far less waste and doesn't require hazardous chemicals such as trimethylsilyl chloride. The biological production of 7-ACA, developed by Hoechst and currently practiced by Biochemie, is another excellent example of the waste reductions possible through implementation of a bioprocess, in this case from 31 to 0.3 kg per kg of product. Overall, the discovery and the development of penicillin and cephalosporin acylases greatly accelerated the commercialization of many semi-synthetic penicillins and cephalosporins and highlight the utility of biocatalysis for the selective modification of complex molecules.

Amino acids. Amino acids are produced by both fermentation and biocatalysis for use in animal feed, fertilizer, as flavor enhancers, dietary supplements, and in pharmaceutical manufacture [103]. By volume, the most important products are L-lysine, L-methionine, L-threonine, and L-tryptophan, most of which are produced by fermentation. Biocatalysis still plays an important role in amino acid supply and has been used to synthesize both L- and D-amino acids, including a variety of nonnatural analogues. The major producers are Degussa, Tanabe Seiyaku, and Kyowa Hakko.

Several enzyme classes have been used to synthesize amino acids, listed below;

- Aminoacylases
- Hydantoinases
- Amidases
- Dehydrogenases
- Racemases
- Lactamases
- Lyases

Aminoacylases (EC 3.5.1.14) were first used for the production of both D- and L-amino acids in the mid-1950s in batch mode processes to effect the selective deacetylation of chemically synthesized racemic *N*-acetyl-DL-amino acids. The most widely used enzyme was derived from *A. oryzae*, selective for the L-isomer. The technology was further developed by Chibata and coworkers at Tanabe to the point where the aminoacylase was immobilized to an ion-exchange resin and used in a packed-bed reactor allowing continuous production [104]. More recently, D-aminoacylases have been isolated and used for the synthesis of D-amino acids. In most schemes, the unwanted isomers are racemized chemically under forcing conditions (pH <3, >80°C) or, alternatively, with a racemase enzyme and subsequently recycled. Examples of amino acids produced through this process include L-methionine, L-valine, L-phenylalanine, and L-tryptophan. The immobilized enzyme is quite stable with a half-life of over 1 month at 50°C. When the total activity of the reactor drops, more of the enzyme can be added to the carrier. Productivity of the continuous process ranges from 70 to over 200 t/year for a 1,000-L column and is approximately 40% less costly than the older batch process. Figure 31.22 depicts the chiral resolution of racemic *N*-acetyl D, L-methionine.

Another widely adopted route to chirally pure amino acids is the hydantoin route. The enzyme hydantoinase

Fig. 31.24 Synthesis of L-*tert*-leucine is carried out using a cofactor recycle process using formic acid as the stoichiometric hydrogen donor

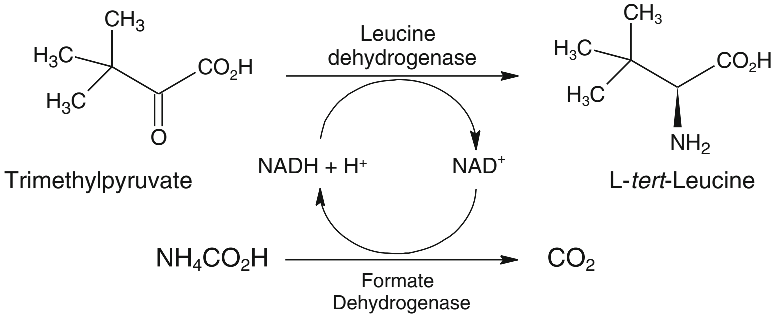
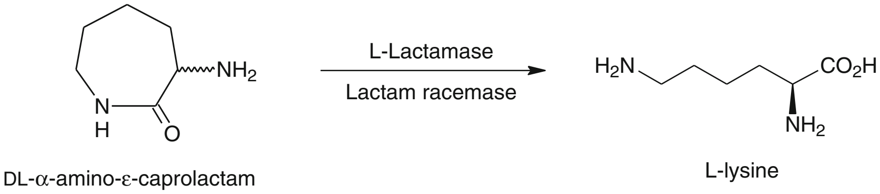


Fig. 31.25 Synthesis of L-lysine from DL-α-amino-ε-caprolactam



allows the selective hydrolysis of L-hydantoins to form a carbamoyl amino acid that is subsequently converted to the free-L-amino acid by carbamoylase (Fig. 31.23). An advantage of this process is the fact that racemization of hydantoins is far more facile than for *N*-acetyl amino acids. This process has been used for many years for the large-scale production (>1,000 t/year) of both D-phenylglycine and 4-hydroxy-D-phenylglycine, used in the synthesis of semi-synthetic antibiotics.

Dehydrogenases are also used for amino acid production, albeit on a far smaller scale. An example is the unnatural amino acid L-*tert*-leucine, manufactured by Evonik (formerly Degussa), used as an intermediate for the manufacture of peptidomimetic drugs such as HIV protease inhibitors. The route used to synthesize this material is noteworthy in that it employs Leucine dehydrogenase, a redox enzyme with cofactor regeneration driven by the decomposition of ammonium formate, in turn catalyzed by formate dehydrogenase (Fig. 31.24). The cofactor in this case is NADH modified with polyethyleneglycol (NADH-PEG) so as to increase its molecular weight and allow retention within a membrane reactor. Total cofactor recycle efficiency is very high at over 125,000 cycles. This elegant approach has also been used to produce other unusual amino acids including β-amino acids.

L-Aspartic acid is used in production of aspartame, in pharmaceuticals, and as a food additive. A bioprocess for the production of this amino acid was first commercialized in 1973 by Tanabe Seiyaku Co. (Japan) and involves the stereoselective addition of ammonia to fumarate catalyzed by L-aspartate ammonia lyase (Aspartase, EC 4.3.1.1). L-Aspartic acid was originally produced by fermentation; however, more recent processes use whole *E. coli* cells immobilized within either an acrylamide or cross-linked polysaccharide matrix. In this case the immobilized cell-based catalyst is far more stable than immobilized forms of the isolated enzyme, with an operational lifetime of over 120 days. The process is typically run in a continuous fashion with recovery of the L-aspartic acid readily achieved through crystallization. The major producers are Tanabe Seiyaku Co. and Kyowa Hakko Co. with a combined production of over 10,000 t per annum.

A similar process is also used for the production of L-malic acid from fumarate, in this case using a hydratase enzyme derived from *Brevibacterium ammoniagenes.* Another variation of the Tanabe technology involves the synthesis of L-alanine from L-aspartic acid through the use of immobilized whole cells (*Pastilla dacunae*) containing aspartate-decarboxylase.

L-Lysine is produced on an enormous scale, over 500,000 t/year, mostly through fermentation using genetically modified organisms. A biocatalytic route was also used to produce L-lysine, the most popular being the Toray process where chiral resolution of DL-α-amino-ε-caprolactam was achieved using a combination of a lactamase selective for the L-lactam and a DL-racemase (Fig. 31.25). A blend of whole-cell biocatalyst strains was used, consisting of *Cryptococcus laurentii* (L-ACL lactamase) and *Achromobacter obae* (ACL racemase). Yields of close to 100% are realized with productivities of over 100 g/L/day. At its peak, this process produced over 10,000 t of L-lysine per year.

Aspartame. Aspartame is a high-intensity sweetener (200-fold that of sucrose) consisting of L-aspartyl-L-phenylalanine methyl ester (L-Asp-L-Phe-OMe). Around 10,000 t are produced per annum, mostly by a chemical process owned by Nutrasweet Corp. Holland Sweetener Company (a DSM/Tosoh joint venture) uses the protease thermolysin (EC

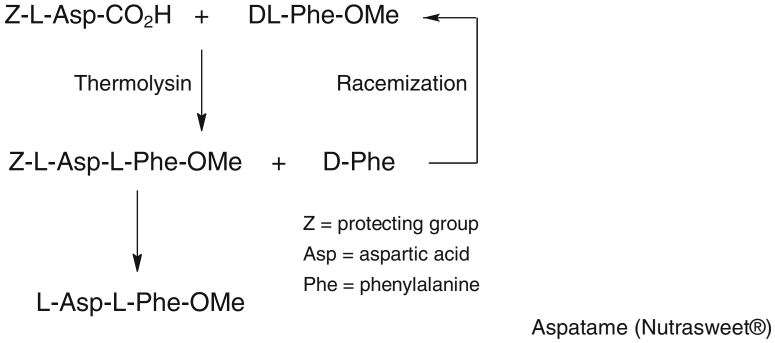


Fig. 31.26 Thermolysin-mediated synthesis of aspartame

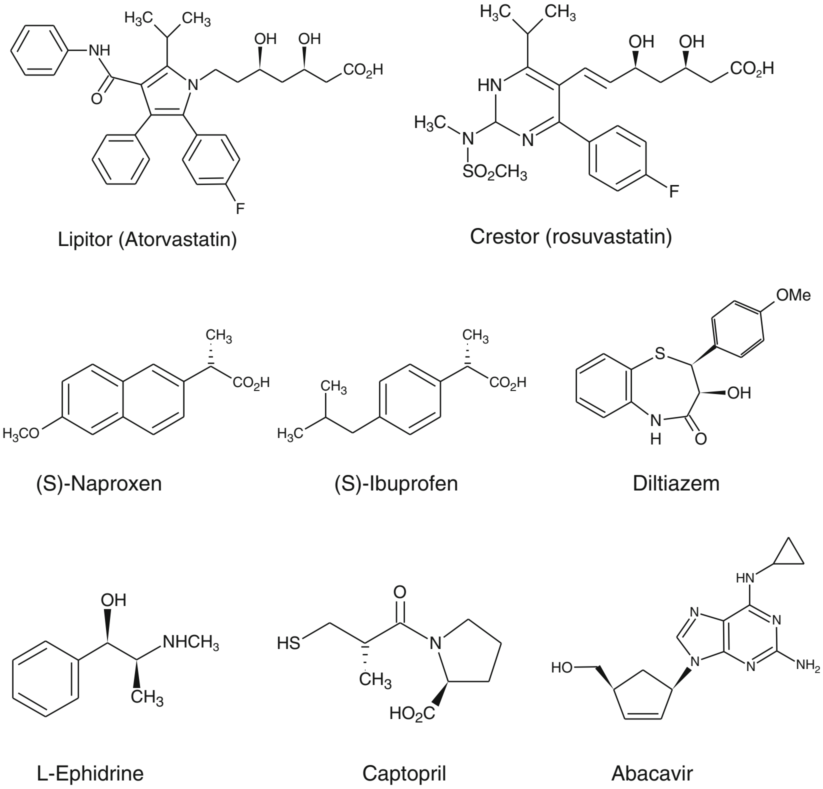


Fig. 31.27 Examples of pharmaceuticals synthesized using biocatalysis

3.4.24.27) to synthesize aspartame from a chemically protected L-aspartic acid derivative (Z-L-Asp) and racemic phenylalanine methyl ester (DL-Phe-OMe). The enzyme only accepts L-phenylalanine-OMe and the unused D-isomer can be isolated, racemized, and recycled. This process is depicted in Fig. 31.26.

Chiral pharmaceutical intermediates. The use of biocatalysis for producing chirally pure pharmaceuticals is now an accepted technology that complements alternative chemical and physical methods [105, 106]. The size of the world market for single enantiomer drugs, over $225 billion in 2005, offers plenty of incentive to develop competitive biocatalytic routes for key intermediates [107]. The market for chiral intermediates themselves, although far smaller, is still considerable and was projected to rise from over $7 billion in 2002 to over $15 billion by 2012 [108]. The majority of these intermediates are generated through traditional chiral pool resolution, 35% by asymmetric chemical synthesis and 15% through the application of biocatalysis (Fig. 31.27) [109]. The number of biocatalytic processes is increasing as more traditional chemical companies begin to view enzymes more as reagents than exotic materials. In most cases biocatalysis is one of many options for the synthesis of given material and must demonstrate economic benefits in order to be adopted.

One of the inherent advantages of enzymes is the ability to discriminate between stereoisomers, often generating products with enatiomeric excesses (i.e., of over 98%). Judicious application of biocatalysis can also reduce the number of chemical steps needed to synthesize certain drugs, leading to hybrid chemoenzymatic processes with lower costs and less waste. The range of enzymes used in

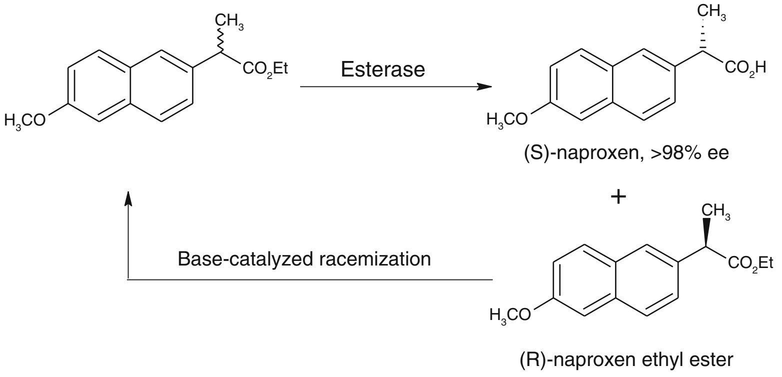


Fig. 31.28 Dynamic chiral resolution of naproxen ethyl ester

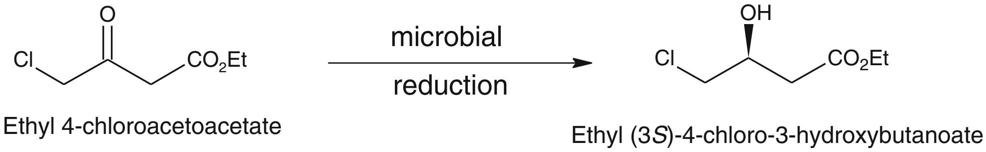


Fig. 31.29 Synthesis of ethyl (3*S*)-4-chloro-3-hydroxybutanoate (ECHB) from ethyl-4-chloroacetoacetate through whole-cell microbial reduction

the synthesis of chiral intermediates has expanded beyond esterases and acylases and now includes oxidoreductases, glycosyltransferases, and C–C bond-forming enzymes (e.g., aldolases). A number of excellent reviews and books have been dedicated to the application of biocatalysis in the production of pharmaceuticals and should be consulted for detailed information [105–111]. A few examples of the use of biocatalysis are discussed below.

Naproxen and ibuprofen are nonsteroidal antiinflammatory drugs (NSAIDs) widely available as OTC medications. In both cases the active (*S*)-enantiomer is far more potent. Selective hydrolysis of racemic esters of these drugs enables the production of both (*S*)-naproxen and (*S*)-ibuprofen. A number of bioroutes have been developed for the synthesis of these drugs. Most commercial processes are based on the selective hydrolysis of racemic esters performed by lipases and esterases, although the selective hydrolysis of nitriles with nitrilases has also been used. One such process for (*S*)-naproxen, was developed by Chirotech using a recombinant esterase, generates titers of around 150 g/L in a batch-mode process. Recovery of the (*R*)-ester is readily achieved by centrifugation and is subsequently racemized by base treatment and returned to the reactor (Fig. 31.28).

The blockbuster anti-cholesterol agents Lipitor® (atorvastatin) and Crestor® (rosuvastatin) generated combined sales of over $15 billion in 2010. Several companies have developed improved biocatalytic methods for the synthesis of ethyl (*S*)-4-chloro-3-hydroxybutyrate (ECHB), a key intermediate for these drugs [112]. One popular route to ECHB involves selective reduction of ethyl 4-chloroacetoacetate with a whole-cell biocatalyst (Fig. 31.29). Daicel Chemical Industries of Japan produces over 100 t/year of ECHB using two enzymes coexpressed in *E. coli* to carry out this reduction. An NADH-dependant carbonyl reductase, originally isolated from *Kluyveromyces aestuarii*, and an engineered formate dehydrogenase dehydrogenose from *Mycobacterium vaccae* resulted in ECHB yields of nearly 50 g/L/h. Alternative biocatalytic routes have been developed by DowPharma in collaboration with Verenium (formerly Diversa), and also by Codexis.

Diltiazem, a benzothiazepin used widely for the treatment of high blood pressure, is derived from a *p*-methoxyphenyl-substituted glycidyl ester with two stereocenters. Lipase-mediated resolution is used to obtain the desired (2*R*, 3*S*)-ester. Both DSM and Sepracor have commercialized processes for this intermediate using lipases including those from *C. rugosa* and *Serratia marcesens*. A typical process is run in a membrane reactor using a two-phase solvent/water mixture. The unwanted isomer is hydrolyzed to a carboxylic acid that subsequently decomposes to an aldehyde and CO_2 (Fig. 31.30). The addition of bisulfite to the reactor prevents inhibition of the lipase by complexing with the aldehyde byproduct and allowing extraction into the aqueous phase.

A number of other important drug intermediates are produced at scale (>1,000 kg/year) by biocatalysis including

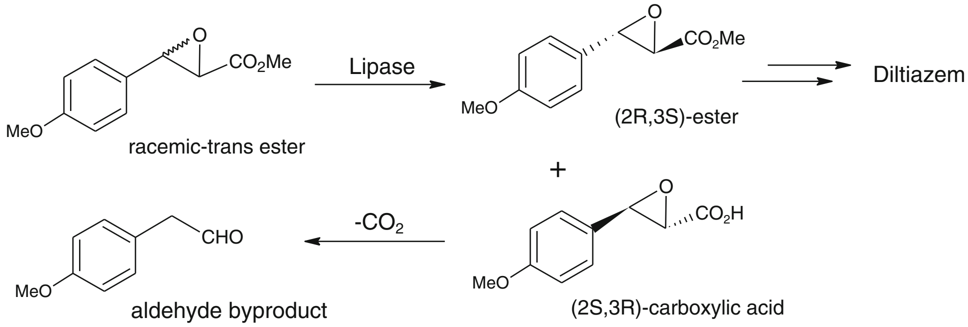


Fig. 31.30 Lipase-mediated synthesis of the key intermediate for Diltiazem

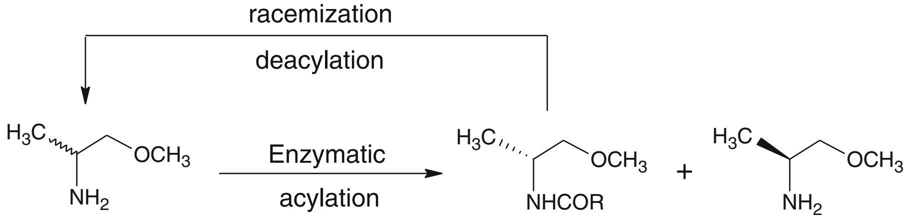
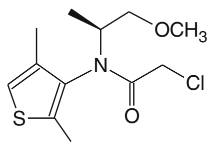


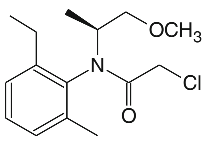
Fig. 31.31 Chiral resolution of 2-methoxyisopropylamine

D-phenylglycine and D-(*p*-hydroxyphenyl)glycine for antibiotics, nicotinamide, 6-hydoxynicotinic acid, (*R*)-glycidol, and D-pantothenic acid. Many fine chemical companies have developed a biocatalysis capability such as Avecia (now a part of Nitto Denko Corp.) and Lonza to address the growing market for chiral intermediates. Others specialize in the development of novel biocatalysts of use for fine chemical synthesis.

Herbicides. Herbicides are used in great quantities globally, accounting for over $18 billion in sales in 2009. A number of herbicides are produced as racemates and, as is the case for many pharmaceuticals, one enantiomer is often more potent than the other. Examples include Frontier, produced by BASF, and Metolochlor, a Syngenta product. The active enantiomer, (*S*)-metolachlor, was introduced in 1999. BASF introduced Outlook^{TM}, the *S*-enantiomer of Frontier, in 2001.



Outlook®



(S)-Metolochlor

Whereas (*S*)-metolochlor is produced chemically by asymmetric hydrogenation, a biocatalytic route is employed by BASF for the production of (*S*)-methoxyisopropylamine on the multithousand-ton scale per annum. Resolution of racemic methoxyisopropylamine is achieved through selective enzymatic acylation of the unwanted (*R*)-isomer. This amide is then separated, racemized, and recycled (Fig. 31.31). A similar approach is used by BASF to produce a range of additional chiral amines as part of the ChiPros® portfolio.

Another chiral small molecule, (*S*)-2-chloropropionic acid (S-CPA), is a key intermediate in the synthesis of the 2-phenoxypropionic acid class of herbicides, including Fusilade (Avecia) and Mecoprop (BASF). Kinetic resolution of racemic CPA is performed with a whole-cell biocatalyst expressing a dehalogenase enzyme selective for the (*R*)-enantiomer and producing (*S*)-lactic acid, itself a useful byproduct. The S-CPA is extracted with a solvent following removal of the biocatalyst by filtration. Removal of the solvent by distillation gives crude S-CPA, which can be further purified. The process was developed by Avecia (purchased by Dr. Reddy's Laboratories in 2005) and produces over 2,000 t per annum.

Carbohydrates. Aside from starch processing and the production of HFCS, there are few biocatalytic processes currently employed for the large-scale synthesis of carbohydrates [113, 114]. A number of carbohydrates are, however, produced on the smaller scale. They include monosaccharides such as xylitol, L-sorbose, and sialic acid, alkyl glycosides, sugar fatty acid esters, cyclodextrins, and oligosaccharides including gluco-, galacto-, and fructooligosaccharides. Many of these processes involve the formation of a glycosidic bond, the carbohydrate version of an acetal or ketal. Three types of enzyme are used for glycoside formation, depicted in Fig. 31.32.

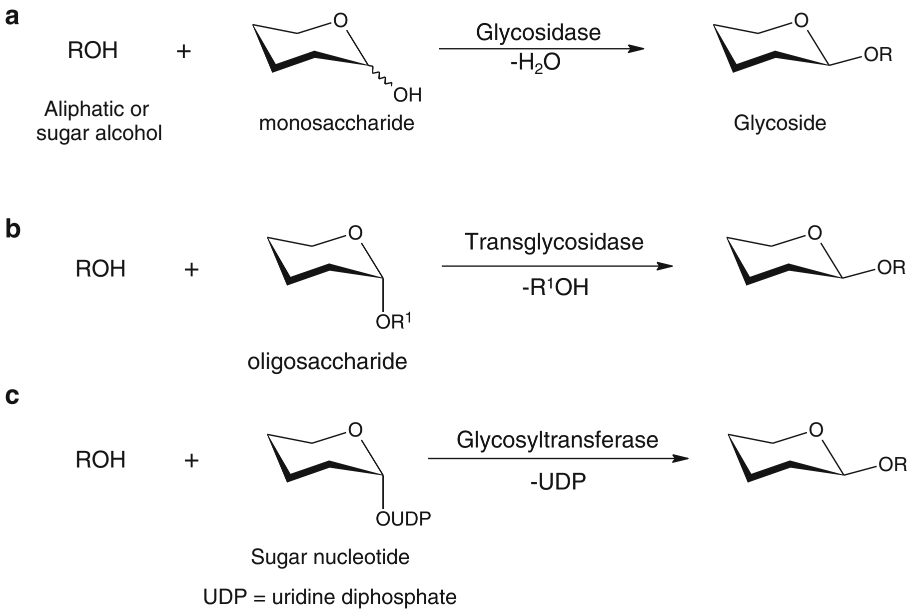


Fig. 31.32 Formation of glycosides using (**a**) a glycosidase and a free sugar; (**b**) a transglycosidase and a glycoside; (**c**) a glycosyltransferase and a sugar nucleotide

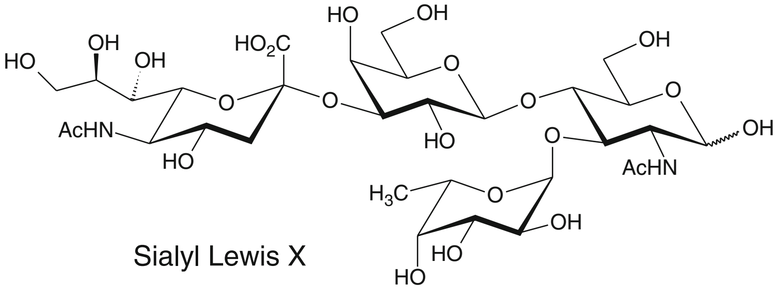
Glycosidases typically hydrolyze glycosidic bonds, but can be made to work in reverse under certain conditions. An example is the β-glucosidase from sweet almonds, which can form glycosides from free (unactivated) sugars and aliphatic alcohols. Transglycosidases are used to interconvert glycosides. Many glycosidases also have transglycosidase activity. The industrial production of butyl-α-D-glucoside involves the transglycosylation of maltose with butyl alcohol in a two-phase system employing the α-transglucosidase from *A. niger* [115].

Glycosyltransferases are the most specific of the enzymes used to form glycosidic bonds and have also been used to synthesize complex oligosaccharides in a far more direct fashion than is possible by chemical means alone. One disadvantage of glycosyltransferases is the need for activated sugar nucleotide donors, such as UDP-glucose (UDP = uridine diphosphate). Recent progress in the practical application of glycosyltransferases has been made possible through the commercial availability of recombinant enzymes and elegant schemes for the in situ regeneration of sugar nucleotide donors using cheaper energy sources such as phosphenolpyruvate.

For example, production of a complex pentasaccharide, Sialyl Lewis X (SLe^x), was achieved on the multikilogram scale by Cytel Corporation using technology developed by Wong and Whitesides whereby all of the required sugar nucleotides were regenerated in a multienzyme scheme [116]. This molecule was originally identified as a possible therapeutic agent for the prevention of reperfusion injury following heart attacks. Although this drug ultimately failed in clinical trials, the technology has since been used in the production of carbohydrates for infant formula, as anti-infectives, and for the production of glycosylated therapeutic proteins.

Future Trends in Biocatalysis

Biocatalysts. Continued discovery of new enzymes and microorganisms will further drive the development of biocatalysis by both enabling new chemistries and lowering the cost of existing technologies. It has been estimated that over 98% of the microbial diversity in Earth remains untapped. The potential of this resource cannot be underestimated. For example, the discovery of extremophilic bacteria



Sialyl Lewis X

existing at temperatures exceeding 120°C and extreme pH values has already led to the development of thermostable enzymes for a variety of applications, of particular note the PCR reaction conducted with DNA polymerases. Libraries of enzymes derived from extremophiles are now commercially available. Table 31.15 lists different classes of extremophiles from which useful enzymes have been derived. Further information on extremophiles and their potential utility for biocatalysis can be found in the reviews by Kumar et al. [117] and Gomes and Steiner [118].

Enzyme engineering will increasingly drive the development of robust enzymes tailored to specific bioprocesses. The strategies used for improving the stability of enzymes will continue to focus on the generation and screening of libraries of mutants produced by random or targeted amino acid substitutions; however, alternative approaches also include the rational design of stabilizing protein motifs (e.g., salt bridges) and even the incorporation of unnatural amino acid analogues [119]. A case in point is the passive incorporation of fluorinated analogues such as hexafluoroisoleucine into proteins by Tirrel and coworkers combined with the observation that fluorinated amino acids can stabilize protein folds in some instances [120]. The ability to design enzymes from unrelated proteins is another long-term goal. Similarly, computational design of enzymes from first principles has great promise, although the field is still in its infancy [121, 122]. Improved production methods for enzymes useful for biocatalysis will also enable the commercialization of many currently uneconomical processes. Gains are also to be expected from refinement of methods used for enzyme immobilization and bioreactor design. The development of improved supports for enzyme immobilization through the application of nanotechnology has particular promise through the enhancement of stability and productivity of immobilized enzymes.

Table 31.15 Classification of extremophiles

Class	Optimal environment
Thermophile	Temperatures above 60°C
Hyperthermophile	Temperatures above 80°C
Psychrophile	Temperatures below 15°C
Halophile	Ionic strength >2 M
Acidophile	pH values below 3
Alkaliphile	pH values over 9
Radiophile	Radioactive environments
Xerophile	Low-water conditions
Piezophile	Pressures-over 100 MPa

Novel media for biocatalysis. The development of novel media in which to conduct biocatalytic transformations has been subject to investigation for almost as long as biocatalysis itself, with greatest focus on the use of organic solvents. In recent years several new directions have gained attention. In particular, the use of ionic liquids (Fig. 31.33) and supercritical carbon dioxide ($scCO_2$) have extended the range of conditions under which biocatalytic processes can be applied, primarily through allowing the solubilization of substrates under conditions amenable to enzymatic catalysis and/or facilitating the isolation of products from the reaction medium [123–126]. The desire to reduce waste generation and eliminate the need for volatile and potentially flammable solvents is also spurring the development of novel media for biocatalysis.

Ionic liquids based on salts of alkylimidazoles and pyridines have been used for the synthesis of esters of ascorbic acid and other sugars on the laboratory scale and show particular advantage in their ability to dissolve otherwise incompatible substrates. One of the key features of these solvents is their extremely low vapor pressures and excellent ability to dissolve both polar and nonpolar materials. Reactions can be run under vacuum allowing the removal of volatile byproducts such as water and enable one to shift the equilibrium of a reaction towards product formation, as in the case of reverse hydrolysis used for the formation of esters, glycosides, and peptide bonds. Product recovery can be facilitated through continuous extraction or phase separation allowing recycling of the solvent. The ability to alter the nature of both the cation and anion of ionic liquids independently affords a degree of tunability, a useful property when attempting to optimize a bioprocess. To date, bioprocesses in ionic liquids have predominantly employed lipases; however, other classes of enzyme such as proteases and glycosidases can also be used. Baker's yeast and other whole-cell biocatalysts have also been used successfully for enantioselective reductions in ionic liquids.

Supercritical fluids (SCFs) and in particular $scCO_2$ have also been shown to provide a unique medium in which to perform biocatalytic transformations. The ability to perform biocatalysis in a SCF was first demonstrated in 1985 by several groups. There are several advantages to the use of

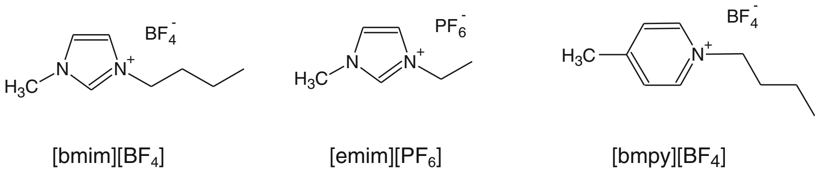


Fig. 31.33 Structures of some ionic liquids

Table 31.16 Critical points of solvents useful for biocatalysis

Solvent	Critical temp (°C)	Critical pressure (MPa)
Carbon dioxide	31.1	7.4
Fluoroform	25.9	4.8
Ethane	32.3	4.9
Sulfur hexafluoride	46	3.8
Ethylene	9.3	5.0

SCFs, as well as some drawbacks that need to be addressed if large-scale processes are to be developed.

Advantages:

- Improved solvation of nonpolar materials
- Improved mass transfer relative to water and other liquids
- Ability to modulate solvent properties through pressure changes
- Ability to remove solvent by depressurization
- Nontoxic and environmentally friendly
- Ease of product recovery

Disadvantages:

- Not compatible with all enzymes
- Need for specialized equipment
- Cosolvents often needed
- Water activity often important
- No large-scale processes developed to date

The conditions under which CO_2 becomes supercritical are mild (Table 31.16) and compatible with many, but not all enzymes. Other SCFs aside from $scCO_2$ have also been used to conduct bioconversions, for example, trifluoromethane, ethane, and propane.

In some cases an enzyme that performs poorly in $scCO_2$ will exhibit greatly improved activity in an alternative SCF. An example is the increased activity of subtilisin in supercritical propane and propane/CO_2 mixtures as compared to $scCO_2$ alone. A particularly interesting property of SCFs is the ability to dramatically alter the solvent properties (e.g., dielectric constant) through changes in temperature and pressure. Increases in pressure can enhance enzyme activity in some cases. Enzyme activity in SCFs is also markedly affected by water activity, as is the case in organic media, and the presence of small amounts of water is often essential for effective catalysis. Lipases are particularly suitable for supercritical biocatalysis and are often used in immobilized form owing to increased stability and ease of handling. Other enzymes have also been employed including proteases, oxidases, and glycosidases. The use of whole cells in supercritical CO_2 has also been reported. Matsuda et al. in 2001 employed the bacterium *B. megaterium* to catalyze the carboxylation of pyrrole at rates exceeding those possible in water [127].

Often a cosolvent is used in order to solubilize particularly polar substrates such as sugars and amino acids. Surfactants or additional solvents may also allow adequate solvation of enzymes. In some cases two-phase systems can be used to conduct bioconversion. For example, Reetz et al. employed both SCFs and ionic liquids in a semi-continuous process for the esterification of alcohols where the enzyme resided in the IL phase and $scCO_2$ enabled substrate solvation and product removal [128].

Products. Biocatalysis will increasingly be used for the production of bulk chemicals including fuels, polymers, and other large-volume materials. Much of this development will be driven by large chemical manufacturers, many of whom have invested heavily in the development of in-house biocatalysis capabilities. Examples include BASF, Dupont, Celanese, DSM, Dow Corning, Lonza, and Evonik. The major factor driving this investment is the realization that in many instances bioprocesses can improve productivity, lower costs, and reduce waste streams. The production and modification of bulk polymers holds particular promise as enzymes may enable the production of polymeric architectures that are not readily made through purely chemical routes, for example, polyesters derived from sugars and other polyols [129].

Production of fine chemicals and active pharmaceutical ingredients (APIs) will also continue to benefit from the judicious application of biocatalysis, in many cases as part of multistep synthetic schemes. Of particular relevance is the increasing demand for chirally pure pharmaceuticals, driven by concerns about the unwanted side-effects often associated with racemic drugs. Another growth area is likely to be the production of biologically active carbohydrates, traditionally requiring complex and expensive chemistries for production, to be used as pharmaceuticals, in infant formula, and as nutritional supplements.

Continued development of the biorefinery concept should eventually lead to sources of renewable materials at costs competitive with those from petrochemical sources through the effective utilization of biomass from various sources (cellulosic, agricultural waste, garbage, vegetable oils). Many of these materials will be amenable to further modification through biocatalytic processes, for example, the production of biodiesel using lipases and fatty acids. Industrial enzymes overall should also see increased use in medicine as diagnostics and as components in biosensors and other devices.

Summary and Conclusion

Enzymes are currently applied across many industries and are used as processing aids, incorporated into consumer products, and used in biocatalysis for chemical production. The global market for industrial enzymes will continue to expand as enzymes find their way into new markets and attract greater interest from the chemical industry at large. Several factors will contribute to this growth:

- Improved knowledge of enzyme mechanisms
- Reduction in the production costs of industrial enzymes
- Improved means for enzyme immobilization and bioprocess engineering
- Novel schemes for cofactor recycling
- Advances in metabolic engineering of whole microorganisms
- Greater awareness of the environmental benefits of enzyme technology

Ultimately, the distinction between biological and chemical catalysis may become less distinct as technologies for modified enzymes, enzyme mimetics, and chemoenzymatic catalysis advance [130]. In the meantime, however, it is clear that further application of enzyme technology within industry will greatly benefit society in the twenty-first century.

References

1. Goodsell DS (2009) The machinery of life. Springer, New York
2. Renneberg R (2008) Biotechnology for beginners. Academic, Amsterdam
3. Jencks WP (1987) Catalysis in chemistry and enzymology. McGraw-Hill, New York
4. Cook PF, Cleland WW (2007) Enzyme kinetics and mechanism. Garland Science, New York
5. Buchholz K, Kasche V, Bornscheuer UT (2005) Biocatalysts and enzyme technology. Wiley-VCH, Weinheim
6. Aehle W (ed) (2007) Enzymes in industry: production and applications, 3rd edn. Wiley-VCH, Weinheim
7. Polaina J, MacCabe AP (eds) (2007) Industrial enzymes, structure, functions and applications. Springer, Dordrecht
8. Whitehurst RJ, Van Oort M (eds) (2010) Enzymes in food technology, 2nd edn. Wiley-Blackwell, Oxford
9. Olempska-Beer ZS et al (2006) Food-processing enzymes from recombinant microorganisms—a review. Reg Toxicol Pharm 45:144–158
10. Dewan SS, Enzymes in industrial applications: global markets (2011) Report BIO030F, BCC Research, Wellesley, MD
11. Sandhya C, Nampoothiri KM Pandey A (2005) Microbial proteases. In: Barredo JL (ed) Microbial enzymes and biotransformations, Methods in biotechnology, vol 17. Humana Press, Totawa, pp 165–179
12. Eggleston G (2007) Advances in the industrial application of enzymes on carbohydrate-based materials. In: Eggleston G, Vercellotti JR (eds) Industrial application of enzymes on carbohydrate-based materials, ACS symposium series, vol 972. American Chemical Society, New York, pp 1–16
13. Aehle W (ed) (2004) Nonindustrial enzyme usage. In: Enzymes in industry: production and applications (Ch. 6), 2nd edn. Wiley-VCH, Weinheim
14. World enzymes (2007) Study #2229, The Freedonia Group, Cleveland, OH
15. Lairson LL et al (2008) Glycosyltransferases: structures, functions, and mechanisms. Annu Rev Biochem 77:521–555
16. Straathof AJJ, Aldercreutz P (eds) (2001) Applied biocatalysis, 2nd edn. Harwood, Amsterdam
17. Bommarius AS, Riebel BR (2004) Biocatalysis—fundamentals and applications. Wiley-VCH, Weinheim
18. Liese A (2000) Industrial biotransformations. Wiley-VCH, Weinheim
19. Kohler R (1972) The reception of Eduard Buchner’s discovery of cell-free fermentation. J Hist Biol 5:327–353
20. Sumner JB (1946) The chemical nature of enzymes. Nobel Lecture, 12 Dec 1946
21. Phillips DC (1967) The hen egg-white lysozyme molecule. Proc Natl Acad Sci U S A 57:484–495
22. Estell DA, Graycar TP, Wells JA (1985) Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation. J Biol Chem 260:6518–6521
23. Manz A, Pamme N, Lossifidis D (2004) Bioanalytical chemistry. World Scientific Publishing Company, Singapore
24. Houg DW, Danson MJ (1999) Extremozymes. Curr Opin Mol Biol 3:39–46
25. Reymond J-L, Fluxa VS, Maillard N (2009) Enzyme assays. Chem Commun 1:34–46
26. Chen LH et al (1992) 4-Oxalocrotonate tautomerase, an enzyme composed of 62 amino acid residues per monomer. J Biol Chem 267:17716–17721
27. Smith S, Witkowski A, Joshi AK (2003) Structural and functional organization of the animal fatty acid synthase. Prog Lipid Res 42:289–317
28. Dill KA, Ozkan SB, Shell MS, Weiki TR (2008) The protein folding problem. Annu Rev Biophys 37:289–316
29. Gevaert K, Vandekerkhove J (2000) Protein identification methods in proteomics. Electrophoresis 21:1145–1154
30. Mittermaier A, Kay LE (2006) New tools provide new insights in NMR studies of protein dynamics. Science 312:224–228
31. Masel RI (2001) Chemical kinetics and catalysis. Wiley-Interscience, Oxford
32. Bugg T (2004) Introduction to enzyme and coenzyme chemistry, 2nd edn. Wiley-Blackwell, Oxford, p 31
33. Benkovic SJ, Hammes GG, Hammes-Schiffer S (2008) Free energy landscape of enzyme catalysis. Biochemistry 47:3317–3321
34. Sauer J et al (2000) Glucoamylase: structure/function relationships and protein engineering. Biochim Biophys Acta 1543:275–293
35. Messerschmidt A (ed) (2011) Handbook of metalloproteins. Wiley-Blackwell, Oxford
36. Eisenmesser EZ et al (2005) Intrinsic dynamics of an enzyme underlies catalysis. Nature 438:117–121
37. Sutcliffe MJ, Scrutton NS (2002) A new conceptual framework for enzyme catalysis. Eur J Biochem 269:3096–3102
38. Menten L, Michaelis MI (1913) Die kinetik der invertinwirkung. Biochem Z 49:333–369
39. Estell DA (1991) Liquid detergent with stabilized enzyme. US Patent 5,039,446
40. Yeh W-K, Yang H-C, McCarthy JR (eds) (2010) Enzyme technologies: metagenomics, evolution, biocatalysis and biosynthesis, Part A. Wiley, New York
41. Martin CH et al (2009) Synthetic metabolism: engineering biology at the protein and pathway scales. Chem Biol 16:277–286
42. Warnecke F, Hess MA (2009) A perspective: metatranscriptomics as a tool for the discovery of novel biocatalysts. J Biotechnol 142:91–95
43. Eggert T (2006) Optimization of industrial enzymes by molecular engineering. In: Liese A, Seelbach K, Wandrey C (eds) Industrial biotransformations (Ch. 4), 2nd edn. Wiley-VCH, Weinheim
44. Rondon MR et al (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Appl Environ Microbiol 66:2541–2547
45. Knietsch A et al (2003) Construction and screening of metagenomic libraries derived from enrichment cultures: generation of a gene bank for genes conferring alcohol oxidoreductase

activity on *Escherichia coli*. Appl Environ Microbiol 69:1408–1416
46. Ward DE et al (2002) Proteolysis in hyperthermophilic microorganisms. Archaea 1:63–74
47. Böttcher D, Bornscheuer UT (2010) Protein engineering of microbial enzymes. Curr Opin Microbiol 13:274–282
48. Dodge T (2009) Production of industrial enzymes. In: Whitehurst RJ, van Oort M (eds) Enzymes in food technology, 2nd edn. Wiley-Blackwell, Oxford, UK
49. Krishna C (2005) Solid-state fermentation system—an overview. Crit Rev Biotechnol 25:1–30
50. Hood EE (2002) From green plants to industrial enzymes. Enzyme Microb Technol 30:279–283
51. Kumar D et al (2008) Microbial proteases and application as laundry detergent additive. Res J Microbiol 3:661–672
52. Outtrup H, Jørgensen ST (2008) The importance of *bacillus* species in the production of industrial enzymes. In: Berkeley R, Heyndrickx M, Logan N, De Vos P (eds) Applications and systematics of Bacillus and relatives. Blackwell Science, Oxford
53. Maurer K-H (2004) Detergent proteases. Curr Opin Biotechnol 15:330–334
54. Choudhary RB, Jana AK, Jha MK (2004) Enzyme technology applications in leather processing. Indian J Chem Technol 11:659–671
55. De Souza PM, Magalhães PO (2010) Application of microbial α-amylase in industry—a review. Braz J Microbiol 41:850–861
56. Gupta R et al (2003) Microbial-amylases: a biotechnological perspective. Process Biochem 38:1599–1616
57. Kumar P, Satyanarayana T (2009) Microbial glucoamylases: characteristics and applications. Crit Rev Biotechnol 29:225–255
58. Ito S et al (2003) Carbohydrate-active enzymes from alkaliphiles. J Appl Glycosci 50:257–262
59. Rausch KD, Belyea RL (2006) The future of co-products from corn processing. Appl Biochem Biotechnol 128:47–86
60. Bhosale SH, Rao MB, Deshpande VV (1996) Molecular and industrial aspects of glucose isomerase. Microbiol Rev 60:280–300
61. Sukumaran RK, Singhania RR, Pandey A (2005) Microbial cellulases—production, applications and challenges. J Sci Ind Res 64:832–844
62. Shallom D, Shoham Y (2003) Microbial hemicellulases. Curr Opin Microbiol 6:219–228
63. Soundari SG, Sashi V (2009) Bacterial xylanases. Asian J Microbiol Biotechnol Environ Sci 11:677–682
64. Hoondal GS et al (2002) Microbial alkaline pectinases and their industrial applications: a review. Appl Microbiol Biotechnol 59:409–418
65. Sun Y, Cheng J (2002) Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresour Technol 83:1–11
66. Shanmugam KT, Ingram LO (2008) Engineering biocatalysts for production of commodity chemicals. J Mol Microbiol Biotechnol 15:8–15
67. Anuradha P et al (2009) Microbial lipases: a potential tool for industrial applications. J Pure Appl Microbiol 3:301–306
68. De Maria L et al (2007) Phospholipases and their industrial applications. Appl Microbiol Biotechnol 74:290–300
69. Singh B, Kunze G, Satyanarayana T (2011) Developments in biochemical aspects and biotechnological applications of microbial phytases. Biotechnol Mol Biol Rev 6:69–87
70. Xu F (2005) Applications of oxidoreductases: recent progress. Ind Biotechnol 1:38–50
71. Bankar SB, Bule MV, Singhal RS, Ananthanarayan L (2009) Glucose oxidase—an overview. Biotechnol Adv 27:489–501
72. Chelikani P, Fita I, Loewen PC (2004) Diversity of structures and properties among catalases. Cell Mol Life Sci 61:19–208
73. Riva S (2006) Laccases: blue enzymes for green chemistry. Trends Biotechnol 24:219–226
74. Majeau J-A, Brar SK, Tyagi RD (2010) Laccases for removal of recalcitrant and emerging pollutants. Bioresour Technol 101:2331–2350
75. Theriot CM, Grunden AM (2010) Hydrolysis of organophosphorus compounds by microbial enzymes. Appl Microbiol Biotechnol 89:35–43
76. Pedreschi F, Kaack K, Granby K (2008) The effect of asparaginase in the formation of acrylamide in French fries. Food Chem 109:386–392
77. Mathews I et al (2007) Structure of a novel enzyme that catalyzes acyl transfer to alcohols in aqueous conditions. Biochemistry 46:8969–8979
78. Ager DJ, May O (2008) Pharmaceutical industry: biocatalysts and chemocatalysts. In: Begley TP (ed) Wiley encyclopedia of chemical biology. Wiley, New York, pp 1–15
79. Takasaki Y, Tanabe O. Enzyme method for converting glucose in glucose syrups to fructose. US Patent 3,616,221
80. Davis BG, Boyer V (2001) Biocatalysis and enzymes in organic synthesis. Nat Prod Rep 18:618–640
81. Kirk O, Christensen MW (2002) Lipases from Candida antarctica—unique biocatalysts from a unique origin. Org Proc Res Dev 6:446–451
82. Greenberg WA et al (2009) Aldolase enzymes for complex synthesis. In: Whittall J, Sutton P (eds) Practical methods for biocatalysis and biotransformations (Ch. 6). Wiley, Chichester
83. Guisan JM (ed) (2006) Immobilization of enzymes and cells (Methods in biotechnology), 2nd edn. Humana Press, Totowa
84. Přenosil JE, Kut ÖM, Dunn IJ, Heinzle E (2009) Biocatalysis, 2. Immobilized biocatalysts. In: Bellussi G et al (eds) Ullmann's encyclopedia of industrial chemistry. Wiley-VCH, Weinheim
85. Berendsen WR, Lapin A, Reuss M (2006) Investigations of reaction kinetics for immobilized enzymes—identification of parameters in the presence of diffusion limitation. Biotechnol Prog 22:1305–1312
86. Fernandez-Lafuente R (2010) Lipase from Thermomyces lanuginosus: uses and prospects as an industrial biocatalyst. J Mol Catal B: Enzym 62:197–212
87. Mateo C et al (2007) Advances in the design of new epoxy supports for enzyme immobilization-stabilization. Biochem Soc Trans 35:1593–1601
88. Kim J, Grate JW, Wang P (2008) Nanobiocatalysis and its potential applications. Trends Biotechnol 26:639–646
89. Betancor L, Luckarift HR (2008) Bioinspired enzyme encapsulation for biocatalysis. Trends Biotechnol 26:566–572
90. Kristenen JB et al (2010) Biomimetic silica encapsulation of enzymes for replacement of biocides in antifouling coatings. Green Chem 12:387–394
91. Becker U et al (2008) Industrial applications of whole-cell biocatalysis. Pharm Technol 41(S6):S8–S11
92. Bechtold M, Panke S (2009) In situ product recovery integrated with biotransformations. Chimia 63:345–348
93. Shuler ML, Kargi F (2002) Bioprocess engineering: basic concepts, 2nd edn. Prentice-Hall, Upper Saddle River
94. Koskinen AMP, Klibanov AM (eds) (1995) Enzymatic reactions in organic media. Springer, Berlin
95. Klibanov AM (2001) Improving enzymes by using them in organic solvents. Nature 409:241–246
96. Adlercreutz P (2008) Fundamentals of biocatalysis in neat organic solvents. In: Carrea G, Riva S (eds) *Organic synthesis with enzymes in non-aqueous media (*Ch. 1*)*. Wiley-VCH, Weinheim
97. Drauz K, Waldmann H (eds) (2002) Enzyme catalysis in organic synthesis, 2nd edn. Wiley-VCH, Weinheim
98. Krieger N et al (2004) Non-aqueous biocatalysis in heterogeneous solvent systems. Food Technol Biotechnol 42:279–286

99. Tufvesson P et al (2010) Process considerations for the scale-up and implementation of biocatalysis. Food Bioprod Process 88:3–11
100. Parker K, Salas N, Nwosu VC (2010) High fructose corn syrup: production, uses and public health concerns. Biotechnol Mol Biol Rev 5:71–78
101. Yamada H, Kobayashi M (1996) Nitrile hydratase and its application to industrial production of acrylamide. Biosci Biotechnol Biochem 60:391–1400
102. Bruggink A (ed) (2001) Synthesis of β-lactam antibiotics: chemistry, biocatalysis & process integration. Kluwer Academic, Dordrecht
103. Sonke T, Kaptein B, Schoemaker HE (2009) Use of enzymes in the synthesis of amino acids. In: Hughes AB (ed) Amino acids, peptides and proteins in organic chemistry: origins and synthesis of amino acids (Ch. 3), vol 1. Wiley-VCH, Weinheim
104. Tosa T, Mori T, Fuse N, Chibata I (1969) Studies on continuous enzyme reactions 6: enzymatic properties of DEAE-sepharose aminoacylase complex. Agr Biol Chem 33:1047–1056
105. Patel R (ed) (2006) Biocatalysis in the pharmaceutical and biotechnology industries. CRC Press, Florida
106. Woodley JM (2008) New opportunities for biocatalysis: making pharmaceutical processes greener. Trends Biotechnol 26:321–327
107. Erb S (2006) Single-enantiomer drugs poised for further market growth. Pharm Technol 30(suppl):s14–s18
108. Meyer H-P, Turner NJ (2009) Biotechnological manufacturing options for organic chemistry. Mini-Rev Org Chem 6:300–306
109. Stinson SC (2001) Chiral pharmaceuticals. Chem Eng News 79:79–97
110. Tao J, Xu J-H (2009) Biocatalysis in development of green pharmaceutical processes. Curr Opin Chem Biol 13:43–50
111. Pollard DJ, Woodley JM (2007) Biocatalysis for pharmaceutical intermediates: the future is now. Trends Biotechnol 25:66–73
112. Liljeblad A, Kallinen A, Kanerva LT (2009) Biocatalysis in the preparation of the statin side chain. Curr Org Synth 6:362–379
113. Chen X, Kowal P, Wang PG (2000) Large-scale enzymatic synthesis of oligosaccharides. Curr Opin Drug Discov Devel 3:756–763
114. Taniguchi H (2008) Biocatalysis-based development of oligosaccharides in Japan. In: Hou CT, Shaw J-F (eds) Biocatalysis and bioenergy (Ch. 17). Wiley, New York
115. Monsan P, Paul F, Pelenc V, Boures E (1996) Enzymatic production of α-butylglucoside and its fatty acid esters. Ann New York Acad Sci 799:633–641
116. Wong C-H, Whitesides G (1994) Enzymes in synthetic organic chemistry. In: Wong C-H et al (eds) Tetrahedron organic chemistry, vol 12. Pergamon, Oxford
117. Kumar L, Awasthi G, Singh B (2011) Extremophiles: a novel source of industrially important enzymes. Biotechnology 2011:1–15
118. Gomes J, Steiner W (2004) The biocatalytic potential of extremophiles and extremozymes. Food Technol Biotechnol 42:279–286
119. Arnold U (2009) Incorporation of non-natural modules into proteins: structural features beyond the genetic code. Biotechnol Lett 31:1129–1139
120. Tang Y, Tirrell DA (2001) Biosynthesis of a highly stable coiled-coil protein containing hexafluoroleucine in an engineered bacterial host. J Am Chem Soc 123:11089–11090
121. Marti S et al (2008) Computational design of biological catalysts. Chem Soc Rev 37:2634–2643
122. Vieceli J, Muellegger J, Tehrani A (2006) Computer-assisted design of industrial enzymes. The resurgence of rational design and in silico mutagenesis. Ind Biotechnol 2:303–308
123. Gorke J, Srienc F, Kazlauskas R (2010) Toward advanced ionic liquids. Polar, enzyme-friendly solvents for biocatalysis. Biotechnol Bioprocess Eng 15:40–53
124. van Rantwijk F, Madeira Lau R, Sheldon RA (2003) Biocatalytic transformations in ionic liquids. Trends Biotechnol 21:131
125. Cantone S, Hanefeld U, Basso A (2007) Biocatalysis in non-conventional media—ionic liquids, supercritical fluids and the gas phase. Green Chem 9:954–971
126. Matsuda T, Harada T, Nakamura K (2005) Biocatalysis in supercritical CO_2. Curr Org Chem 9:299
127. Matsuda T, Ohashi Y, Harada T, Yanagihara R, Nagasawa T, Nakamura K (2001) Conversion of pyrrole to pyrrole-2-carboxylate by cells of *Bacillus megaterium* in supercritical CO_2. Chem Commun 21:2194–2195
128. Reetz MT, Wiesenhoefer W, Francio G, Leitner W (2002) Biocatalysis in ionic liquids: batchwise and continuous flow processes using supercritical carbon dioxide as the mobile phase. Chem Commun 9:992–993
129. Cheng HN, Gross RA (2008) Polymer biocatalysis and biomaterials: current trends and developments. In: Cheng HN, Gross RA (eds) Polymer biocatalysis and biomaterials II, ACS symposium series, vol 999, pp 1–20
130. Wohlgemuth R (2007) Interfacing biocatalysis and organic synthesis. J Chem Technol Biotechnol 82:1055–1062

Industrial Production of Therapeutic Proteins: Cell Lines, Cell Culture, and Purification

32

Marie M. Zhu, Michael Mollet, and Rene S. Hubert

The biotechnology and pharmaceutical industries have seen a recent surge in the development of biological drug products manufactured from engineered mammalian cell lines. Since the hugely successful launch of human tissue plasminogen activator in 1987 and erythropoietin in 1988, the biopharmaceutical market has grown immensely. Global sales in 2003 exceeded US $30 billion [1]. Currently, a total of 108 biotherapeutics are approved and available to patients (Table 32.1). In addition, 324 medically related, biotechnology-derived medicines for nearly 150 diseases are in clinical trials or under review by the US Food and Drug Administration [2]. These biopharmaceutical candidates promise to bring more and better treatments to patients. Compared to small molecule drugs, biotherapeutics show exquisite specificity with fewer off-target interactions and improved safety profiles. Protein engineering technologies have advanced to create protein drugs with improved efficacy, specificity, stability, pharmacokinetics, and solubility. Strategies that have been employed to implement these changes include mutagenesis, recombination, and other directed evolution methods, as well as rational design and structure-based computational approaches [3–7]. These advanced protein engineering technologies are creating novel drug designs and clever treatment strategies that are fuelling the biopharmaceutical market growth.

Currently, 60–70% of all biotherapeutics based on recombinant proteins are produced in cultivated mammalian cells [8]. Mammalian systems are often preferred over other hosts, such as bacteria, plants, and yeast, because of their capability for proper protein folding, assembly, and posttranslational modification. The quality and efficacy of a protein can thus be enhanced when expressed in mammalian cells. With the recent expansion in the biotherapeutics market, the demand for proteins derived from mammalian cells continues to grow, and the biopharmaceutical industry is faced with the challenge of efficiently producing these proteins in large quantity. To keep up with demand while driving manufacturing costs down, mammalian cell production expectations are rising every year, with product titers reaching 4 g of protein/L of culture [8]. The high yields obtained today are the result of combined efforts made in improving host cells, expression vectors, screening methods, medium development, and process development. In this chapter, we outline the common methods applied to mammalian cell line development and describe the typical industrial processes used in cell culture and purification for the production of recombinant proteins. Recent advances in this field are also presented.

Cells Used For Industrial Production

This section focuses on the generation of engineered mammalian cell lines that stably produce therapeutic proteins. Hybridomas, transient expression systems, and insect and bacterial cell line development are outside the scope of this section. The commonly used cell lines, expression systems and vectors, as well as cell banking and stability are described.

Host Cell Lines

Mammalian cell lines commonly used to manufacture therapeutic proteins include NS0 mouse myeloma, baby hamster kidney (BHK), and Chinese hamster ovary (CHO) derived lines such as CHO-S, CHO-K1, CHODUXB11, and CHO-DG44. The latter two cell lines are engineered to have single or double knockout mutations, respectively, for the dihydrofolate reductase (DHFR) gene. NS0 and CHO cell lines are also available with the glutamine synthetase gene

---

M.M. Zhu (✉) • R.S. Hubert
Technical Operations, Xencor Inc., Monrovia, CA, USA
e-mail: mzhu@agensys.com

M. Mollet
Technical Operations, Xencor Inc., Monrovia, CA, USA

Process Development, Medimmune Inc., Gaithersburg, MD, USA

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_32,

Table 32.1 Approved biotechnology products

Trade name	Proper name	Manufacturer	Approval Date	Application	Cell type
Abbokinase	Urokinase	Abbott Laboratories	10/2002	Lysis of acute massive pulmonary emboli	Human neonatal kidney
Actimmune	Interferon gamma-1b	InterMune Pharmaceuticals	12/1990	Chronic granulomatous disease/osteopetrosis	NA
Activase	Alteplase recombinant	Genentech	11/1987	Acute myocardial infarction/acute massive pulmonary embolism/ischemic stroke	CHO
AcuTect	Tc-99 m apcitide	Berlex	1998	Scintigraphic imaging of acute venous thrombosis	NA
Adagen	Injection pegademase	Enzon	1990	Severe combined immunodeficiency disease	NA
Aldurazyme	Laronidase	Biomarin	4/2003	Hurler and Hurler–Scheie forms of mucopoly saccharidosis I	CHO
Alferon N injection	Interferon alfa-n3	Hemispherx Biopharma	10/1989	Genital warts	NA
Amevive	Alefacept	BiogenIdec	1/2003	Moderate to severe chronic plaque psoriasis	CHO
Apligraf	Graftskin	Novo Novartis	5/1998	Venous leg ulcers	NA
Aranesp	Darbepoetin alfa	Amgen	9/2001	Anemia associated with chronic renal failure/ chemotherapy-induced anemia	CHO
Argatroban	–	GlaxoSmithKline	6/2000	Heparin-induced thrombocytopenia syndrome	NA
Avastin	Bevacizumab	Genentech	2/2004	Metastatic colorectal cancer	CHO
Avonex	Interferon beta-1a	BiogenIdec	5/1996	Multiple sclerosis	CHO
BeneFIX	Human factor IX	Wyeth	2/1997	Hemophilia B	CHO
Betaseron	Interferon beta-1b	Chiron/Berlex	7/1993	Multiple sclerosis	E. coli
Bexxar	Tositumomab and Iodine I-131 tositumomab	Corixa	6/2003	CD20 positive, follicular, non-Hodgkin's lymphoma	Mammalian cells
BioTropin	Human growth factor	Bio General - Technology	5/1995	Growth hormone deficiency in children	NA
Botox	Botulinum toxin type A	Allergan	4/2002	Cervical dystonia/moderate to severe glabellar lines	C. botulinum
Campath	Alemtuzumab	Millennium/Ilex	5/2001	B-cell chronic lymphocytic leukemia	CHO
Carticel	cultured Autologous chondrocytes	Genzyme	8/1997	Symptomatic cartilaginous defects of the femoral condyle	Human cartilage
CEA-Scan	Arcitumomab	Immunomedics	6/1996	Imaging agent for detection of colorectal carcinoma	NA
Cerezyme	Imiglucerase for injection	Genzyme	5/1994	Type 1 Gaucher disease	CHO
Comvax	Haemophilus b conjugate	Merck	10/1996	Vaccination against both invasive haemophilus influenzae b and hepatitis B	S. cerevisiae
DACS SC	Stem cell enrichment device	Dendreon	9/1999	Rescue therapy following high-dose chemotherapy	NA
Elitek	Rasburicase	Sanofi-Synthelabo	7/2002	Management of plasma uric acid levels in patients with leukemia, lymphoma, and solid tumor malignancies	S. cerevisiae
Enbrel	Etanercept	Amgen	11/1998	Moderately to severely active rheumatoid arthritis/ polyarticular course juvenile rheumatoid arthritis/ psoriasis	CHO
Engerix-B	Hepatitis vaccine B	GlaxoSmithKline	8/1989	Hepatitis B	S. cerevisiae; CHO
Epogen	Epoetin alfa	Amgen	6/1989	Multiple types of anemia	CHO
Erbitux	Cetuximab	ImClone Systems	2/2004	Advanced colorectal cancer	NS0
Fabrazyme	Agalsidase beta	Genzyme	4/2003	Fabry disease	CHO
Follistim	Follicle-stimulating hormone	Organon	9/1997	Infertility	CHO

(continued)

Table 32.1 (continued)

Trade name	Proper name	Manufacturer	Approval Date	Application	Cell type
Fortéo	Teriparatide	Eli Lilly	12/2002	Osteoporosis	E. coli
Fuzeon	Enfuvirtide	Roche	3/2003	HIV infection/AIDS	NA
Genotropin	Somatropin	Pfizer	8/1995	Short stature in children due to growth hormone deficiency	E. coli
Geref	Human growth hormone	Serono	12/1990	Pediatric growth hormone deficiency	NA
GlucaGen	Glucagon	Novo Nordisk	6/1998	Hypoglycemia	S. cerevisiae
Gonal-F	Follicle-stimulating hormone	Serono	9/1997	Female infertility	CHO
Helixate FS	factor Antihemophilic	Bayer	6/2000	Hemophilia A	BHK
Herceptin	Trastuzumab	Genentech	9/1998	Metastatic breast cancer	CHO
Humalog	Insulin lispro	Eli Lilly	10/1982	Diabetes	NA
Humatrope	Somatropin	Eli Lilly	3/1987	Human growth hormone	E. coli
Humira	Adalimumab	Laboratories Abbott	12/2002	Moderately to severely active rheumatoid arthritis	CHO
Humulin	Human insulin	Eli Lilly	10/1982	Diabetes	E. coli; S. cerevisiae
Infergen	Interferon alfacon-1	Amgen	10/1997	Chronic hepatitis C virus (HCV) infection	E. coli
Infergen	Interferon alfa-2a	Hoffman-LaRoche	9/2003	Removal of Kaposi's sarcoma indication	NA
Intron A	Interferon alfa-2	Schering-Plough	6/1986	Hairy cell leukemia/genital warts/AIDS-related Kaposi's sarcoma/hepatitis C/follicular lymphoma in conjunction with chemotherapy/hepatitis B	E. coli
Kineret	Anakinra	Amgen	11/2001	Moderately to severely active rheumatoid arthritis	E. coli
KoGENate	Antihemophiliac factor	Bayer	2/1993	Clotting treatment of hemophilia A	BHK
KoGE-Nate-FS	rFVIII	Bayer	6/2000	Clotting treatment of hemophilia A	BHK
Leukine	Sargramostim	Berlex	3/1991	Colony autologous bone marrow transplantation	Yeast
LymErix	Lyme disease vaccine	GlaxoSmithKline	12/1998	Prevention of lyme disease	E. coli
Mylotarg	Gemtuzumab ozogamicin	Wyeth	5/2000	CD-33-positive acute myeloid leukemia	Micromonospoa echinospora. sp. calichensis
Myoblocs	Botulinum toxin type B	ElanPharmaceuticals	12/2000	Cervical dystonia to reduce the severity of abnormal head position and neck pain	C. botulinum
Myoscint	Imciromab pentetate	Centocor	7/1996	Cardiac imaging agent for detection of myocardial necrosis	Mammalian cells
Natrecor	Nesiritide	Scios	8/2001	Acute decompensated congestive heart failure	E. coli
Neulasta	Pegfilgrastim	Amgen	1/2002	Chemotherapy-induced neutropenia	E. coli
Neumega	Oprelvekin	Genetics Institute	11/1997	Prevention of chemotherapy-induced thrombocytopenia	E. coli
Neupogen	Filgrastim	Amgen	2/1991	Chemotherapy-induced neutropenia/autologous bone marrow transplantation/chronic severe neutropenia/acute myeloid leukemia	E. coli
Norditropin	Somatropin	Novo Nordisk	5/1995	Growth failure in children due to inadequate growth hormone secretion	E. coli, CHO
Novolin 70/30	Human insulin	Novo Nordisk	6/1991	Insulin-dependent diabetes mellitus	S. cerevisiae; E. coli
Novolin L	Human zinc insulin	Novo Nordisk	6/1991	Insulin-dependent diabetes mellitus	S. cerevisiae; E. coli
Novolin N	Human insulin isophane	Novo Nordisk	7/1991	Insulin-dependent diabetes mellitus	S. cerevisiae; E. coli
Novolin R	Regular, human insulin	Novo Nordisk	6/1991	Insulin-dependent diabetes mellitus	S. cerevisiae; E. coli
NovoLog	Insulin aspart	Novo Nordisk	11/2001	Insulin-dependent diabetes mellitus	S. cerevisiae; E. coli

(continued)

Table 32.1 (continued)

Trade name	Proper name	Manufacturer	Approval Date	Application	Cell type
NovoSeven	Factor VIIa	Novo Nordisk	3/1999	Bleeding episodes in hemophilia A or B	BHK
Nutropin	Somatropin	Genentech	3/1994	Growth failure in children due to chronic renal insufficiency	E. coli
Nutropin AQ	Somatropin	Genentech	12/1995	Growth failure in children due to chronic renal insufficiency	E. coli
Oncospar	PEG-L-asparaginase	Enzon	2/1994	Acute lymphoblastic leukemia	NA
Ontak	Denileukin diftitox	Ligand Pharmaceuticals	2/1999	Persistant or recurrent cutaneous T-cell lymphoma	E. coli
Orthoclone Okt3	Muromonab-CD3	Ortho Biotech	6/1991	Reversal of acute kidney transplant rejection	NA
Ovidrel	Choriogonado -tropin alfa	Serono	10/2003	Induction of final follicular maturation and early luteinization	CHO
Pegasys	Peginterferon alfa-2a	Hoffman-La Roche	10/2002	Chronic hepatitis C	E. coli
PEG-Intron	Peginterferon alfa-2b	Schering-Plough	1/2001	Chronic hepatitis C	E. coli
Procrit	Epoetin alfa	Ortho Biotech	12/1990	Anemia	CHO
Proleukin	Aldesleukin	Chiron Corporation	1/1998	Metastatic melanoma/renal cell carcinoma	E. coli
ProstaScint	Capromab Pendetide	Cytogen Corp	10/1996	Labeled with Indium 111 for imaging of prostatic carcinoma	Mammalian cells
Protropin	Somatrem	Genentech	10/1985	Growth hormone deficiency in children	E. coli
Pulmozyme	Dornase alfa	Genentech	12/1993	Cystic fibrosis/advanced cystic fibrosis	CHO
Raptiva	Efalizumab	Genentech	10/2003	Chronic plaque psoriasis	CHO
Rebetron	Ribavirin/ interferon alfa-2b combination	Schering-Plough	6/1998	Chronic hepatitis C	E. coli
Rebif	Interferon beta-1a	Serono	5/2003	Relapsing forms of multiple sclerosis	CHO
Recombinat	rAHF	Baxter Healthcare	12/1992	Control of bleeding episodes in hemophilia A	NA
ReFacto	Antihemophilic factor VIII	Wyeth	3/2000	Clotting hemophilia A	CHO
Refludan	Lepirudin	Berlex Laboratories	3/1998	Heparin-induced thrombocytopenia	NA
Regranex	Becaplermin	Ortho-McNeil	12/1997	Diabetic ulcers	S. cerevisiae
Remicade	Infliximab	Centocor	08/1998	Moderately to severely active Crohn's disease/ rheumatoid arthritis	NS0
ReoPro	Abciximab	Centocor/Eli Lilly	12/1994	Antiplatelet prevention of blood clots/refractory unable angina	Mammalian cells
Retavase	Reteplase	Centocor	10/1996	Acute myocardial infarction in adults	CHO
Rituxan	Rituximab	Genentech, BiogenIdec	11/1997	Relapsed or refractory low-grade or follicular, B-cell non-Hodgkin's lymphoma	CHO
Roferon-A	Interferon alfa-2a	Hoffman-LaRoche	6/1986	Hairy cell leukemia/AIDS-related Kaposi's sarcoma/ hepatitis C/chronic myelogenous leukemia	E. coli
Saizen	Somatropin	Serono	10/1996	Pediatric growth hormone deficiency	Mouse C127
Serostim	Somatropin	Serono	8/1996	AIDS-associated catabolism	Mouse C127
Simulect	Basiliximab	Novartis	5/1998	Prevention of renal transplant rejection	Myeloma cell
Synagis	Palivizumab	Medimmune	6/1998	Prophylaxis of serious lower respiratory tract disease caused by respiratory syncytial virus	NS0
Thymoglobulin	Thymocyte globulin	SangStat	12/1998	Prevention of kidney transplant rejection	NA
Thyrogen	Thyrotropin alfa	Genzyme	11/1998	Thyroid cancer follow-up	CHO
TNKase	Tenecteplase	Genentech	6/2000	Reduction of mortality associated with acute myocardial infarction (AMI)	CHO
Trisenox	Arsenic trioxide	Cell Therapeutics	6/2000	Acute promyelocytic leukemia	NA
Velosulin BR	Human insulin	Novo Nordisk	7/1999	Diabetes mellitus	S. cerevisiae

(continued)

Table 32.1 (continued)

Trade name	Proper name	Manufacturer	Approval Date	Application	Cell type
Verluma	Nofetumomab	DuPont Pharma	8/1996	Detection of small cell lung cancer	NA
Visudyne	Verteporfin	QLT	4/2000	Minimally classic age-related macular degeneration	NA
Vitravene	Fomiviren sodium	Ciba Vision	8/1998	Cytomegalovirus retinitis	NA
Xifaxan	Ritaximin	Salix Pharmaceuticals	5/2004	Traveler's diarrhea	NA
Xigris	Drotrecogin alfa	Eli Lilly	11/2001	Severe sepsis	NA
Xolair	Omalizumab	Genentech	6/2003	Moderate to severe persistent asthma	CHO
Zenapax	Daclizumab	Hoffman-La Roche	12/1997	Prophylaxis of acute renal allograft rejection	SP_2/O
Zevalin b	Ibritumoma Tiuxetan	BiogenIdec	2/2002	Relapsed or refractory low-grade, follicular, or transformed B-cell non-Hodgkin's lymphoma	CHO
Zorbtive	Somatropin	Serono	12/2003	Short bowel syndrome	Mouse C127

NA not available; *BHK* baby hamster kidney; *CHO* Chinese hamster ovary

knocked out. Stable cell lines are generated by integrating the genes that encode the desired proteins into the genome of one of the above host cell lines. The CHO-DHFR system uses a plasmid that contains a variety of promoters driving the production of the genes for the proteins and dihydrofolate reductase, which acts as a selectable marker. Amplified expression can be achieved by applying methotrexate, the selective agent for dihydrofolate reductase. The NS0-GS and CHO-GS systems use a vector containing the glutamine synthetase gene, which functions as the selection marker. This type of system can be amplified to boost expression levels by using methionine sulphoximine. In addition to being engineered for selection, host cells have also been genetically constructed by inserting growth factor genes, anti-apoptosis genes, and cell cycle control genes to generate superior production hosts [9].

Expression Systems

The goal of cell line development is to engineer cells to express a large amount of a recombinant protein and to stably maintain this level of production for many cell doublings. The basic schema is:

1. Generate plasmid expressing recombinant protein.
2. Introduce plasmid into cell.
3. Identify cell line expressing high levels of recombinant protein.
4. Select cell line maintaining high levels of expression through scale-up and bioreactor processes.

Plasmids. The plasmid contains all genetic elements necessary for the expression of recombinant protein and for the selection of the cells generating the desired product. For stable cell lines, the plasmids must integrate into transcriptionally active chromosomal regions. The vector promoter elements drive the expression of recombinant protein. The strong cytomegalo virus (CMV) promoter is in general use, although the weaker simian virus 40 (SV40) and rouse sarcoma virus (RSV) promoters are also used. Recently, the strong promoter, CHO-EF1 (CHEF1) [10] has been generating CHO cell lines with high production levels. The inclusion of a chimeric intron in the primary transcript leads to higher stable expression of recombinant protein through enhanced transport and processing of the mRNA from the nucleus into the cytoplasm where it is translated [11, 12].

To allow the selection of cells expressing the desired protein, the plasmid also contains selectable markers such as the *neo* gene generating aminoglycoside 3-phosphotransferase (APH 3 II) for G418 geneticin resistance, the *hph* phosphotransferase gene for hygromycin resistance, the *Sh ble* gene for zeocin resistance, or puromycin N-acetyl-transferase encoded by the *pac* gene for puromycin resistance. Selection can also occur using a complementation gene such as DHFR or GS, which can be included in plasmids when using DHFR-negative or GS-negative cell lines, such as CHO-DG44 and NS0-GS, respectively. The use of an internal ribosomal entry site (IRES) element can facilitate the coexpression of selectable markers and protein product when integrated into the genome [13]. This system generates a single transcript accessible to ribosomes at two locations just prior to the start site of each gene. The fact that the selectable marker and the product gene are under the control of a single promoter, which generates one transcript, is likely to improve cell line stability. Elements such as matrix attached regions (MARs) [14, 15] and ubiquitous

Table 32.2 Expression systems for establishing stable cell lines

Specific productivity				
Prescale-up Expression system	Cell line	(pg/Cell/Day)	Development (months)	Length of Industrial references
Unamplified	CHO	0.5–4	4	Various
Unamplified	NS0	2–14	6	Various
Amplified (GS or DHFR)	CHO	5–15	6 (GS) 9 (DHFR)	Lonza (GS) Various (DHFR)
CHEF1	CHO	6–15	3	ICOS
MARs	CHO	15	2–3	Selexis
UCOE	CHO	4–9	NA	Cobra
Sequential transfection	CHO	10–15	6–8	Xoma, Sunol, ICOS
Retroviral transduction	CHO	17–35*	3	Gala Biotech

NA not available

Data are from Gala Biotech, a wholly owned subsidiary of Cardinal Health, Middleton, WI. Significance of * is "Data are from Gala Biotech, a wholly owned subsidiary of Cardinal Health, Middleton, WI"

chromatin opening elements (UCOEs) [16] can also be included in plasmids, as they are known to generate transcriptionally active genomic environments once integrated into the cell genome.

DNA delivery. Several methods are in use for introducing plasmid DNA into cells. $Ca_2(PO4)_2$ precipitation is the earliest method, but has been surpassed by the more convenient and consistent lipid-base reagents (lipofectamine, fugene, transfectin) and electroporation [17]. Polyethyleneimine (PEI) [18] is a low-cost alternative DNA delivery method. Retroviral transduction has a long history for inserting DNA into cells, but use for manufacturing is only now emerging (Gala Biotech, Middleton, Wisconsin). Transfection efficiencies can vary from 5 to 100% depending on cell line and DNA delivery method. CHO cells can achieve 5–40% with $Ca_2(PO4)_2$, 20–60% with lipids, and 100% with retroviral transduction.

Amplification systems. The DHFR and GS amplification systems have successfully generated manufacturing cell lines with high protein titers (Lonza, Basel, Switzerland) [19, 20]. These systems employ a DHFR or GS cell line that is transfected with the plasmid encoding product of interest along with DHFR or GS, respectively. The use of methotrexate and methionine sulphoximine results in chromosomal amplification events that increase the DHFR or GS gene copy number, respectively, to overcome the drug toxicity. The gene encoding the product of interest is usually coamplified with the DHFR or GS genes as they are inserted into the genome in the same locations. Tenfold or greater improvements in expression can be achieved with this amplification system. The DHFR amplification system has the potential to experience the loss of transgene copy number [21–23], so stability studies are especially important to characterize cell lines derived from drug-induced genomic amplification approaches.

Sequential transfection. A simple approach to boosting expression is to repeat the transfection on previously transfected cells but with selectable markers not used in the first transfection. Xoma (Berkeley, CA), Sunol Molecular (Miramar, FL), and ICOS (Bothell, WA) have used this approach successfully. Fivefold or greater improvements in expression can be achieved in a single sequential transfection. ICOS reports the added advantage of balancing heavy and light chain ratios to improve the secretion and expression of recombinant monoclonal antibodies.

Table 32.2 summarizes expression systems that are currently used in the biotechnology industry for generating stable cell lines. Before one makes a decision as to which expression system is preferred, it is important to evaluate all aspects including productivity and time needed for cell line development, as well as potential license fee and cost.

Identifying High-Expressing Cells

For industrial cell line development, moderate to high-throughput methods are needed to identify the cells expressing the highest levels of the protein pharmaceutical. Typically, the cells are grown in selective or amplifying environments in several 96-well plates. For secreted proteins, enzyme-linked immunosorbant assays (ELISA) on conditioned media can identify the cells expressing the highest protein levels. AlphaScreen™ (Perkin-Elmer, Boston, MA) is a recent homogeneous assay that is well suited for high-throughput quantification of protein production. The Guava Inc. (Hayward, CA) personal cell analyzer (PCA) economically and conveniently generates fluorescence-activated cell sorting-like (FACS) expression profiles of cells with moderate throughput in 96-well microtiter plates. Using Guava PCA, clonal populations and cells with the highest average productivity can be identified.

A critical element of generating stable cell lines is identifying clonal populations of expressing cells. Pools of expressing cells tend to express lower levels of desired protein; they can drift to lower expression levels, and are more difficult to adapt to serum-free suspension. Limited dilution methods have been used for years whereas FACS sorting of live cells has also proven successful. FACS can be used to simultaneously clone and enrich for the highest expressing cells [24]. Recent automated approaches for identifying clones include picking high-expressing colonies in semisolid media (Genetix, New Milton, UK) and enriching for high-expressing cells by laser ablation of low and nonproducing clones (Cyntellect, San Diego, CA).

Cell Banking

A stock of cells must be preserved to ensure continuity for research, development, and manufacturing production programs. For a small research program, only a small number of frozen vials may be needed. However, to continue to supply a cell line for the manufacture of therapeutic proteins, it is usually best to prepare two-tiered cell banks: a master cell bank (MCB) and working cell banks (WCB). A single cell line demonstrating suitable expression levels and stability is used to generate an MCB, and a WCB is derived from one vial of the MCB. Each MCB and WCB usually includes 100–300 vials. As a WCB is depleted during manufacturing runs, another frozen vial of MCB is used to generate a new WCB.

Making cell banks involves the process of cryopreserving cells. During cryopreservation, cells can be damaged by the formation of intracellular ice crystals or by osmotic effects that occur with decreased water content. To minimize cell damage, the rate of freezing must be controlled and cryoprotectants must be used. Dimethylsulfoxide (DMSO) at 7.5–10% is routinely used as a cryoprotectant. In some cases, a low percentage of fetal calf serum or serum albumin is added to the freezing medium. However, when freezing cell lines for the production of therapeutic proteins, it is generally preferable to eliminate any animal-derived material to minimize the risk of disease transmission from animal to human. To control the rate of freezing, Nalgene Cryo freezing containers (Nalge Company, Rochester, NY) are routinely used in making small cell banks, whereas programmable controlled-rate cell freezers are needed for large cell banks.

Cell Stability

The properties of a cell line are likely to change during a long period of continuous passage. For example, cell lines can lose their expression and can generate undesired proteins clipped from the product. Therefore, it is critical to characterize the cells to ensure consistency for large-scale production and to guarantee that the properties of the protein derived from the cells are maintained. For a production cell line, an acceptable level of stability of the desired characteristics must be established and a maximum passage number must be defined so that comparison of the cells' characteristics can be made after low and extensive passages. Tests such as peptide mapping, amino acid sequencing, DNA fingerprinting, and determination of gene copy number and phenotype markers must be conducted to ensure the cells' genetic stability. In addition, productivity and product identity must be examined to assess the stability of the cell line. A good production cell line should be able to maintain its productivity and product quality through the many generations required to reach the end of large-scale production. In most instances, stability retained for 50 generations will satisfy the rigors of large-scale manufacturing.

Media

Mammalian cell culture is the most important source of therapeutic proteins and monoclonal antibodies. Just as mammalian cells are more complicated than most other microorganisms, the media required for their growth is also more complex. The extracellular medium must provide the same nutrients and growth factors that mammalian cells are exposed to in vivo in order for them to survive, proliferate, and differentiate. Serum contains many important components that support the growth of mammalian cells including growth factors, hormones, transport and binding proteins, attachment factors, protease inhibitors, and lipids. Serum was therefore commonly supplemented in the media in early mammalian cell culture and for large-scale production of therapeutic proteins and monoclonal antibodies in the 1980s. However, the use of serum in mammalian cell culture has many disadvantages: (1) it is a potential source of bacterial, mycoplasmal, and viral contamination; (2) it is the most expensive additive to cell culture media; (3) it has a high degree of batch variability, making production consistency difficult; (4) it contains a high concentration of proteins that can interfere with product recovery. In the early 1990s, these drawbacks, especially the serious concern about the risk of transferring diseases from animal to human, led to an important regulatory-driven trend to eliminate serum and animal-derived components from mammalian cell growth media. This trend sparked an industry-wide interest in developing serum-free media.

Commercial Serum-Free Media

Many new medium companies formed in the early 1990s. Today, more than a decade later, the medium development industry has matured. Currently, a variety of serum-free (SF)

Table 32.3 Examples of commercial serum-free media

Media	Type of medium	Cell type	Applied vendor
CD CHO	PF and CD	CHO	Invitrogen
CHO-S-SFM II	SF with low-protein	CHO	Invitrogen
HyQ® CDM4CHO™	CD and ACF	CHO	Hyclone
HyQ SFM4CHO	PF	CHO	Hyclone
HyQ PF CHO	PF	CHO	Hyclone
EX-CELL™ 302 SF	SF with low-protein	CHO	JRH Biosciences
EX-CELL™ 325 PF CHO	PF	CHO	JRH Biosciences
EX-CELL™ CD CHO	CD and ACF	CHO	JRH Biosciences
CHO animal- component-free	ACF	CHO	Sigma-Aldrich
CHO DHFR-medium	ACF	DHFR CHO	Sigma-Aldrich
CHO CD-3 Chemically defined	CD and ACF	CHO	Sigma-Aldrich
BD Cell™ MAb medium (220509)	SF	Sp_2/O, CHO, and myelomas	BD Biosciences
BD Cell™ MAb medium (220513)	ACF	Sp_2/O, CHO, and myelomas	BD Biosciences
IS CHO™ medium	SF	CHO	Irvine Scientific
IS CHO-CD™	CD and ACF	CHO	Irvine Scientific
CD hybridoma	PF and CD	Hybridomas and myelomas	Invitrogen
Hybridoma-SFM	SF with low-protein	Hybridomas and myelomas	Invitrogen
HyQ® CDM4NS0™	CD and ACF	Hybridomas and myelomas	Hyclone
HyQ® SFM4MAb™	SF with low-protein	Hybridomas and myelomas	Hyclone
HyQ® ADCF-MAb™	ACF	Hybridomas and myelomas	Hyclone
EX-CELL™ 620-HSF	SF with low-protein	NS0 hybridomas	JRH Biosciences
EX-CELL™ NS0	PF and CD and ACF	NS0 hybridomas	JRH Biosciences
Hybridoma medium	CD and ACF	Hybridomas and myelomas	Sigma-Aldrich
hybridoma medium	SF and PF	Hybridomas and myelomas	Sigma-Aldrich
IS MAB-V™	ACF	Hybridomas and myelomas	Irvine Scientific
>IS MAB-CD™	CD and ACF	Hybridomas and myelomas	Irvine Scientific

Information is from the vendors' Web sites

media are available commercially. Table 32.3 lists some SF media produced in the United States. Many of these media are also made in powder form to facilitate use in large-scale production. As listed in Table 32.3, there are different types of SF media, which can be categorized into protein-free (PF), chemical-defined (CD), and animal component-free (ACF) media. The relationships among these different types of media are illustrated in Fig. 32.1.

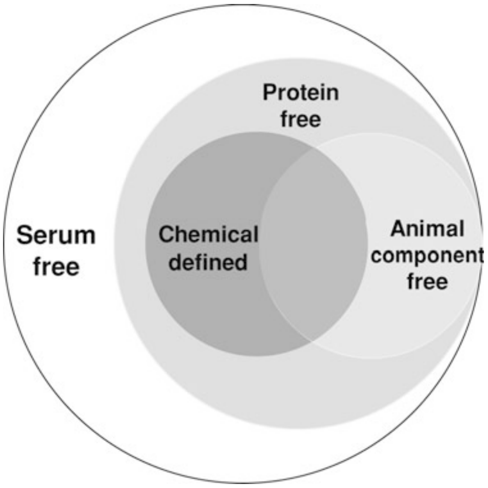


Fig. 32.1 Relationships among serum-free, protein-free, chemical defined and animal component-free media

SF, PF, ACF, and CD media tend to be highly specific to one cell type and sometimes even to one particular cell line. It is not uncommon that a different optimal medium is required for a particular cell line. Developing SF, PF, ACF, or CD media requires considerable experience and expertise and can be very time-consuming. One approach is to start with a commercially available SF medium and add necessary nutrients to optimize growth and production for a particular cell type. This approach can shorten the time-line needed for in-house medium development.

However, using commercial media has two major drawbacks: commercial SF media are expensive, which can lead to a high cost of goods for large-scale production; the composition of a commercial SF medium is proprietary to the medium company and the buyer will not know its formulation. This makes it difficult to gain full knowledge about the growth and production characteristics of your cell line. To fully understand the metabolism and characteristics of your cell line, there is no substitute for developing your own media.

Approaches for Serum-Free Medium Development

Medium development is part of the process of cell culture optimization. It can be very complicated and time consuming, and will require expertise and resources. In this section, we outline the common and novel methods used for SF medium development. Details of each strategy can be found in the referenced literature.

In the early 1980s, efforts were initiated to eliminate serum or animal-derived components from the culture media used to produce human therapeutics. At that time, different strategies were recommended for the development of SF media:

1. Limiting factor method: [25] Starting with an existing formulation, the serum concentration is lowered until cell growth becomes limited; then the concentration of each component of the medium is optimized until cell growth recovers.
2. Synthetic method: [26] A variety of growth factors is added to the existing basal media to replace the serum's functions.

The last decade saw the development of many new genomic and automated screening tools. These advancements, as well as an improved understanding of mammalian cell culture, allowed novel concepts and approaches to be applied to the development of SF and CD media. Some representative strategies are briefly summarized below.

1. Rational design: [27] This method uses four complementary methods, including component titration (CT), media blending (MB), spent media analysis (SMA), and automated screening (AS), to achieve the best SF culture medium in the shortest timeline. An example of how to use this strategy is outlined in Fig. 32.2.
2. Factorial design: [28–30] This method identifies key growth factors in a lean basal medium by performing experiments using a full factorial matrix. It allows calculating the maximum number of interaction effects and gives insights into growth factor biology.
3. Genomic tools: [31] This method takes advantage of select genomic (microarrayor PCR-based approaches) and proteomic (antibody array analysis) tools to identify the receptors for growth factors, hormones, cytokines, and other components of cell signaling pathways expressed by a culture of interest.

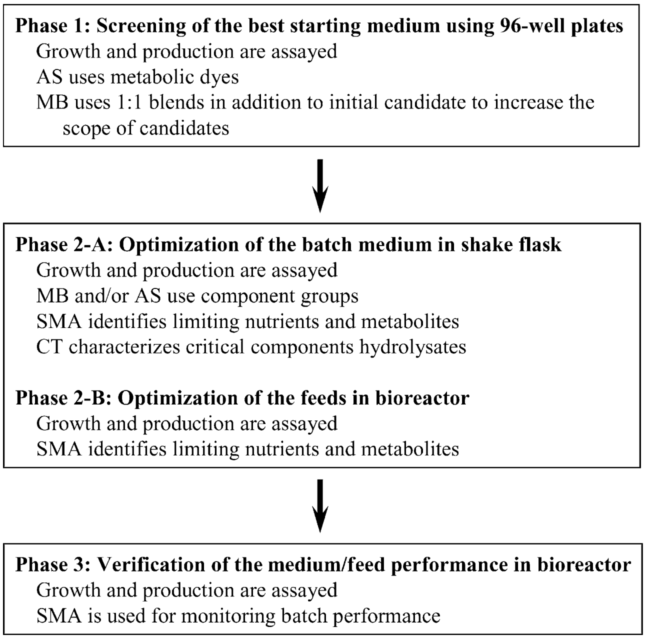


Fig. 32.2 Procedure of using rational design

Serum-Free Adaptation

In most cases, cells are grown in a serum-containing environment during the early stages of cell line development, such as transfection and selection. Once an SF medium with a good nutrient balance is chosen, the next step is to adapt cells to SF growth. Two adaptation strategies, sequential adaptation and starve/save adaptation, are described.

Sequential Adaptation. This method weans cells off serum gradually (e.g., from 10 to 5% to 2.5 to 1.25% to 0.5 to 0.1%) until the serum is completely removed. This strategy is conservative and easy to follow, and therefore commonly used in the biotechnology industry. A drawback is that it takes a relatively long time (up to 6 months) before achieving a full SF adaptation. A typical protocol includes these steps:

1. Grow the cells in a targeted SF medium containing 10% serum and split the cells until the same doubling time is achieved as in the old serum-containing medium.
2. Reduce the serum concentration to 5% and continue to split the cells until a relatively consistent doubling time is achieved. Reduce the serum concentration to 2.5% and repeat the process.
3. Transfer the cells to the medium containing 1.25% serum. At this level of serum, it will be more difficult for the cells to adapt. The cells may become stressed and show diminished survival. It is therefore wise to maintain the cells at the previous serum concentration while spliting them to a lower serum concentration.
4. Continue to reduce the serum to 0.5, 0.25, 0.1, and 0% by repeating step 3. At the low serum concentrations, seeding the cells at a high cell density (5×10^5 cells/mL) will facilitate adaptation. This can be achieved by centrifuging the cells out from the medium containing the current serum concentration and resuspending them into a

smaller volume of medium containing a lower serum concentration.

5. Once the cells are able to survive and grow in an SF medium with a high seeding density, split them several times at a lower density (2 × 10^5 cells/mL) to ensure that the cells are truly adapted.

Starve and save adaptation. This method uses an SF medium to starve the cells of components contained in serum, thereby selecting for adaptable cells. Switching to a serum-containing medium then saves the cells that became stressed during starvation. This process is repeated allowing selection and adaptation to continue until full SF adaptation is achieved. The advantage of this method is its effectiveness in selection, so it usually only takes 4–6 weeks to achieve the goal [32]. The following steps are used.

1. Define an SF medium and a serum-containing medium that have exactly the same formulation except for the serum.
2. Grow cells in the serum-containing medium to reach a high cell density.
3. Split and grow the cells in the SF medium, monitoring changes in cell density and viability.
4. When cell viability drops to about 50%, save the cells by transferring them back to the serum-containing medium.
5. Switch back to the SF medium to further select the cells and repeat steps 3–5 until the cells grow successfully in the SF medium.

Bioreactor Systems

Many therapeutic proteins are produced using genetically modified mammalian cells, as described in the preceding sections. This section describes the basic design and function of bioreactors used for suspended mammalian cell culture. Bioreactors should provide a sterile environment, adequate mixing, ease of operation, and control of temperature, pH, and dissolved oxygen. Traditionally, these requirements were met using glass or stainless steel stirred-tank systems. At production scale, therapeutic proteins are primarily produced in stirred-tank bioreactors. However, single-use, disposable systems such as the Wave bioreactor are gaining acceptance for certain applications.

Stirred-Tank Systems

Stirred-tank bioreactors are generally glass or stainless steel tanks with an impeller to provide mixing. Air or oxygen is usually bubbled through the media to supply oxygen to the cells. An example of a stirred-tank bioreactor is depicted in Fig. 32.3. Sterility is obviously an important issue in mammalian cell culture, and therefore stirred-tank systems require extensive cleaning and autoclaving. A significant portion of the labor involved with operating these systems is related to cleaning and sterilizing.

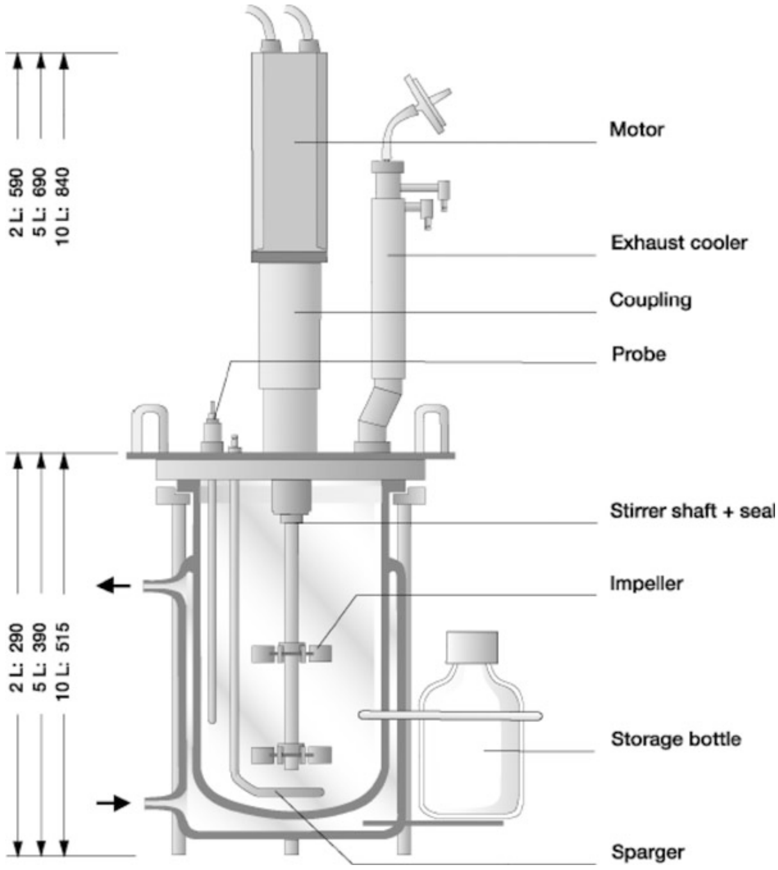


Fig. 32.3 Standard bench-scale, water-jacketed stirred tank bioreactor (Permission of Sartorius BBI Systems Inc.)

Mixing in stirred-tank bioreactors. Adequate mixing is essential to suspend the cells and to facilitate heat and mass transfer. Historically, however, due to concerns regarding mammalian cells' sensitivity to hydrodynamic stress, most stirred-tank bioreactors were agitated just enough to keep the cells in suspension [33]. This low level of mixing can result in large concentration gradients of pH, oxygen, and other nutrients. Ozturk demonstrated this experimentally by adding base to a poorly mixed bioreactor [34]. The base was added to the top of the bioreactor, and because of poor mixing, a high pH region was created at the top. The cells in this region ruptured, forming a "snowball" of cellular debris in the vicinity of the base inlet.

Proper impeller selection and sizing will improve mixing. Generally, the diameter of the impeller should be approximately one-third to one-half of the tank diameter. The lowest impeller should produce a radial flow pattern to aid in gas dispersion, and the upper impeller(s) should produce axial flow to eliminate "zones of mixing." Paddle or Rushton impellers produce radial flow, and hydrofoils and pitched blade impellers produce axial flow. Retrofitting existing bioreactors with different impellers, however, may prove difficult because the motor driving the impellers may not produce enough torque to turn them. Adequate baffling will also improve mixing; baffles prevent solid body rotation and vortex formation [35].

Aeration. Oxygen can be introduced to the culture in many different ways. Membrane aeration provides efficient oxygen transfer with minimum shear damage to the mammalian cells and minimal foaming. However, due to the design complexity and the difficulty involved in cleaning and sterilizing membrane reactors, membrane aeration has limited utility in large-scale bioreactor systems. Sparger aeration offers high oxygen transfer rates and is widely used in both bench and production-scale bioreactors. Because oxygen is only sparingly soluble in water, a large surface area is needed to maximize diffusion of oxygen into the cell culture media. Frit spargers with micropores provide a large surface area for diffusion, but this type of sparger can cause foaming problems at large scale. Traditional large-hole-ring spargers reduce foaming and are therefore often used in production-scale stirred-tank bioreactors [36].

Biosensors. Sensors are required to adequately monitor bioreactor performance. Ideally, one would like to have online sensors to minimize the number of samples to be taken from the bioreactor and to automate the bioreactor process. Most bioreactors have autoclavable pH and dissolved oxygen (DO) electrodes as online sensors, and use offline detectors to measure other critical parameters such as glucose and glutamine concentration, cell density, and carbon dioxide partial pressure (pCO_2). An online fiber-optic-based pCO_2 sensor is commercially available and appears to be robust [37]. Probes are also commercially available that determine viable cell density by measuring the capacitance of a cell suspension. Data from perfusion and batch cultures indicate that these probes are reasonably accurate at cell concentrations greater than 0.5×10^6 cells/mL [38, 39].

All of the aforementioned biosensors are designed for a standard stirred-tank bioreactor; they will not work with small-scale vessels such as shake flasks. Small noninvasive sensors have been developed to measure DO and pCO_2 inside shake and T-flasks [40, 41]. A DO or pCO_2 detecting "patch" is placed in the flask and this patch contains an oxygen or carbon dioxide luminescent dye. The color shifts of these dyes are detected using an external fluorometer.

Disposable Bioreactors

One of the major drawbacks of the stirred-tank bioreactors is the difficulty of cleaning and sterilizing the vessels. Disposable bioreactors are now commercially available that are based on the use of presterile plastic bags. The sterile bag is partially filled with cell culture media and is then pressurized with a mixture of carbon dioxide and air. The bag is placed on a platform that rocks back and forth creating waves inside the bag. The wave motion provides sufficient mixing and gas transfer to easily support cell growth of over 20×10^6 cells/mL [42]. Currently, wave bags are available in sizes up to 500 L culture volume and have been proven for the good manufacturing practice (GMP) production of human therapeutics. Wave bioreactors have been primarily used as batch culture for inoculum preparation and transient production; however, internal perfusion filters can be applied, allowing the option of use for high-density perfusion culture.

Modes of Bioreactor Operation

Industrial mammalian cell culture can be divided into three primary modes of operation: batch, fed-batch, and perfusion. Each mode of operation has its attributes and drawbacks, which are described in detail below.

Batch. Batch culture is the simplest of the three modes to operate. Cells are inoculated into media, and with the exception of agitation, temperature, pH, and DO control, essentially left alone for 3–5 days; no additional nutrients or media are added. Batch processes are easy to operate and require the least optimization effort of the three modes. However, cell densities and protein titer are significantly lower for batch processes, typically less than 5 × 10^6 cells/mL and 0.5 g/L, respectively [43]. Also, some protein products are degraded in the media during the batch process.

Fed-Batch. Fed-batch processes start out as batch cultures; after a few days of growth—when a crucial nutrient is depleted—a concentrated solution of nutrients is added to the media. Fed-batch cultures persist for one to two weeks and may produce high cell density and product titers, typically greater than 10 × 10^6 cells/mL and 1.5 g/L, respectively [43].However, one needs to optimize the contents of the feed solution as well as the feeding strategy. Similar to batch cultures, fragile proteins may be degraded during the course of the culture.

Perfusion. Perfusion cultures can be considered continuous cultures. Broth is continually removed from the bioreactor, the cells are separated using a cell retention device, the cells are returned to the bioreactor, and the supernatant is collected for later protein purification. There is a variety of cell retention devices available, including spin filters, acoustic separators, continuous centrifuges, and gravity settlers. Spin filters and, to a lesser extent, acoustic settlers are the two most commonly used cell retention devices for large-scale perfusion cultures [44]. Perfusion cultures usually last many weeks, but require a longer time for process optimization, and more effort in bioreactor operation than either batch or fed-batch cultures. The increased effort is primarily due to continuous feeding/harvesting and fouling of the cell retention devices. The cell densities attained in perfusion cultures are usually on the order of 50 × 10^6 cells/mL, and the product titer, in most cases, is lower than that from fed-batch culture. However, the volumetric productivity is typically ten times that of fed-batch cultures. Because broth is continually removed from the bioreactor, fragile proteins can be separated immediately from cell proteases and other components that can cause degradation of the product.

Cell Culture Process and Control

Mammalian cell culture processes must be tightly controlled to attain acceptable cell density and maximize product titer. Slight deviations in pH, temperature, nutrient, or catabolite concentrations can cause irreparable damage to the cells. This section covers the effects of pH, shear stress, catabolite, and carbon dioxide accumulation on cell growth and product formation, and discusses the importance of controlling glucose and glutamine concentrations in fed-batch and perfusion cultures. A brief discussion of scale-up heuristics in mammalian cell culture is also included.

Table 32.4 General control parameter set points for mammalian cell culture

Control parameter	Optimal range
Temperature	37°C
pH	7.0–7.4
Osmolality	280–320 mOsm/kg
O_2 demand	0.5–2 mM/h

Process Parameters

General parameter values. The optimal pH range for mammalian cell growth is 7.0–7.4, which is typically maintained using carbon dioxide and sodium bicarbonate. The optimal osmolality of the media is between 280 and 320 mOsm/kg. In general, mammalian cells achieve high cell densities most quickly with the temperature set to 37°C; however, a lower temperature may be advantageous in some cases for extending cell life, thereby increasing product titer. The oxygen demand for mammalian cells is 0.5–2 mM/h and is one of the more critical control parameters in cell culture. The optimal ranges of common control parameters in cell culture are summarized in Table 32.4.

Mitigating Effects of Physical and Chemical Stress

Shear Stress. Because mammalian cells lack a cell wall and are larger than bacteria, they are more susceptible to hydrodynamic forces, or shear stress. Several studies have investigated the effects of shear stress on mammalian cells [45–48]. Many indicate that the action of the impeller alone does not decrease the viability of suspension-adapted mammalian cells [46, 48, 49]. Some bioprocess engineers in industry have seen a few cell lines that appear to be less robust, and anecdotally might have been damaged by the impeller. However, bubble rupture does cause sufficient hydrodynamic force to kill all the cells attached to the bubble [48]. The effects of bubble rupture can be greatly reduced by the addition of surfactants, such as Pluronic F-68. Pluronic F-68 renders cell-bubble adhesion thermodynamically unfavorable, so the cells do not adhere to bubbles [50]. However, Pluronic F-68 offers very little shear protection; cells rupture at the same level of hydrodynamic force regardless of the Pluronic F-68 concentration [51]. Cells attached to microcarriers, however, are very susceptible to shear stress and can quite easily be removed and killed by the action of the impeller [47].

pH Perturbations. As mentioned earlier, significant pH gradients within the bioreactor are common, due to

Table 32.5 Effects of shear force, pH deviation, and accumulation of lactate, ammonia, and carbon dioxide on cell culture

Stress	Cause	Effects	Possible solution
Shear force	Bubble rupture, impeller	Decreased viability	Add surfactants such as Pluronic F-68
pH deviations	Inadequate mixing	Decreased growth rate, viability	Increase agitation/aeration rate; optimize vessel configuration
Lactate accumulation	High glucose concentration	Decreased growth rate, increased glucose/glutamine consumption	Optimize feed strategy/media to reduce (glucose)
Ammonia accumulation	Metabolism/deamination of glutamine	Apoptosis, decreased growth rate, altered glycosylation	Optimize feed strategy/media to reduce (glutamine); use cell line with GS system
CO_2 accumulation	Cellular respiration, sparged gas component	Decreased growth and protein production	Increase agitation/aeration rate; optimize vessel configuration

inadequate mixing. Cell lysis occurs at extreme pH; however, even moderate deviations from the optimal pH may be detrimental. Osman and coworkers found that pH values greater than 8.0 or less than 7.0 cause a considerable decrease in cell viability and a reduced culture time [52]. Antibody titers increased when the pH set point was reduced from 7.2–7.3 to 7.0 [52, 53]. The best way to eliminate deviations from optimal pH is to improve mixing (reduce the mixing time); this can be achieved by increasing the agitation or aeration rate, adding baffles, or optimizing the impeller design or placement.

Catabolite accumulation. Several byproducts of cellular metabolism accumulate during the course of a bioreactor run. Many of these catabolites, such as lactate, ammonia, and carbon dioxide, are detrimental to cell growth and protein production. Lactate is a product of glycolysis and lactate accumulation greater than 2 g/L tends to inhibit cell growth and increase glucose and glutamine consumption [54, 55]. Lao and Toth pointed out the difficulty in completely decoupling the effects of lactate accumulation from the effects of increased osmolality (osmolality increases with increasing lactate concentration) [54]. The effects of lactate accumulation can be mitigated by keeping glucose levels low, which can be achieved by optimizing media composition and feeding strategies.

In cell culture, ammonia is produced as a cellular metabolite and is converted from glutamine in the media through deamination. Less than 20 mM ammonia inhibits cell growth, induces apoptosis, and alters glycosylation in certain cell lines [56–58]. The primary method of reducing ammonia concentrations in fed-batch and perfusion cultures is to optimize feeding strategies. In addition, cell lines containing the glutamine synthesis gene (GS systems) can be used to eliminate ammonia production; in GS systems, the cells are able to synthesize glutamine based on need. A recent report shows that substituting glutamine with pyruvate can greatly reduce ammonia production [57].

Carbon dioxide is a product of cellular respiration. In mammalian cell culture, carbon dioxide and sodium bicarbonate are normally used to control bioreactor pH. Elevated partial pressure of carbon dioxide (pCO_2) hinders cell growth and protein production [59–61]. As with lactate accumulation, the effect of elevated pCO_2 is difficult to completely decouple from the effect of elevated osmolality [60, 61]. Generally, one can reduce pCO_2 by increasing the volumetric mass-transfer coefficient (kLa), typically by increasing the sparge rate and/or the agitation rate. Mostafa and Gu were able to reduce pCO_2 in a 1,000 L culture and nearly double the titer by increasing the sparge rate and using an open pipe instead of a sparger [62]. Table 32.5 summarizes the effects of shear force, pH deviation, and accumulation of lactate, ammonia, and carbon dioxide on cell culture and lists methods to minimize these effects.

Temperature shifts. Culture temperature is one of the primary control parameters in mammalian cell culture. As stated previously, most mammalian cells grow optimally at 37°C. However, reducing the temperature set point slows cellular growth rate and metabolism, extends the period of high cell viability, arrests cells in the G0/G1 phase of the cycle, and possibly reduces intracellular protease activity [63–66]. By decreasing the temperature to 31–33°C, the period with high cell viability is extended for several days, leading to higher product titers. Fox et al. [65] and Bollati-Fogolin et al. [66] reaped the benefits of a cooler temperature without excessively extending culture time by shifting from 37 to 32 or 33°C, respectively. The temperature shift occurred toward the end of the cells' exponential growth phase, approximately 3–4 days after inoculation in a batch culture. Different cell lines and culture conditions may have different optimal time points for the temperature shift. However, finding the optimal time point is important and worth investing the time required. Operating cell cultures at a reduced temperature will most likely increase production of the target protein, but to avoid substantially increasing the culture time, the culture is generally started at 37°C, and then decreased to 31–33°C at a predetermined time point.

Fed-Batch Process Control and Optimization

In general, fed-batch cultures are initially operated in batch mode; when a key nutrient(s) is exhausted, a solution containing the nutrient(s) is added to the media. As mentioned earlier, the feed solution and feeding strategy should be optimized. A common approach is to use partial concentrates (i.e., 10×) of the basal media as the feed solution; typically most of the salts found in the basal media are not included in the feed solution. The bioreactor is sampled at regular intervals and when one or more of the key nutrients (usually glucose and/or glutamine) is below a certain concentration, a precise amount of feed solution is added to raise the nutrient concentration up to its set point. A useful rule for choosing the nutrient set point concentrations is to provide enough nutrients to support cell growth and production while avoiding the formation of toxic levels of catabolites due to excess feeding. Another method is to determine which of the media's several nutrients become depleted during the culture and add those particular nutrients to the bioreactor independently [67]. Wong et al. selected glutamine and glucose as key control nutrients, and used two feed solutions (glucose and glutamine solutions) to maintain them at specified concentrations [68]. Using this strategy, they maximized cell viability and density while decreasing accumulation of lactate and other catabolites.

Perfusion Process Control and Optimization

In perfusion bioreactors, supernatant is removed from the bioreactor at certain times, the cells are separated from the supernatant, the supernatant is harvested, and the cells are returned to the bioreactor. Perfusion bioreactors can be operated in a variety of modes. The simplest mode is to consistently remove a certain amount of broth each day (i.e., one bioreactor volume/day) and replace with fresh media. This mode is relatively easy to control. However, as the cell density increases, the required nutrient level may not be met. Also, the protein product tends to become diluted in the collected supernatant [34, 38]. Another mode of perfusion operation is to remove relatively small volumes of media at a time and replace the volume with a concentrated solution of nutrients. This mode is similar to fed-batch bioreactors. One might think of this as a perfusion/fed-batch hybrid, and as with fed-batch bioreactors, considerable effort goes into determining the optimal nutrient set point concentrations and the feeding strategy. Sophisticated analysis and control schemes have been developed for this type of perfusion bioreactor [69, 70]. A third mode of perfusion bioreactor operation attempts to maintain a pseudo-steady-state cell concentration, after an initial growth period. Dowd and coworkers employed an online cell density meter and were able to maintain a relatively constant cell concentration by altering the perfusion rate, the amount of media removed from the bioreactor [38]. This mode of perfusion reduces the frequency of sampling and analysis required to uphold a set nutrient concentration, but requires a well-characterized online cell density probe and a well-calibrated pump control scheme.

Scale-Up of Mammalian Cell Bioreactors

When scaling up a process to large scale, it is important to maintain the same physical and chemical conditions as in small scale. The chemical conditions include pH, oxygen level, concentration of medium components, and concentrations of toxic metabolites; these must be monitored and controlled to keep the cells in the proper physiological environment. The physical conditions include the bioreactor configuration and the power provided to the bioreactor. In scale-up, it is critical to preserve a similar geometrical configuration of the bioreactor in order to facilitate duplication of mixing patterns.

Impellers are an important physical component in a stirred bioreactor; they convert electronic energy to hydrodynamic motion and generate the turbulence required to keep the cells in suspension and achieve good mass transfer. At large scale, efficient oxygen delivery and carbon dioxide stripping become increasingly difficult due to poor mixing. At the same time, the energy generated at the tip of the impeller blades must be limited, as certain cell lines can be damaged by the elevated shear force. Therefore, to achieve good mixing and minimize possible damage of the cells from high shear force, one must determine the proper impeller shape, ratio of impeller to vessel diameter, and impeller tip speed.

The mixing time as well as the oxygen and carbon dioxide mass transfer rates can be correlated to the power per unit volume of the reactor, also known as the average or overall energy dissipation rate [71]. Maintaining constant power per unit volume is a commonly used scale-up strategy because of its simplicity. Other strategies include keeping the average shear force experienced by the cells constant, or keeping the maximal shear force constant (the shear force experienced by the cells when passing the impeller tip).

In addition to scale-up models, scale-down models are widely used to establish the operating ranges of critical large-scale process variables [72, 73]. Conducting the many experiments required to define a validated range for

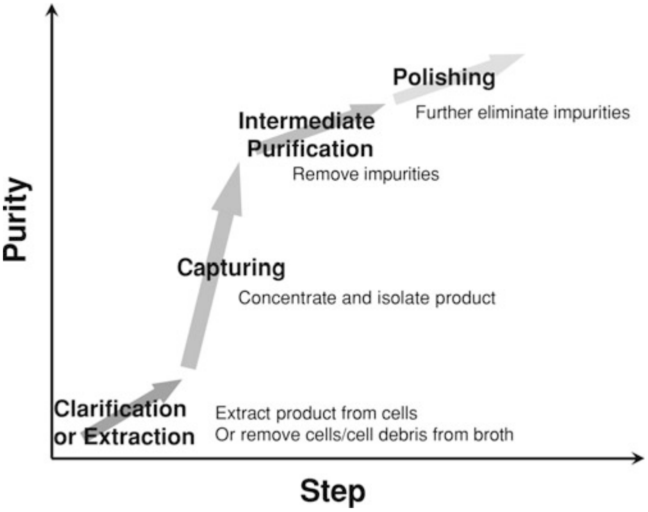


Fig. 32.4 General strategies for protein purification (Permission of Amersham Biosciences, GE Healthcare)

each parameter is not feasible at large scale. A scale-down model based on an existing large-scale process therefore becomes an efficient and economic tool in reaching this goal.

Purification Process

The process of purification, also called "downstream processing," depends on the product and the degree of purification required [74]. Current strategies used for purification of therapeutic proteins generally involve these steps: (1) sample preparation (clarification or extraction), (2) product capture (product concentration), (3) intermediate purification (removal of bulk impurities), and (4) polishing (removal of trace impurities) as shown in Fig. 32.4.

Microbial cells such as *E. coli* and yeast are limited in their ability to make glycoproteins or correctly structured glycoproteins. Therefore, therapeutic proteins (including monoclonal antibodies) that require glycosylation for their biological activity are often produced via cultivation of mammalian cells. Compared to microbial fermentation, mammalian cell culture generates complex broths with a variety of impurities. Typical impurities found in culture broths include:

1. Host cells and cell debris
2. Host cell protein (HCP) and DNA released by the cells
3. Aggregated proteins or cleaved proteins produced by the cells
4. Medium additives (such as serum and protein used to support cell growth)

These impurities pose risks for the safety of proteins used as therapeutics and must be removed to a final concentration below their target limit. In addition, the product stream contacts materials such as filters and resins. Extractables, such as leachates from protein A resins, can pose an immunogenic risk to the patient and must be eliminated [75]. Finally, adventitious agents such as viruses and bacterial pathogens or related contaminants such as endotoxins can lead to serious problems with the safety of the protein preparation and therefore must be minimized. Table 32.6 lists concentrations for the above impurities that are generally considered acceptable in a final protein product [76].

Table 32.6 Concentrations of impurities generally acceptable in final protein product

Impurity	Concentration in final product
Cell debris	0
Host cell protein and serum	<0.1–10 mg/L
Aggregated/cleaved proteins	Various
DNA	10 ng/dose
Leachates	0.1–10 mg/L
Virus pathogens	$<10^{-6}$/dose
Endotoxins	0.25 EU/mL

This section describes some general processes used for protein purification, including methods and tools currently employed by the bioseparation industry to achieve clarification, capture, and removal of impurities.

Generic Processes

As shown in Fig. 32.3, each purification step has different goals. The purification problems associated with a particular step will greatly depend on the properties of the starting material. Thus, the goal of a purification step will vary according to its position in the process.

Chromatography is one of the most important tools in protein purification. Chromatographic purification techniques include affinity chromatography (AC), ion exchange chromatography (IEX), hydrophobic interaction chromatography (HIC), and gel filtration (GF). These techniques separate proteins according to differences in specific protein properties. The protein property used for separation, the attributes of each technique, and its suitability for different purification steps are summarized in Table 32.7.

Selection of the purification strategy will depend on the specific properties of the sample and the required level of purification. Due to the widely differing properties of proteins, a final purification strategy that is most suitable for one protein may be unsuitable for another. A logical combination of chromatographic steps can usually achieve the final goal of protein purification. Figure 32.5 shows two flow diagrams commonly used for the purification of proteins expressed by mammalian cells.

Table 32.7 Attributes and purification suitability of chromatographic techniques (Permission of Amersham Biosciences, GE Healthcare)

Technique	Protein property	Main attributes	Capture	Intermediate Purification	Polishing
Affinity chromatography (AC)	Ligand specificity	High capacity High resolution High speed	+ + +	+ + +	+ +
Ion exchange chromatography (IEX)	Charge	High capacity High resolution High speed	+ + +	+ + +	+ + +
Hydrophobic interaction chromatography (HIC)	Hydrophobicity	High capacity High resolution High speed	+ +	+ + +	+
Gel filtration (GF)	Size and shape	Low capacity High resolution	–	+	+ + +

+ + + = highly suitable; + + = very suitable; + = suitable; – = not suitable

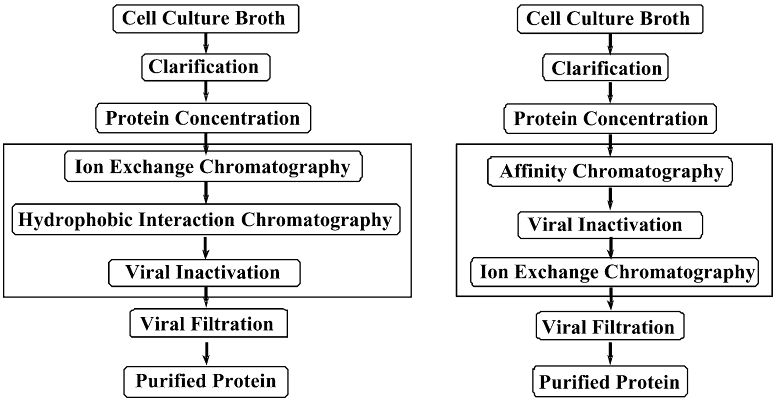


Fig. 32.5 Flow diagrams for protein purification

Clarification of Broth

For proteins secreted into broth, the first objectives after cell culture are to remove cell mass and debris, to reduce processing volume, and to bring the product to a stable holding point for further purification steps.

At bench-scale purification, a two-step operation of depth filtration followed by sterile filtration (or centrifugation followed by sterile filtration) can efficiently remove large particles, colloidal particles, and bacteria. At industrial scale, however, the clarification step is usually completed by three stages in series as shown in Fig. 32.6.

Centrifuges can function in batch-mode and continuous mode. Batch-mode works well for laboratory-scale processes, whereas continuous-mode can handle larger volumes and is thus conducive to large-scale processes. Both depth and tangential flow filtration are commonly used in the clarification. For large-scale purifications, depth filters may have volume and throughput limitations. Tangential flow filtration becomes a better choice when processing large harvest volumes.

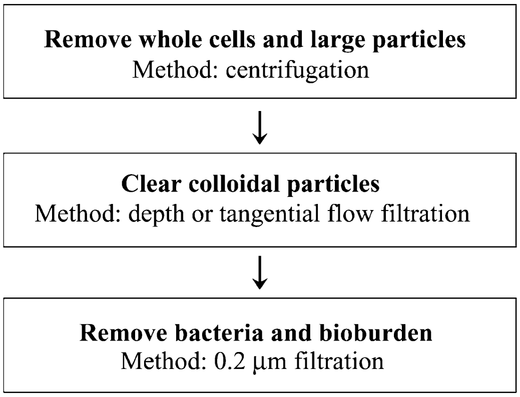


Fig. 32.6 Procedures for broth clarification at large scale

Capture of Product

The goal in the capture step is to concentrate and isolate a protein product. This step is usually accomplished with ion exchange or affinity chromatography. Ion exchange separates protein molecules based on interactions between charged side chains on the protein surface and oppositely charged groups, such as ammonium or sulfate, covalently linked to the chromatography matrix. The charge characteristics of a protein can be altered by changing the pH of the separation. Based on the charge differences of a targeted protein and other impurities, conditions at the capture step are selected to maximally bind the targeted protein and minimally bind the impurities to achieve isolation of the product. Ion exchange is probably the most frequently used chromatographic technique for the separation and purification of proteins. It offers many advantages such as high capacity, long resin lifetime, and low cost. Its operation is relatively simple and easy to control. A major disadvantage is that it usually only leads to moderate purity [77].

Affinity chromatography separates proteins based on a reversible interaction between a protein and a specific ligand coupled to a chromatography matrix. One of the most important attributes of this method is its excellent selectivity. As a result, affinity purification offers immense time savings over less selective multistep procedures. A single-step operation with affinity chromatography can achieve purification levels on the order of several thousandfold with high yield. The most successful example of affinity chromatography in protein purification is the use of Protein A to capture monoclonal antibodies. One-step purification with Protein A not only results in high yields of targeted antibodies, but also removes bulk contaminants such as host cell protein, DNA, and virus, leading to purity greater than 95%. Another advantage of using Protein A is that it works for multiple antibodies. Major disadvantages are: (1) high cost, (2) cleaning and sanitization difficulties, and (3) ligand leakage.

There are many different types of Protein A affinity media: MabSelect™ (GE Healthcare, Chalfont St. Giles, UK) and PROSEP Ultra (Millipore, Billerica, MA) are the two most commonly used at production scale. Recently, GE Healthcare released a new type of Protein A called MabSelect SuRe™. This latest MabSelect resin is resistant to deterioration from cleaning and sanitization-in-place with 0.1–0.5 NaOH.

Removal of Impurities

Although the capture step dramatically enriches a targeted protein and removes some impurities, bulk impurities such as host cell protein, DNA, endotoxin, virus, and leaching ligand remain in the eluted pool. Additional procedures are needed to eliminate these impurities.

Depending on the required level of purity, the step used for removing impurities can be a single or multiple-step chromatography operation. Table 32.8 summarizes chromatographic methods that can be used to remove impurities following the capture step [76, 78–80].

Table 32.8 Chromatography methods used to remove impurities

Impurity	Methods used to remove impurity
Host cell protein	AEX, CEX, HIC, AC
Aggregated/cleaved proteins	CEX, HIC
DNA	AEX, CEX, HIC, AC
Leachates	AEX, CEX, HIC
Virus pathogens	AEX, CEX, HIC, AC
Endotoxins	AEX, CXE, AC

For monoclonal antibody purification in which Protein A is primarily used for protein capture, cation and anion exchange chromatography are employed to reduce the remaining impurities [81]. Because the majority of monoclonal antibodies have a pI greater than 7 and predominating impurities (HCP and protein A leachates) have pIs lower than 7, cation exchange can retain the product and allow the impurities to flow through. Anion exchange can be employed as a noproduct-binding step to remove residual DNA and endotoxin. In this mode, the trace impurities will be retained on the column and the antibody will flow through. This step typically clears 2–5 logs endotoxin and 3–5 logs DNA [76].

Separation by HIC is based on the reversible interaction between a protein and the hydrophobic surface of chromatographic medium. This interaction is enhanced by high ionic strength buffers. Thus, HIC usually follows an ion exchange step where a high salt buffer is used for elution. It can be used to remove impurities such as host cell protein, DNA, and virus. However, at production scale, it is not as widely used as IEX.

Gel filtration separates proteins based on size and shape and can also be applied to remove the trace impurities. However, because this method is usually very slow, its application is often limited to bench or pilot scale.

In addition to chromatography, membrane filters have also proven to be effective in removing virus and endotoxin. For example, application of Sartobind™ membrane adsorbers made by Sartorius (Goettingen, Germany) results in up to 5 logs reduction of endotoxin. The Ultipor® VR grade DV50 membrane by Pall Corporation (East Hills, NY) demonstrated a greater than 7 log clearance in removal of retroviruses [82].

Strategies for Scaling Up Purification Processes

The great demand for high-quality therapeutic proteins requires efficient manufacturing processes, both upstream and downstream. When bioreactors are scaled up from bench to pilot to production scale (10,000–20,000 L), downstream processes must be scaled up accordingly. Equally complicated as upstream processes, scale-up of downstream processes requires more than simply increasing the size and volume of the laboratory equipment. The use of larger piping dimensions, larger filters, different types of pumps, and larger dead volumes can introduce variation to the large-scale process. We've limited our description here to common strategies used by the industry to scale up chromatographic columns.

Scale-up of affinity chromatography [78]

1. Prior to scale-up, different protein A resins should be screened. To achieve high production rate, defined as the amount of protein purified per unit time and per unit column volume, a resin must have a high dynamic binding capacity and be able to operate at low back-pressures.
2. Before scale-up, the column volume based on the required binding capacity must be calculated and the optimum bed height that allows high flow rates and high dynamic capacities must be determined.
3. During scale-up, the residence time, which is equal to the bed height divided by the linear velocity of the fluid, must be kept constant.
4. The operation flow rate should not exceed 70% of the maximum flow rate specified by the resin's vendor.

Scale-up of ion exchange column [78].

1. Prior to scale-up, screen resins to obtain those that offer a long lifetime, lot-to-lot consistency, and long-term availability.
2. During scale-up, the column volume should be increased by keeping the height of the resin bed constant and increasing the column diameter.
3. During scale-up, the total load of protein per unit of resin should be the same and the linear flow rate should be identical.
4. Scale-up usually involves two steps: the first step results in an increase of 50- to 100-fold, from laboratory to pilot scale; the second step results in an increase of 10- to 50-fold, from pilot scale to full production scale [83].

Acknowledgements The authors would like thank all the colleagues who helped in many invaluable ways in the production of this chapter, in particular, Marie Ary, Kenton Abel, Bassil Dahiyat, Joyce Morrison, and Christopher O'Brien.

References

1. Walsh G (2003) Biopharmaceutical benchmarks—2003. Nat Biotechnol 21:865–870
2. PhRMA (2004) 324 biotechnology medicines in testing promise to bolster the arsenal against disease. New medicines in development. Accessed Oct 25, 2004. http://www.phrma.org/newmedicines/biotech/
3. Graddis TJ, Remmele RL Jr, McGrew JT (2002) Designing proteins that work using recombinant technologies. Curr Pharm Biotechnol 3:285–297
4. Brekke OH, Loset GA (2003) New technologies in therapeutic antibody development. Curr Opin Pharmacol 3:544–550
5. Lazar GA, Marshall SA, Plecs JJ, Mayo SL, Desjarlais JR (2003) Designing proteins for therapeutic applications. Curr Opin Struct Biol 13:513–518
6. Marshall SA, Lazar GA, Chirino AJ, Desjarlais JR (2003) Rational design and engineering of therapeutic proteins. Drug Discov Today 8:212–221
7. Vasserot AP, Dickinson CD, Tang Y, Huse WD, Manchester KS, Watkins JD (2003) Optimization of protein therapeutics by directed evolution. Drug Discov Today 8:118–126
8. Wurm FM (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. Nat Biotechnol 22:1393–1398
9. Arden N, Nivtchanyong T, Betenbaugh MJ (2004) Cell engineering blocks stress and improves biotherapeutic production. Bioprocessing 3:23–28
10. Running Deer J, Allison DS (2004) High-level expression of proteins in mammalian cells using transcription regulatory sequences from the Chinese hamster EF-1alpha gene. Biotechnol Prog 20:880–889
11. Brinster RL, Allen JM, Behringer RR, Gelinas RE, Palmiter RD (1988) Introns increase transcriptional efficiency in transgenic mice. Proc Natl Acad Sci U S A 85:836–840
12. Palmiter RD, Sandgren EP, Avarbock MR, Allen DD, Brinster RL (1991) Heterologous introns can enhance expression of transgenes in mice. Proc Natl Acad Sci U S A 88:478–482
13. Pestova TV, Lomakin IB, Hellen CU (2004) Position of the CrPV IRES on the 40S subunit and factor dependence of IRES/80S ribosome assembly. EMBO Rep 5:906–913
14. Zahn-Zabal M, Kobr M, Girod PA, Imhof M, Chatellard P, de Jesus M, Wurm F, Mermod N (2001) Development of stable cell lines for production or regulated expression using matrix attachment regions. J Biotechnol 87:29–42
15. Kim JM, Kim JS, Park DH, Kang HS, Yoon J, Baek K, Yoon Y (2004) Improved recombinant gene expression in CHO cells using matrix attachment regions. J Biotechnol 107:95–105
16. Harland L, Crombie R, Anson S, deBoer J, Ioannou PA, Antoniou M (2002) Transcriptional regulation of the human TATA binding protein gene. Genomics 79:479–482
17. Wong TK, Newmann E (1982) Electric-field mediated gene-transfer. Biochem Biophys Res Commun 107:584–587
18. Kichler A (2004) Gene transfer with modified polyethylenimines. J Gene Med 6(Suppl 1):S3–10
19. Cockett MI, Bebbington CR, Yarranton GT (1990) High level expression of tissue inhibitor of metalloproteinases in Chinese hamster ovary cells using glutamine synthetase gene amplification. Biotechnol (N Y) 8:662–667
20. Lucas BK, Giere LM, DeMarco RA, Shen A, Chisholm V, Crowley CW (1996) High-level production of recombinant proteins in CHO cells using a dicistronic DHFR intron expression vector. Nucleic Acids Res 24:1774–1779
21. Kim SJ, Lee GM (1999) Cytogenetic analysis of chimeric antibody-producing CHO cells in the course of dihydrofolate reductase-mediated gene amplification and their5t stability in the absence of selective pressure. Biotechnol Bioeng 64:741–749
22. Fann CH, Guirgis F, Chen G, Lao MS, Piret JM (2000) Limitations to the amplification and stability of human tissue-type plasminogen activator expression by Chinese hamster ovary cells. Biotechnol Bioeng 69:204–212
23. Yoshikawa T, Nakanishi F, Ogura Y, Oi D, Omasa T, Katakura Y, Kishimoto M, Suga K (2000) Amplified gene location in chromosomal DNA affected recombinant protein production and stability of amplified genes. Biotechnol Prog 16:710–715
24. Brezinsky SC, Chiang GG, Szilvasi A, Mohan S, Shapiro RI, MacLean A, Sisk W, Thill G (2003) A simple method for enriching

populations of transfected CHO cells for cells of higher specific productivity. J Immunol Methods 277:141–155

25. Ham RG (1981) Tissue growth factors. In: Baserga R (ed) Handbook of experimental pharmacology. Springer, New York, p 13
26. Sato GH, Pardee A, Sirbasku DA (eds) (1982) Growth of cells in hormonally defined media, cold spring harbor conference on cell proliferation, vol 9. Cold Spring Harbor Press, Cold Spring Harbor, NY
27. Fletcher T (2005) Designing culture media for recombinant protein production: A rational approach. BioProcess Int 3:30–36
28. Lee GM, Kim EJ, Kim NS, Yoon SK, Ahn YH, Song JY (1999) Development of a serum-free medium for the production of erythropoietin by suspension culture of recombinant Chinese hamster ovary cells using a statistical design. J Biotechnol 69:85–93
29. Liu C, Chu I, Hwang S (2001) Factorial designs combined with the steepest ascent method to optimize serum-free media for CHO cells. Enzyme Microb Technol 28:314–321
30. Chun C, Heineken K, Szeto D, Ryll T, Chamow S, Chung JD (2003) Application of factorial design to accelerate identification of CHO growth factor requirements. Biotechnol Prog 19:52–57
31. Allison DW, Aboytes KA, Fong DK, Leugers SL, Johnson TK, Loke HN, Donahue LM (2005) Development and optimization of cell culture media: Genomic and proteomic approaches. BioProcess Int 31:38–45
32. Zhu MM, Lee ES, Hermans WR, Wasilko DJ (2001) Overview and serum-free medium development for mammalian cell culture. Fourth Conference on recent advances in fermentation technology (RAFTIV), Nov 11–13, Long Beach, CA
33. Varley J, Birch J (1999) Reactor design for large scale suspension animal cell culture. Cytotechnol 29:177–205
34. Ozturk SS (1996) Engineering challenges in high density cell culture systems. Cytotechnol 22:3–16
35. Myers KJ, Reeder MF, Fasano JB (2002) Optimize mixing by using the proper baffles. Chem Eng Progress 98:42–47
36. Chisti Y (1993) Animal cell culture in stirred bioreactors: Observations on scale-up. Bioprocess Eng 9:191–196
37. Pattison RN, Swamy J, Mendenhall B, Hwang C, Frohlich BT (2000) Measurement and control of dissolved carbon dioxide in mammalian cell culture processes using an in situ fiber optic chemical sensor. Biotechnol Prog 16:769–774
38. Dowd JE, Jubb A, Kwok KE, Piret JM (2003) Optimization and control of perfusion cultures using a viable cell probe and cell specific perfusion rates. Cytotechnol 42:35–45
39. Noll T, Biselli M (1998) Dielectric spectroscopy in the cultivation of suspended and immobilized hybridoma cells. J Biotechnol 63:187–198
40. Gupta A, Rao G (2003) A study of oxygen transfer in shake flasks using a non-invasive oxygen sensor. Biotechnol Bioeng 84:351–358
41. Ge X, Kostov Y, Rao G (2005) Low-cost noninvasive optical CO_2 sensing system for fermentation and cell culture. Biotechnol Bioeng 89:329–334
42. Singh V (1999) Disposable bioreactor for cell culture using wave-induced agitation. Cytotechnol 30:149–158
43. Ozturk SS (2005) Batch versus perfusion: a real case comparison of highly developed cell culture processes for the production of monoclonal antibodies. 229th National Meeting American Chemical Society, Mar 13–17, San Diego, CA
44. Voisard D, Meuwly F, Ruffieux PA, Baer G, Kadouri A (2003) Potential of cell retention techniques for large-scale high-density perfusion culture of suspended mammalian cells. Biotechnol Bioeng 82:751–765
45. Alrubeai M, Singh RP, Emery AN, Zhang Z (1995) Cell cycle and cell size dependence of susceptibility to hydrodynamic forces. Biotechnol Bioeng 46:88–92
46. Alrubeai M, Singh RP, Goldman MH, Emery AN (1995) Death mechanisms of animal cells in conditions of intensive agitation. Biotechnol Bioeng 45:463–472
47. Gregoriades N, Clay J, Ma N, Koelling K, Chalmers JJ (2000) Cell damage of microcarrier cultures as a function of local energy dissipation created by a rapid extensional flow. Biotechnol Bioeng 69:171–182
48. Ma N, Koelling KW, Chalmers JJ (2002) Fabrication and use of a transient contractional flow device to quantify the sensitivity of mammalian and insect cells to hydrodynamic forces. Biotechnol Bioeng 80:428–437
49. Kunas KT, Papoutsakis ET (1990) Damage mechanisms of suspended animal cells in agitated bioreactors with and without bubble entrainment. Biotechnol Bioeng 36:476–483
50. Chattopadhyay D, Rathman JF, Chalmers JJ (1995) Thermodynamic approach to explain cell adhesion to air-medium interfaces. Biotechnol Bioeng 48:649–658
51. Ma N, Chalmers JJ, Aunins JG, Zhou W, Xie L (2004) Quantitative studies of cell-bubble interactions and cell damage at different Pluronic F-68 and cell concentrations. Biotechnol Prog 20:1183–1191
52. Osman JJ, Birch J, Varley J (2001) The response of GS-NS0 myeloma cells to pH shifts and pH perturbations. Biotechnol Bioeng 75:63–73
53. Sauer PW, Burky JE, Wesson MC, Sternard HD, Qu L (2000) A high-yielding, generic fed-batch cell culture process for production of recombinant antibodies. Biotechnol Bioeng 67:585–597
54. Lao MS, Toth D (1997) Effects of ammonium and lactate on growth and metabolism of a recombinant Chinese hamster ovary cell culture. Biotechnol Prog 13:688–691
55. Ozturk SS, Thrift JC, Blackie JD, Naveh D (1997) Real-time monitoring and control of glucose and lactate concentrations in a mammalian cell perfusion reactor. Biotechnol Bioeng 53:372–378
56. Martinelle K, Westlund A, Haggstrom L (1996) Ammonium ion transport: A cause of cell death. Cytotechnol 22:251–254
57. Genzel Y, Ritter JB, Konig S, Alt R, Reichl U (2005) Substitution of glutamine by pyruvate to reduce ammonia formation and growth inhibition of mammalian cells. Biotechnol Prog 21:58–69
58. Yang M, Butler M (2000) Effect of ammonia on the glycosylation of human recombinant erythropoietin in culture. Biotechnol Prog 16:751–759
59. Gray DR, Chen S, Howarth W, Inlow D, Maiorella BL (1996) CO_2 in large-scale and high-density CHO cell perfusion culture. Cytotechnol 22:65–78
60. deZengotita VM, Schmelzer AE, Miller WM (2002) Characterization of hybridoma cell responses to elevated pCO_2 and osmolality: Intracellular pH, cell size, apoptosis, and metabolism. Biotechnol Bioeng 77:369–380
61. Zhu MM, Goyal A, Rank DL, Gupta SK, Vanden Boom T, Lee SS (2005) Effects of elevated pCO_2 and osmolality on growth of CHO cells and production of antibody-fusion protein B1: A case study. Biotechnol Prog 21:70–77
62. Mostafa SS, Gu X (2003) Strategies for improved dCO_2 removal in large-scale fed-batch cultures. Biotechnol Prog 19:45–51
63. Chen ZL (2004) Temperature shift as a process optimization step for the production of pro-urokinase by a recombinant Chinese hamster ovary cell line in high-density perfusion culture. J Biosci Bioeng 97:239–243
64. Clark KJ, Chaplin FW, Harcum SW (2004) Temperature effects on product-quality-related enzymes in batch CHO cell cultures producing recombinant tPA. Biotechnol Prog 20:1888–1892
65. Fox SR, Patel UA, Yap MG, Wang DI (2004) Maximizing interferon-gamma production by Chinese hamster ovary cells through temperature shift optimization: Experimental and modeling. Biotechnol Bioeng 85:177–184
66. Bollati-Fogolin M, Forno G, Nimtz M, Conradt HS, Etcheverrigaray M, Kratje R (2005) Temperature reduction in cultures of hGM-CSF-expressing CHO cells: Effect on productivity and product quality. Biotechnol Prog 21:17–21

67. Dempsey J, Ruddock S, Osborne M, Ridley A, Sturt S, Field R (2003) Improved fermentation processes for NS0 cell lines expressing human antibodies and glutamine synthetase. Biotechnol Prog 19:175–178
68. Wong DCF, Wong KTK, Goh LT, Heng CK, Yap MGS (2005) Impact of dynamic online fed-batch strategies on metabolism, productivity and N-glycosylation quality in CHO cell cultures. Biotechnol Bioeng 89:164–177
69. Dowd JE, Kwok KE, Piret JM (2000) Increased t-PA yields using ultrafiltration of an inhibitory product from CHO fed-batch culture. Biotechnol Prog 16:786–794
70. Dowd JE, Kwok KE, Piret JM (2001) Predictive modeling and loose-loop control for perfusion bioreactors. Biochem Eng J 9:1–9
71. Nienow AW, Langheinrich C, Stevenson NC, Emery AN, Clayton TM, Slater NKH (1996) Homogenisation and oxygen transfer rates in large agitated and sparged animal cell bioreactors: Some implications for growth and production. Cytotechnol 22:87–94
72. Gardner AR, Smith TM (2000) Identification and establishment of operating ranges of critical process variables. In: Sofer G, Zabriskie DW (eds) Biopharmaceutical process validation. Marcel Dekker, New York, pp 61–76
73. Moran EB, McGowan ST, McGuire JM, Frankland JE, Oyebade IA, Waller W, Archer LC, Morris LO, Pandya J, Nathan SR, Smith L, Cadette ML, Michalowski JT (2000) A systematic approach to the validation of process control parameters for monoclonal antibody production in fed-batch culture of a murine myeloma. Biotechnol Bioeng 69:242–255
74. Lyddiatt A (1981) Downstream processing: Protein recovery. In: Butler M (ed) Mammalian cell biotechnology: a practical approach. IRL Press at Oxford University Press, New York, pp 187–206
75. Bennan J, Bing F, Boone H, Fernandez J, Seely B, van Deinse H, Miller D (2002) Evaluation of extractables from product-contact surfaces. Biopharm Int 15:22–34
76. Kemp G, O'Neil P (2004) Large-scale production of therapeutic antibodies: Considerations for optimizing product capture and purification. In: Subramanian G (ed) Antibodies, vol 1, Production and purification. Kluwer, Boston, pp 75–100
77. Burton S (2002) A generic approach to the purification of monoclonal antibodies: An alternative to protein A. IBC Conference: antibody production & downstream processing, Feb 13–15, San Diego, CA
78. Jacob LR, Frech M (2004) Scale-up of antibody purification. In: Subramanian G (ed) Antibodies, vol 1, Production and purification. Kluwer, Boston, pp 101–131
79. Gattschalk U (2005) Large scale manufacturing of mAbs and the backlog in bioseparation technologies. IBC Conference: antibody production & downstream processing, March 8–11, San Diego, CA
80. Hubbard B (2005) Platform approaches to monoclonal antibody purification. IBC Conference: antibody production & downstream processing, Mar 8–11, San Diego, CA
81. Fish B (2002) Taking a monoclonal antibody from mg to kg scale: Production strategies, issues and successes. IBC Conference: Scaling-up From Bench to Clinic and Beyond, Aug 14–16, San Diego, CA
82. Aranha-Creado H (1998) Clearance of murine leukaemia virus from monoclonal antibody solutions by a hydrophilic PVDF microporous membrane filter. Biologicals 26:167–172
83. Rathore A, Velaydhan A (2003) Guidelines for optimization and scale-up in preparative chromatography. Biopharm Int 16:34–42

Biomass Conversion

33

Stephen R. Decker, John Sheehan, David C. Dayton, Joseph J. Bozell,
William S. Adney, Bonnie Hames, Steven R. Thomas, Richard L. Bain,
Stefan Czernik, Min Zhang, and Michael E. Himmel

An Introduction to Biomass and the Biorefinery

What Is Biomass?

In its simplest terms, biomass is all the plant matter found on our planet. Biomass is produced directly by photosynthesis, the fundamental engine of life on earth. Plant photosynthesis uses energy from the sun to combine carbon dioxide from the atmosphere with water to produce organic plant matter. More inclusive definitions are possible. For example, animal products and waste can be included in the definition of biomass. Animals, like plants, are renewable; but animals clearly are one step removed from the direct use of sunlight. Using animal rather than plant material thus leads to substantially less efficient use of our planet's ultimate renewable resource, the sun. So, we emphasize plant matter in our definition of biomass. It is the photosynthetic capability of plants to utilize carbon dioxide from the atmosphere that leads to its designation as a "carbon neutral" fuel, meaning that it does not introduce new carbon into the atmosphere. In reality—as discussed later in the description of life cycle assessments of biomass use—we find that biomass fuels are not quite carbon neutral, because somewhere in the life cycle of their production, conversion, and distribution, some fossil energy carbon is released.

Few people understand the scale of energy and organic matter captured by photosynthesis in the form of biomass on the planet. Consider, for example, that the sun sends 3×10^{24} J of energy per year to our planet [1]. On average, over the entire globe, plant photosynthesis captures only 0.1% of the solar energy bombarding our world, storing 3×10^{21} J of energy in biomass annually. How does that compare with world energy production? The U.S. Department of Energy (DOE) estimates that, in 2000, global primary energy production in the form of oil, coal, natural gas, nuclear power, hydroelectric power, and other forms (including a small amount of renewable energy) was 400 quadrillion Btus annually or 4.2×10^{20} J/year [2]. Thus, energy stored in biomass each year worldwide is seven times greater than humankind's annual energy production. Can we expect to tap all of this energy to meet the ever-growing demand for energy? Could we be taking much greater advantage of this renewable source of organics and stored energy?

The Structure and Composition of Biomass

Why, then, isn't biomass the major source of our energy supply? The answer is that—in comparison to the relatively recent discovery and use of fossil energy sources—biomass

---

S.R. Decker

J. Sheehan

D.C. Dayton

J.J. Bozell

W.S. Adney

B. Hames

S.R. Thomas

R.L. Bain

S. Czernik

M. Zhang

M.E. Himmel (✉)
National Renewable Energy Laboratory, National Bioenergy Center,
1617 Cole Boulevard, Golden, CO 80401, USA
e-mail: mike_himmel@nrel.gov

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_33,

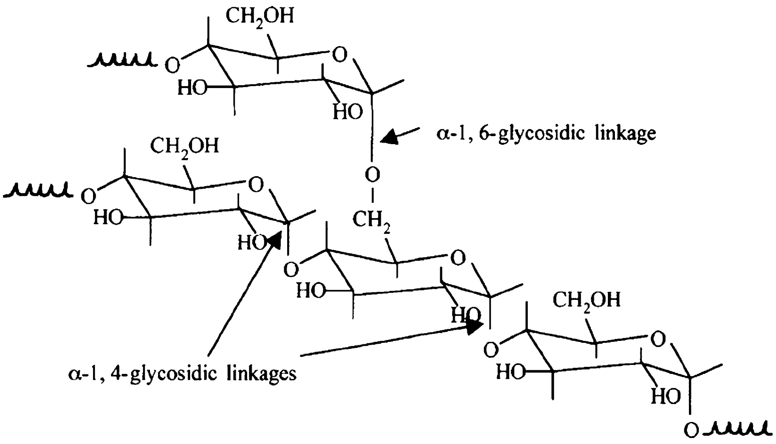


Fig. 33.1 Starch: a biopolymer of glucose molecules used today in the United States as the basis for production of fuel ethanol from corn grain

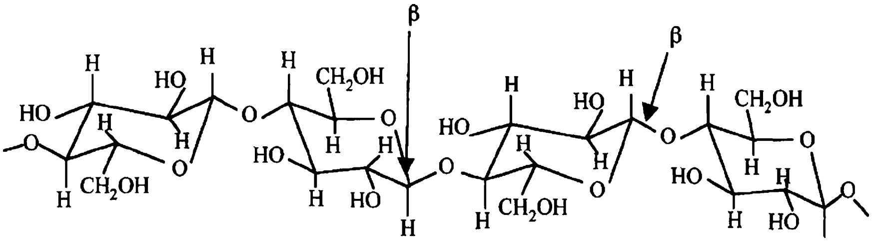


Fig. 33.2 Cellulose: a biopolymer of glucose found in the structural components of all plant cell walls

has proved more difficult or at least more costly to convert into convenient forms of energy. Thus, although biomass has a very long history as humankind's first primary source of energy, it has tended to find its most important use in modern times as a source of higher-value (societally and economically) food and fiber products.

There are, of course, many forms of biomass. In modern agriculture, many crops are grown for the starch, sugars, protein, natural oils, and fiber they contain. Sugars from sugarcane are used today in Brazil to produce fuel ethanol [3]. They were also the primary source of fermentable sugars in the United States until molasses became too costly [4]. In the United States today, ethanol made from sugars derived from starch in corn is the largest source of renewable transportation fuel. Starch consists of glucose molecules strung together by α-glycosidic linkages. These linkages occur in chains of α-1,4 linkages with branches formed as a result of α-1,6-linkages (see Fig. 33.1). The ability to release sugars from starch is common to many animals, including humans. It is thus no surprise that starch was the first major carbohydrate biopolymer to be used for energy production other than combustion. Indeed, the use of enzymes to biologically release sugars from starch is one of the earliest examples of modern industrial enzyme technology. It stands as an example of how biotechnology might be used to process other forms of biomass [5].

To foster a new bioindustry, we need to turn to less used—but more abundant—forms of biomass. This means focusing on the lignocellulosic components of the cell walls of all higher plants. Here we find highly complex material, both in chemical composition and structure. Plant cell walls are divided into two sections, the primary and the secondary cell walls [6]. The primary cell wall, which provides structure for expanding the cell is composed of the major polysaccharides and a group of basic glycoproteins, primarily extensins [7]. The predominant polysaccharide in the primary cell wall is cellulose; the second most abundant is hemicellulose; and the third is pectin. Because cellulose is made up only of β-(1,4)-linkages, it has a highly linear structure that encourages the formation of strong hydrogen bonds between chains of cellulose (see Fig. 33.2). The high level of hydrogen bonding among the chains makes it difficult to attack or depolymerize, either chemically or biologically. Once depolymerized, however, it is relatively easy to ferment, because it consists of the single sugar, glucose. Hemicelluloses, in contrast, are biopolymers of six- and five-carbon sugars that are almost always branched with a wide spectrum of substituents, including acetyl esters, uronic acids, sugars, hydroxycinnamic acids, and other moieties distributed along the backbone polysaccharide. The branched and amorphous nature of hemicellulose is vulnerable to degradation, but organisms in nature do not readily utilize some of its various sugars. Hemicelluloses are thought to hydrogen bond to cellulose, as well as to other hemicelluloses, which helps stabilize the cell wall matrix and renders the cell wall insoluble in water.

The secondary cell wall, produced after the cell has completed growing, also contains polysaccharides and is strengthened by polymeric lignin covalently cross-linking to hemicellulose [6]. Lignin is a high-energy content biopolymer rich in phenolic components. It provides structural integrity to plants. The combination of hemicellulose and lignin provide a protective sheath around the cellulose and this sheath must be modified or removed before efficient hydrolysis of cellulose can occur.

The picture of biomass sketched here offers important insights. The complex structure of lignocellulosic biomass provides protection and structural integrity to biomass. This makes things more difficult for industrial scientists and engineers seeking to convert biomass into useful forms of energy, chemicals, and products. On the other hand, the diversity of biomass composition offers opportunities for processing biomass into a wide range of new and existing chemicals. The challenge to the industrial processor is to develop sophisticated and robust approaches to optimizing the recovery and conversion of each component from this highly complex structure.

Converting Biomass into Energy and Other Products

Biological Versus Thermochemical Processing

The distinguishing feature of the various paths from biomass to useful energy is the choice of conversion technology. These technologies fall into two main categories: thermochemical and biochemical conversion. Thermochemical processing, as its name suggests, relies on heat and chemical catalysis to produce useful energy and products from biomass. Likewise, biochemical processing relies on biological organisms and biological catalysts to transform biomass into energy and products. They are complementary, rather than competing, technology paths that are part of an integrated scheme for making a wide range of fuels, products, heat, and power from biomass.

Thermochemical conversion of biomass. The simplest form of thermochemical conversion is direct combustion of biomass to produce heat and power. Direct combustion of biomass for electricity production in the United States currently supports more than 10,000 MW of generating capacity, with the majority of the biomass coming from landfill gas, municipal solid waste, and wood fuel [8]. The Energy Information Administration of the DOE estimates that 3 gigawatts of electric capacity could currently be supplied by biomass at prices competitive with coal-derived electricity [9]. Direct combustion of biomass in existing power plants is a very effective means of offsetting fossil energy use because it displaces conventional fossil fuels in electric generating stations on essentially a one-to-one basis.

At the heart of most advanced thermochemical conversion processes is the conversion of biomass into simple chemical intermediates in a process known as gasification, as shown in Fig. 33.3. Gasification is a process in which biomass is heated in the presence of air, oxygen, and/or steam to produce a gas mixture that can be used as an energy source itself or can be converted into a variety of fuels and products. Gasification dates back to the early 1800s. It has its roots in technology originally used for the production of gas from coal for lighting and heating [10]. There are many possible variations for the production of a gas mixture often referred to as a synthesis gas or "syngas." The core steps, however, are gas production in a gasification reactor, tar cracking (integrated or separate from the gasifier), and cleanup of the syngas. The clean syngas can then be used to generate power by direct combustion or in a gas turbine/combined cycle system or be synthesized to other fuels or products.

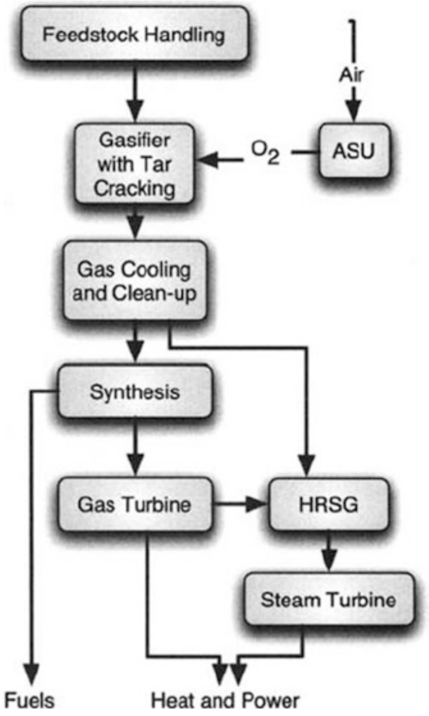


Fig. 33.3 General schematic of a process for thermochemical conversion of biomass to fuels, heat, and power. *ASU* Air Separation Unit; *HRSG* Heat Recovery Steam Generator

The severity of the conditions in this process affects the distribution of products, which can include pyrolysis oils, chars, and gases. The chemistry of gasification is relatively simple [11]. The key reaction that takes place in gasifiers is the "water gas shift" reaction. In this reaction, water (provided by steam) can combine with carbon to form carbon monoxide and hydrogen:

$$C + H_2O = CO + H_2$$

The carbon monoxide, hydrogen, and steam in the gasifier can then undergo more reaction to produce more hydrogen, carbon dioxide, and methane. The relative composition of the syngas will depend on the reactor conditions, choice of

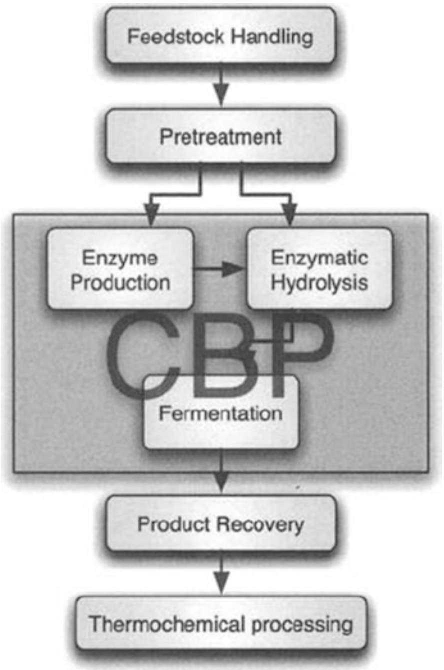


Fig. 33.4 General schematic of a biological process for converting lignocellulosic biomass to ethanol

catalysts, and the relative amounts of air, oxygen, and steam used. A wealth of chemistry is possible with this collection of simple compounds. Syngas can, therefore, serve as a platform for producing all sorts of fuels and other products. Syngas from coal, for example, is used to produce gasoline or diesel fuel products in large-scale commercial operations using so-called "Fischer–Tropsch" catalysts that can convert these simple compounds into hydrocarbons and oxygenates. Selection of conditions and catalysts in the water gas shift reaction also provides an opportunity to produce hydrogen from biomass for use in fuel cells [12].

Experience with biomass gasification, however, has been limited to heat and power generation. In the United States, the DOE helped to test gasification of wood, coupled with electricity generation, at a power plant in Vermont, successfully demonstrating this technology at the 200-t/day (5-MW) scale [13].

Biological conversion of biomass. Ethanol is typically the major fuel product of biological conversion. The basic steps of biological conversion are shown in Fig. 33.4. Lignocellulosic biomass can be converted into mixed-sugar solutions plus lignin-rich solid residues by the sequential use of thermochemical pretreatment and enzymatic saccharification. Sugars from hemicellulose and cellulose can then be fermented to ethanol for fuel production. Not shown in Fig. 33.4 is the possibility of biologically or chemically converting these sugars into other products.

There is a long and rich history of using acid and base catalysts to release the sugars found in cellulose and hemicellulose dating back to the discovery of wood sugars in the nineteenth century. The technology was commercialized during World War I in the United States, during World War II in Germany, and later in the twentieth century in the Soviet Union [14–34]. More advanced schemes for biological processing are under development today; however, they rely on this chemical hydrolysis step only as a pretreatment for removal of hemicellulose. Biologically mediated hydrolysis of cellulose is now viewed as the most selective and efficient means of hydrolyzing or depolymerizing the cellulose biopolymer to release its glucose sugar monomers. Many workers in the field agree that cellulose decrystallization and depolymerization are indeed the rate-limiting steps in the conversion of lignocellulosic biomass.

Removal of hemicellulose by dilute-acid pretreatment has been the classic means of rendering biomass more amenable to cellulase action [35]. In a hallmark study, Soltes and coworkers [36] showed that biomass with reduced acetylation responded significantly more favorably to cellulase action than did native biomass. Selig and co-workers more recently showed that enzymatic acetyl removal was effective in enhancing biomass hydrolysis after a variety of pretreatments. [679]. Although still controversial, there is some indication that biomass with reduced lignin content is also more readily hydrolyzed by cellulase action [37, 38]. At least one study has shown that lignin removal enhances enzymatic hemicellulose hydrolysis, the removal of which allows for greater hydrolysis of cellulose [680]. One key to understanding cellulase action on biomass is the fact that the structural and reactive chemical components of the substrate—primarily defined as acetyl and lignin contents—strongly affect enzyme access to cellulose. Another is that once cellulase component enzymes are available in sufficient ratio and concentration at the site of hydrolysis, the degree of cellulose crystallinity controls the hydrolytic rate [37, 39]. For the engineer seeking to improve this natural process, the key challenge is to make biomass depolymerization a more rapid and less costly conversion.

To be cost-effective, the process must use organisms capable of fermenting the full spectrum of five- and six-carbon sugars released from cellulose and hemicellulose. The advent of efficient genetically engineered organisms equipped with metabolic pathways to handle both types of sugars is an important improvement in the process that has occurred over the past decade or so [40–45]. The first generation of fuel ethanol technology for lignocellulosic biomass will include the production and use of enzymes for the hydrolysis of cellulose that work with these new versatile fermenting organisms. A future advance in the technology may involve the development of organisms capable of carrying out both the enzymatic hydrolysis of cellulose and the fermentation of the resulting sugars. This configuration—sometimes referred to as consolidated bioprocessing—would represent a

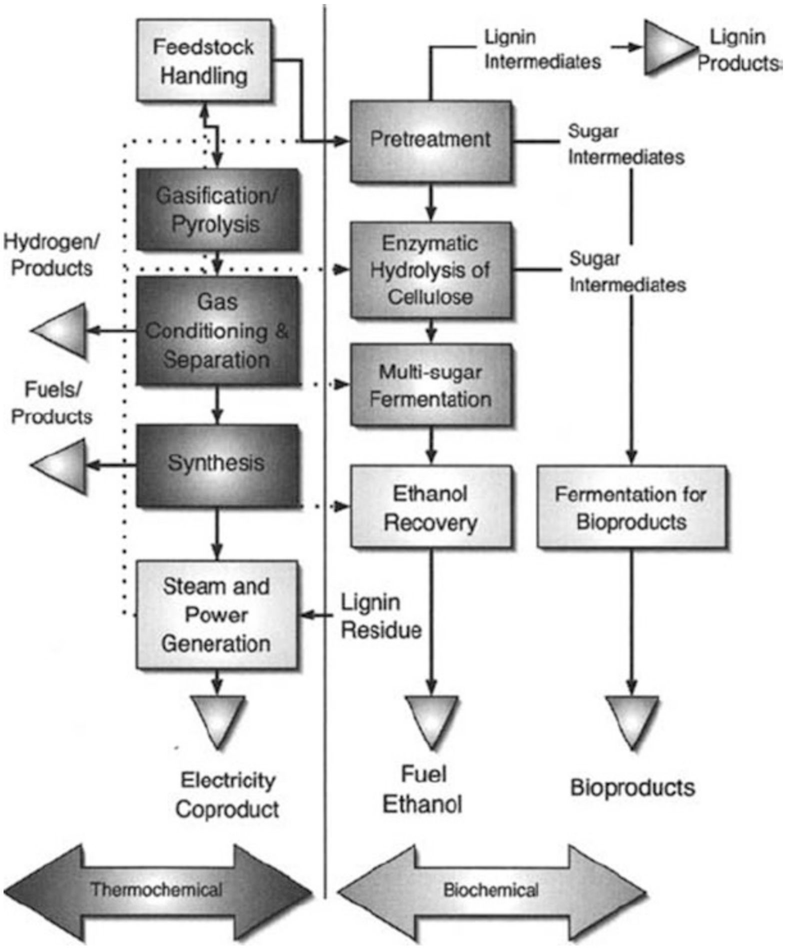


Fig. 33.5 General schematic of an ideal biorefinery combining biological and thermochemical processes for production of fuels, chemicals, heat, and power

significant step toward simplification and cost reduction of the process [46].

A purely standalone biological process is unlikely without significant improvement in both feedstock convertibility and microbial degradative and conversion capacity. Even after a combination of thermochemical pretreatment and biologicval conversion, there is a significant amount of residue left over after biological conversion of the carbohydrates. As shown in Fig. 33.4, this residue, containing mostly lignin, is usually sent to some form of thermochemical processing, even if that involves nothing more than direct combustion for production of heat, power, or both.

The Biorefinery

The integrated biorefinery is a conceptual framework that capitalizes on the synergies of integrating technologies from both biological and thermochemical process schemes. Furthermore, like the petroleum refinery, the biorefinery has the potential to combine production of low-volume, high-value products with high-volume, low-value fuels or energy production. The ideal biorefinery employs a combination of conversion technologies that maximizes the value of fuels, chemical, materials, and power made from biomass, as shown in Fig. 33.5.

The biorefinery should benefit from lessons learned during the evolution of modern-day petroleum refineries. These combine use of fluid catalytic cracking, thermal cracking, and hydrocracking technology to convert the higher-boiling-range fractions of crude oil into more useful lower-boiling-range products. Just as few petroleum refineries use all available conversion technologies, biorefineries too will use only those technology platforms that are most cost-effective for converting a certain type of biomass into a certain collection of desired endproducts.

Biomass, Fossil Energy Savings, and Greenhouse Gas Mitigation

The two most often touted benefits of biomass use are the ability to avoid the use of nonrenewable fossil energy resources and the concomitant ability to reduce net greenhouse gas generation, characteristics that can only be

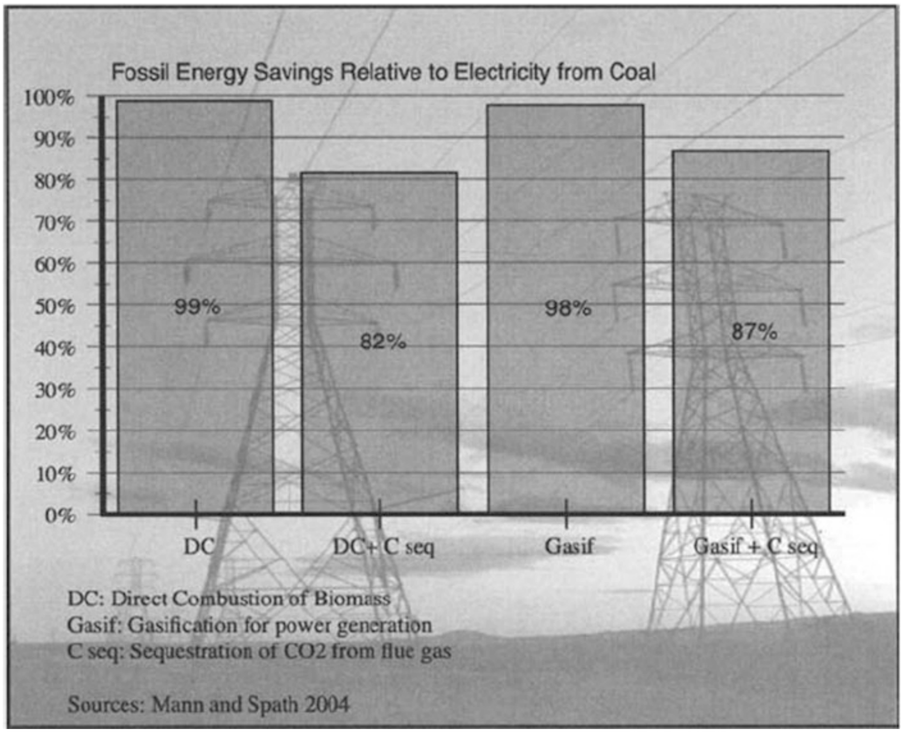


Fig. 33.6 Fossil energy savings for different biomass-to-power scenarios

assessed by using life-cycle assessment. Life-cycle assessment [47–52] is an analytical methodology for understanding the full impact of a given product or service on the environment and on the overall demand placed on our natural resources. In general, this involves looking at all of the stages of production and use of a product or service, from the first extraction of raw materials from the environment to the ultimate disposition of the product. For energy, this includes production and transport of a feedstock (whether that is coal, natural gas, oil, or biomass) to a conversion facility, conversion of the feedstock to a useful energy form, and distribution and use of the final energy product.

Life-Cycle Assessments of Biopower

Figure 33.6 summarizes the results of life-cycle assessments of the impact of various pathways for electricity generation from biomass on fossil energy requirements [53]. Direct combustion and gasification of biomass for power production provides 99 and 98% savings in fossil energy use, respectively, compared to combustion of coal for electricity generation. When carbon sequestration is added to either of these options, fossil energy savings are reduced because of the added energy demand for sequestering the carbon dioxide from the stack of the power plant. Greenhouse gas savings for these same technology options are also very good, often exceeding 100% reductions due to the effects sequestration of carbon in the soil that occurs when energy crops such as switchgrass are grown (see Fig. 33.7).

Life-Cycle Assessments of Biofuels

Figure 33.8 shows fossil-energy savings associated with the production of hydrogen, ethanol, and Fischer–Tropsch liquids from biomass. Savings range from 91 to 102%. Ethanol and Fischer–Tropsch fuels are both liquids that can be used in existing internal combustion engine vehicles, whereas hydrogen—which is a gaseous fuel—requires significant changes in vehicle technology, whether it is burned directly for power or used in a fuel cell. Furthermore, although the energy savings for hydrogen are comparable to the other two liquid fuels, these savings may be offset by higher energy requirements for distribution of hydrogen, which are not included in these energy balance calculations.

Estimates of greenhouse gas emissions for hydrogen and Fischer–Tropsch liquids made from biomass available. A number of estimates are available for have recently become ethanol made from biomass. These are shown in Fig. 33.9 for the use of E85 (a blend of ethanol, 85%, and gasoline, 15%) in a flexible fuel vehicle. The choice of feedstock can have a significant impact, with corn grain ethanol having the lowest benefits of those studied. The big difference between corn grain and the other biomass resources is that the former relies on fossil fuels to provide energy in the conversion facility, whereas the others make use of residual lignin from the biomass to meet all of the energy needs for conversion, with some left over for cogeneration of electricity.

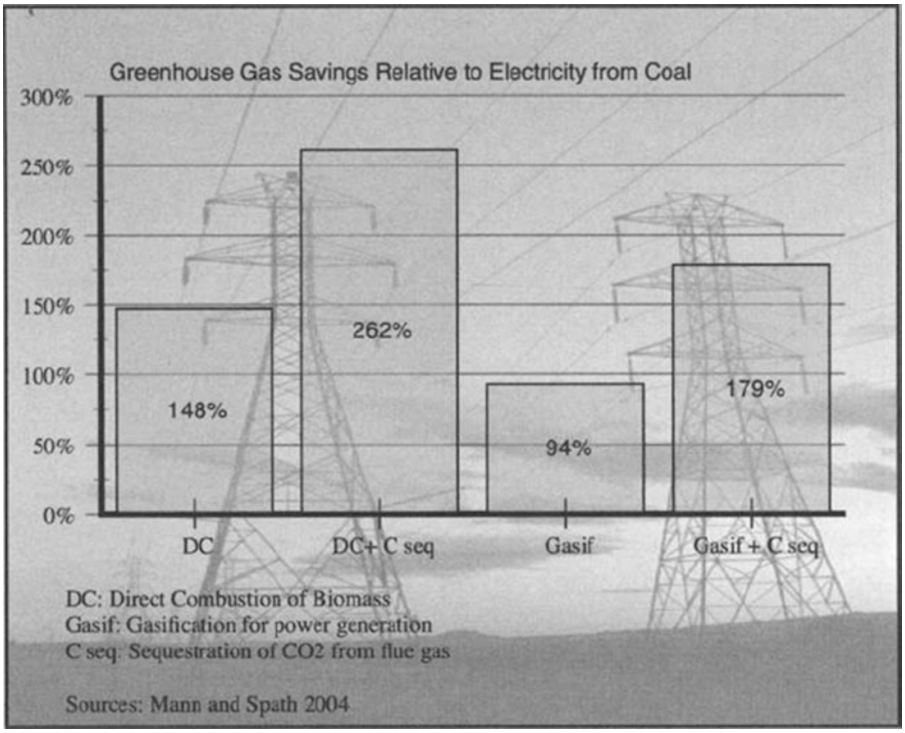


Fig. 33.7 Greenhouse gas savings for different biomass-to-power scenarios

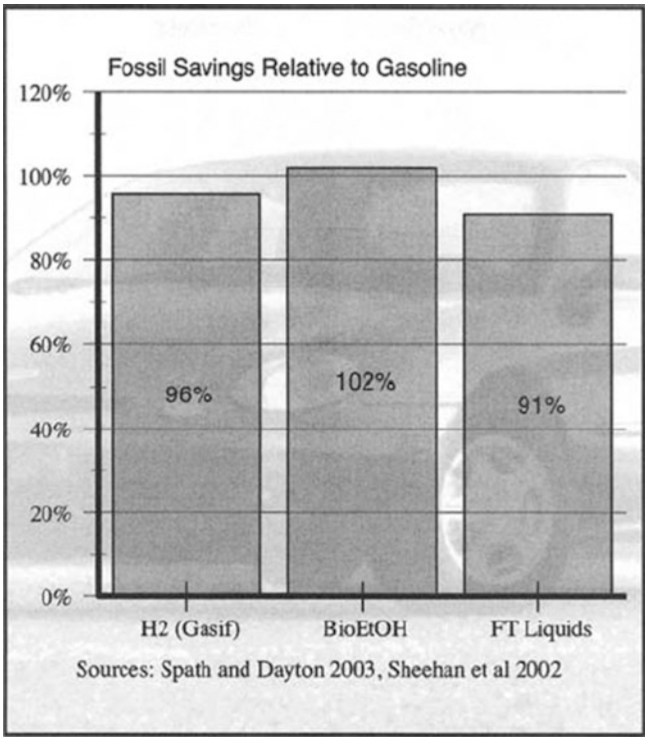


Fig. 33.8 Fossil energy savings for different biofuels

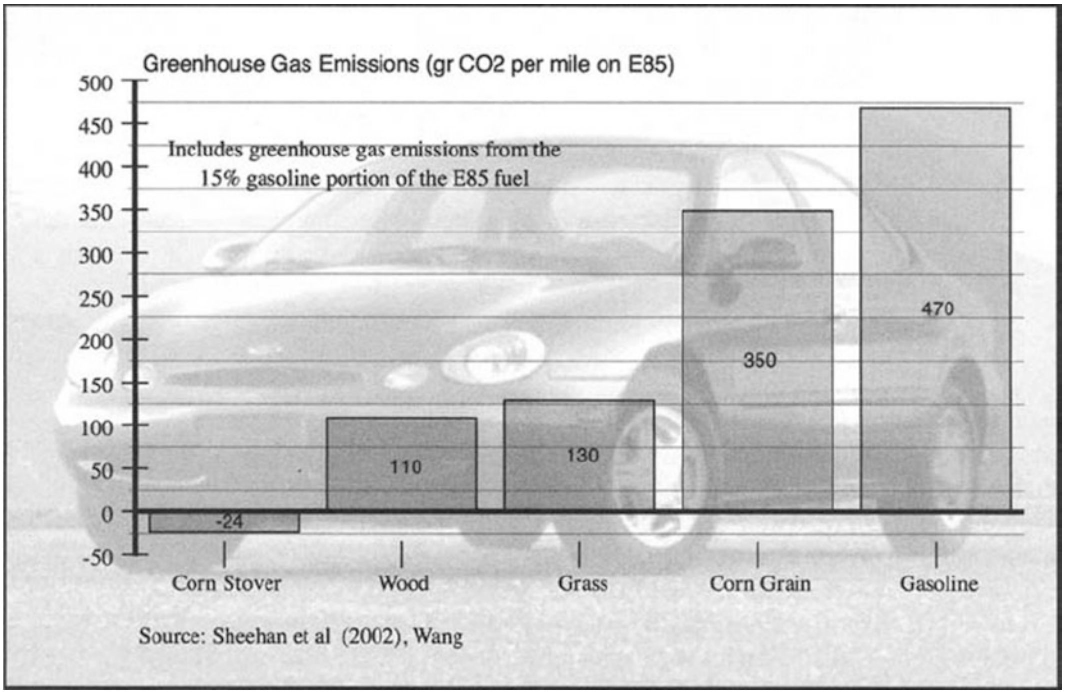
Cost Projections for Fuels and Power from Biomass

The Competitiveness of Biopower

Figure 33.10 compares the range and midpoint of costs for power production from coal and natural gas with the projected range and midpoint of costs of several biomass technology options [54]. As the numbers indicate, coal remains the lowest-cost option by far for producing electricity. Biomass technologies are from two to four times more costly and with predictions that coal prices are actually going to decline [55], this makes the prospects for competitive electricity generation from biomass a goal that will require advances in technology that have not, as yet, even been thought about, much less planned for.

The Competitiveness of Biofuels

Figure 33.11 provides a similar comparison for published estimates of the cost of transportation fuels made from biomass. None of these technologies is ready to compete with gasoline strictly on the basis of delivering energy to a vehicle. In the case of ethanol, the DOE has developed plans for reducing its cost to about $1.76 per gallon ethanol, or $2.62/gallon gasoline equivalent for 2012, which puts this

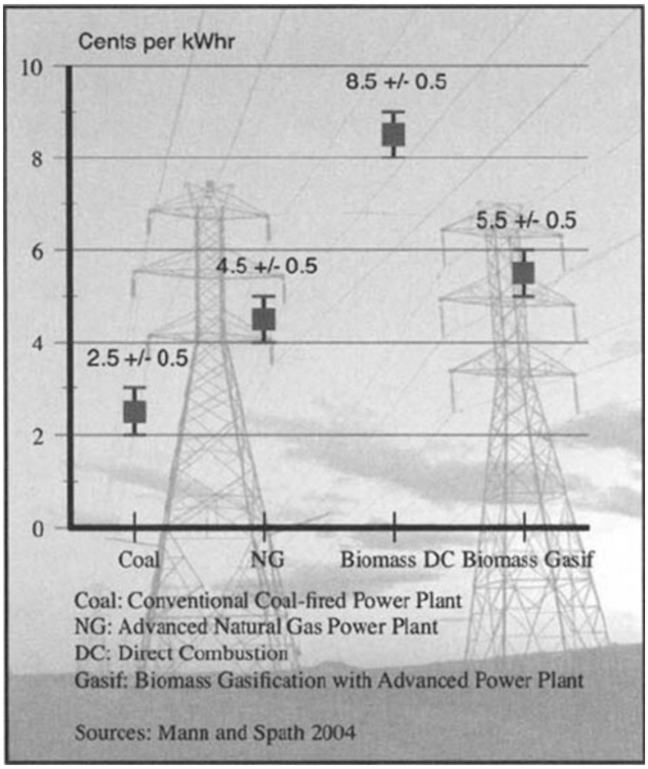


Greenhouse Gas Emissions (gr CO2 per mile on E85)

Includes greenhouse gas emissions from the 15% gasoline portion of the E85 fuel

Source: Sheehan et al (2002), Wang

Fig. 33.9 Greenhouse gas emissions for E85 made from different biomass feedstocks and gasoline

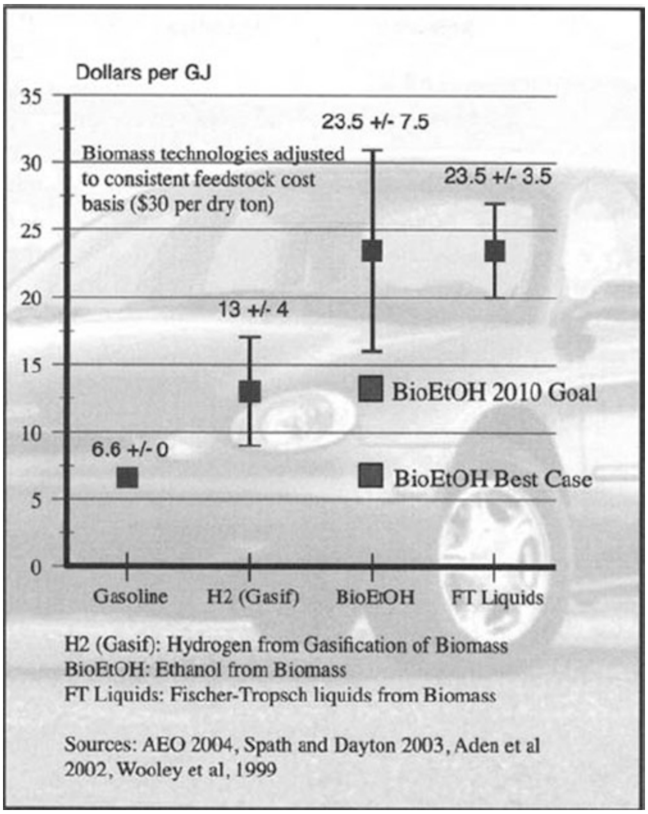


Cents per kWhr

Coal: Conventional Coal-fired Power Plant
NG: Advanced Natural Gas Power Plant
DC: Direct Combustion
Gasif: Biomass Gasification with Advanced Power Plant

Sources: Mann and Spath 2004

Fig. 33.10 Comparison of costs for conventional fossil fuel electricity generation and biomass power options



Dollars per GJ

Biomass technologies adjusted to consistent feedstock cost basis ($30 per dry ton)

H2 (Gasif): Hydrogen from Gasification of Biomass
BioEtOH: Ethanol from Biomass
FT Liquids: Fischer-Tropsch liquids from Biomass

Sources: AEO 2004, Spath and Dayton 2003, Aden et al 2002, Wooley et al, 1999

Fig. 33.11 Comparison of costs for gasoline and various biomass derived fuels for transportation new technology in a competitive range with ethanol made from corn grain. Recent published long-term (mature technology) estimates of the price that ethanol could reach are hard to find. In the early 1990s, the DOE published estimates for a best-case ethanol cost of $0.67 per gallon, or $8.4 per GJ. In 1996, Lynd reported a "best-parameter" case that brings ethanol cost down to $0.50 per gallon, or $6 per GJ, which is in line with gasoline's wholesale price [56]. These long-term projections show that it is possible for a bio-based fuel such as ethanol to approach parity with its fossil fuel counterpart, but that achieving such parity represents a significant stretch for the technology, with respect to its current or even potential cost.

Biomass as a Sustainable and Substantial Energy Source

Biomass as a Substantial Source of Energy: Balancing the Demands on Our Land

The estimates of global photosynthetic capacity described at the beginning of this chapter suggest that biomass could be a truly substantial source of renewable energy. The life-cycle assessments reported in this chapter also point to biomass as a highly sustainable source of energy.

The potential of biomass as a large source of fuels and chemicals, however, hinges on how much of the huge global capacity for biomass production can be harnessed for that purpose, after considering other demands placed on our land. And the question has been nothing if not controversial. Why? Because we recognize that the role land plays as the primary source of food and fiber is its most critical job for society. No other resource can provide us with food. So, when it comes to choosing between the use of land for producing food or energy, the choice is pretty obvious. Forget about the use of our land for providing energy to drive our cars and even to heat and light our homes, if it comes at the price of starvation.

Uncertainties of understanding sustainable management of our land, debates about what level of energy use is both adequate and sustainable for our society, and the unknown potential for future advances in science and technology all fuel controversy over the appropriate use of land. The public, policymakers, and even experts in energy find themselves bombarded with widely divergent perspectives. Here are just a few examples of the "pessimists":

- In 2002, a prominent collection of scientists wrote, "Biomass plantations can produce carbon-neutral fuels for power plants or transportation, but photosynthesis has too low a power density (~6 W/m^2) for biofuels to contribute significantly to climate stabilization." These researchers find the scale of our growing energy demand so daunting that they conclude that none of today's current options for renewable energy or energy efficiency are up to the task. They call for an unprecedented scientific effort to tackle what they call "the technology challenge of the century" [57].
- In 2002, a group of researchers led by David Pimentel dismissed biofuels for transportation outright because of their negative energy balance, and concluded that burning biomass for heat and power might achieve 5.28×10^{18} J/year worldwide by 2050. This represents, according to Pimentel's estimates, a 39% increase over current bioenergy production over five decades, and only 1.25% of world energy output in 2000.

Then, there are the optimists, exemplified below:

- In 2000, an ad hoc committee of leading scientists and engineers established by the National Research Council found that "there is enough unused biomass (in the United States) to satisfy all domestic demand for organic chemicals that can be made from biological resources. . .." Furthermore, the committee proposed a goal of ". . . eventually meeting over 90% of U.S. organic chemical consumption and up to 50% of U.S. liquid fuel needs with biobased products. . . " [58].
- Johansson et al. projected aggressive scenarios for a renewable energy-intensive economy in which 40% of direct fuel use (for everything except power generation) comes from renewables, mostly in the form of biomass, by 2050. Direct use of biofuels exceeds direct use of oil in their scenario [59].

Quantifying the controversial and uncertain factors that go into estimating biomass energy potential can shed light on these apparently contradictory findings. Lynd et al. [60] developed a generic equation to describe "*R*," the ratio of land required for meeting U.S. transportation energy demand with biofuels to the land available for biofuels production. A simplified version is shown below.

$$R = \frac{[\text{Energy Demand}] \times [\text{Land Required per Delivered Energy}]}{[\text{Total Land} - \text{Land Required for Food Production}]}$$

Accounting for all of the factors that go into energy demand (population, vehicle miles traveled per capita, vehicle efficiency) and land required for energy production (biomass land yields, biomass conversion yields, etc.), they found that the numerator in this equation can vary by a factor of 400 for a range of possible high and low values. The denominator can vary tenfold for a range of possible high and low values. This explains the wide diversity of conclusions that has been published regarding the potential for biomass to really affect our energy problem. More important, it points out the need for conducting more rigorous assessments of all these factors. Some of the factors involve social choices, whereas others involve prognostications on future technology developments.

In 2004, several new studies appeared that reconsidered the role of biomass in a future sustainable energy supply. Each has looked at some or all of the above factors, with an eye toward honing in on a more rational assessment of biomass energy supply potential. The Rocky Mountain Institute completed a major study seeking ways of eliminating the United States' dependence on oil. Among five major strategies that, taken together do just that, was an investment in research, development, and deployment of biofuels. They estimated that such a strategy could lead to a 25% reduction in our dependence on oil in the year 2025. This corresponds to around 57 billion gallons of gasoline equivalent supplied in the form of ethanol [61]. The Natural Resources Defense Council estimated in its recent assessment of biofuels that around 30% of total transportation energy demand could be met with ethanol from lignocellulosic biomass by 2050 [62]. This corresponds to about 130 billion gallons per of gasoline equivalent replaced with ethanol.

What sets these reports above others that have been published over the past two decades? First, the organizations themselves are different. They are both nongovernmental organizations (NGOs) that have never been particular advocates of biofuels. That is, they bring the perspective of a "disinterested party" to the table. Second, each looked at biofuels in a rational, future-oriented way. Third, they avoided the "single solution" trap so often found among technology advocates. Their results reflect this. Biomass comes forward as only part of the solution. Gone are the days of single energy solutions, or at least so it would seem from their analyses. Oil was indeed a one-time wonder.

$$R = \frac{[\text{Energy Demand}] \times [\text{Land Required per Delivered Energy}]}{[\text{Total Land} - \text{Land Required for Food Production}]}$$

Meanwhile, a third report has been issued in 2005 that looks at the potential supply of biomass from a different point of view. It comes from the U.S. Department of Energy's Office of the Biomass Program and from three offices within the U.S. Department of Agriculture (USDA). These are certainly not "disinterested parties." They are the key organizations in the federal government charged with looking at biomass and the role that agriculture can play in a sustainable energy future. So, it is not surprising that they turned the question of biomass supply on its head. Their starting question: are there a billion tons of biomass out there for conversion to energy and fuels? Or, put a little differently, could biomass replace 30% of today's petroleum demand? The motivation for the study was, quite simply, to see if we could prove that there is enough potential biomass to make energy companies and others take a second look at biomass. Figure 33.12 summarizes their findings. They found 1.3 billion tons of plant biomass for the taking each year, or around 80–100 billion gallons per year of gasoline equivalent, depending on the efficiency of the biomass-to-fuels conversion.

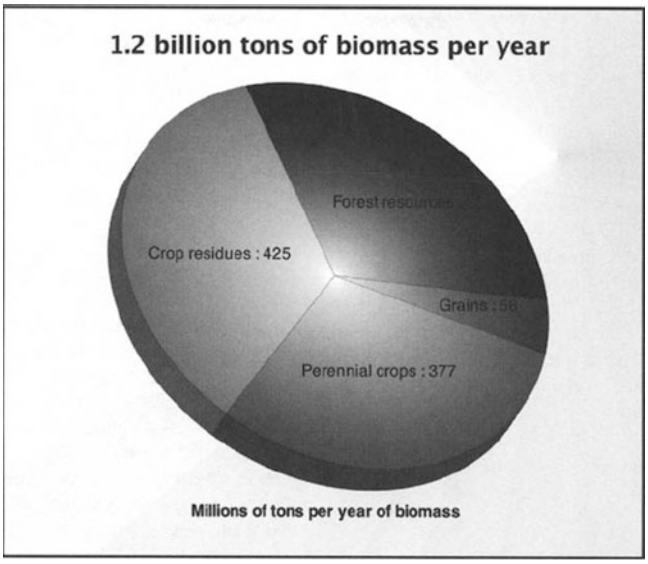


Fig. 33.12 Biomass feedstock potential finds from billion-ton vision study

What can we conclude thus far about the role that biomass can play as a substantial source for our future energy supply? It's worth aiming at. Competing with petroleum-derived energy is still a stretch, but there is still potential for improvement in the technology. As we move forward, we may find disappointments about biomass technology that force us to turn in other directions.

Microbial Bioethanol Production

The primary carbohydrate components of lignocellulosic biomass consist of D-glucose, D-xylose, L-arabinose, D-galactose, and D-mannose. Glucose (from cellulose) and xylose (from hemicellulose) are the two principal carbohydrates present in most biomass feedstocks. The levels of the minor carbohydrates L-arabinose, D-galactose, and D-mannose (also derived from hemicellulose) vary considerably with biomass type. Softwoods typically contain more galactose and mannose than hardwoods, whereas hardwoods, herbaceous plants, and agricultural residues generally contain higher levels of arabinose and xylose. In some herbaceous crops and agricultural residues, arabinose levels are high enough that conversion of arabinose (in addition to glucose and xylose) is required to achieve overall economic viability.

Ethanol Fermentation Schemes

Conversion efficiency and robust fermentation of mixed-sugar lignocellulose-derived hydrolysates are critical for

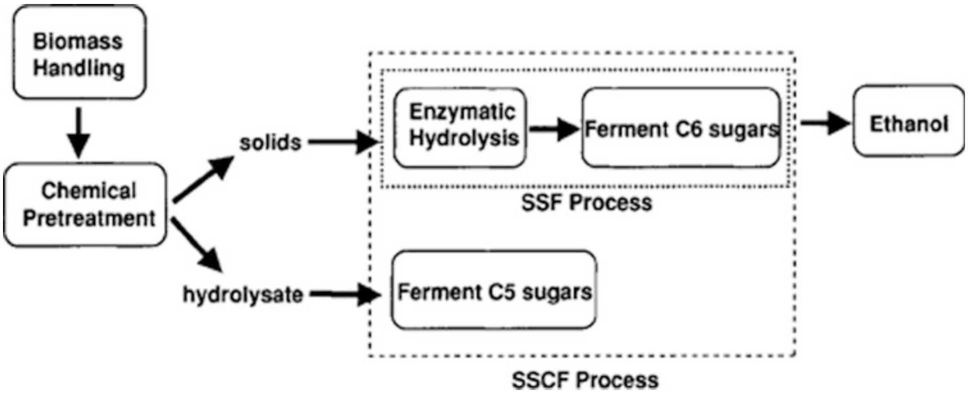


Fig. 33.13 Simultaneous saccharification fermentation (SSF) and simultaneous saccharification and co-fermentation (SSCF) concepts

producing ethanol at low cost to realize a commercially viable biorefinery. Biomass sugars are typically released by thermochemical pretreatment followed by enzymatic hydrolysis of chopped or milled biomass. The pretreated soluble fraction of biomass is called the "hydrolysate" and the hydrolysate containing the insoluble material is referred to as the "slurry." In diluted acid pretreatment, most of the hemicellulosic sugars (xylose, arabinose, galactose, and mannose) are solubilized; however, the glucose component remains in the solid form as cellulose, where it is later depolymerized by cellulases. This step is often combined with the subsequent microbial fermentation of the sugars to relieve the product inhibition of cellulases, the so-called simultaneous saccharification and fermentation (SSF) process. A process based on the fermentation of pentose sugars (derived from the hydrolysate) combined with the saccharification of cellulose and fermentation of glucose (derived from simultaneous enzymatic saccharification) is referred to as a simultaneous saccharification and cofermentation (SSCF; Fig. 33.13). To be successful, this scheme requires that the microorganisms are capable of fermenting hexose and pentose sugars equally well, though not necessarily simultaneously. Alternatively, a hybrid process with partial enzymatic hydrolysis (to obtain high cellulose hydrolysis rate by operating at high temperature) and co-fermentation may be used to achieve high overall conversion rates of biomass sugars to ethanol. Additionally, microorganisms are often susceptible to inhibitors, such as acetic acid, furfural, and phenolic compounds librated from lignocellulose during chemical pretreatment [63, 64]. Because of this, a detoxification step, such as the "over-lime process" is generally applied to reduce the toxicity of the hydrolysate. Alternatively, other adsorbants may be used or the microbes themselves can be adapted or engineered for increased tolerance to some of these inhibitors or the pretreatment process can be tweaked to minimize their production.

Although a number of microorganisms can efficiently ferment glucose to ethanol, only recently has conversion of the pentose sugars in the hemicellulosic fraction become feasible [65]. The few organisms that were known to utilize either D-xylose or L-arabinose typically grow slowly on pentoses and achieve relatively low ethanol yields and productivities [66]. Because of this, the identification and development of microorganisms capable of selectively converting D-glucose, D-xylose, and L-arabinose to ethanol at high yield has been the focus of extensive research during the past 10–15 years. In the past decade, the sophistication of molecular biology has grown tremendously and numerous attempts have been made to use recombinant DNA technologies to engineer superior microorganisms for bioethanol production. Only a few of these efforts have been provisionally successful and considerable work is yet to be done.

Metabolic Pathway Engineering

Metabolic pathway engineering is increasingly recognized as a powerful approach for developing microorganisms capable of efficiently converting biomass sugars to ethanol. In broad terms, superior ethanol-producing microorganisms can be developed by either of these metabolic engineering approaches:

1. Broadening the substrate range to include biomass sugars (e.g., xylose, arabinose, galactose, mannose) in strains exhibiting good product selectivity, but not capable of fermenting sugars other than glucose to ethanol
2. Increasing ethanol product selectivity in strains exhibiting broad substrate range

Of course, beyond these two basic approaches, significant metabolic engineering may also be required to stabilize "improved" strains or to enable such strains to achieve high ethanol yields and fermentation productivities.

Following the first approach, *Escherichia coli* and *Klebsiella oxytoca* have been engineered to be highly effective ethanol producers by introducing the genes for ethanol production from *Zymomonas mobilis* [67–69]. Extensive evaluation of these "ethanologenic" strains have been carried out, both in media containing pure sugars and in pretreatment hydrolysates derived from a variety of feedstocks [68–73].

The second approach, broadening the substrate utilization range of strains that are highly efficient ethanol producers, has been demonstrated by introducing the xylose assimilation and pentose phosphate pathway genes from *E. coli* into *Z. mobilis*. The result was an engineered *Z. mobilis* strain able to ferment xylose to ethanol at high yield [74–76]. An arabinose-fermenting *Z. mobilis* strain was also developed by introducing the arabinose assimilation and pentose phosphate pathway genes from *E. coli* into *Z. mobilis* [77].

More recently, a long-term effort to develop xylose-fermenting *Saccharomyces* sp. has also been successful. For example, xylose fermentation was reported for *Saccharomyces* strains transformed with the xylose reductase and xylitol dehydrogenase genes from *Pichia stipitis* (additionally, overexpression of native xylulokinase was found useful) [78, 79]. An effort to introduce xylose isomerase from *Pimicus* into *Saccharomyces cerevisiae* was also demonstrated [80–82]. Other noteworthy achievements in the metabolic engineering of superior ethanol producers include initial success in improving the performance of xylose-fermenting yeasts by optimizing the expression of genes encoding the xylose assimilation and ethanol production pathways [83, 84]. Successful transformation of the pentose-fermenting *Clostridium thermosaccharolyticum* has also been reported [85], providing a key tool for further developing this microorganism by altering product selectivity to favor ethanol production. Since the previous edition of this chapter, consolidated bioprocessing research has expanded dramatically, including metabolic and enzyme expression engineering of single microbes as well as cooperative multi-microbe systems. Furthermore, recombinant *E. coli*, *K. oxytoca*, *Z. mobilis*, and *Saccharomyces* were reported to be capable of anaerobically fermenting arabinose to ethanol [77, 86, 87], unlike the wild-type xylose-fermenting yeasts, such as *P. stipitis*. These yeasts can grow on arabinose aerobically, but cannot ferment arabinose anaerobically.

Performance Assessment

Efficiency of microorganisms for conversion of biomass to ethanol can be evaluated by three most critical performance factors: yield, productivity, and final product concentrations. Achieving high ethanol yield is the most important factor for the biorefinery, because the cost of feedstocks can be as high as 40% of the process cost [88]. Ethanol yield can be referred to either as metabolic yield or process yield. Metabolic yield is calculated as ethanol produced based on sugars consumed, which provides an indication whether microorganisms produce ethanol selectively. The maximum metabolic yields for both hexoses and pentoses are 0.51 g ethanol per gram sugars used. Formation of byproducts, such as lactic acid, glycerol, and acetic acid reduces the metabolic ethanol yield, consequently reduces process yield as well. Process yields calculated as "ethanol produced based on total sugars available" provides information not only how efficient the microorganisms can produce ethanol from the sugars, but also the degree to which the microorganisms are capable of utilizing all the sugars available. This parameter is especially important when considering high biomass sugar streams. Similarly, specific ethanol productivity, calculated as ethanol produced per gram of cell biomass per hour, reflects the effectiveness of the catalytic capability of cells. Volumetric ethanol productivity is used to reflect efficiency of the overall process. Kinetic parameters from various recombinant microorganisms have been summarized by Dien et al. [65].

Future Directions

Despite initial success in demonstrating microorganisms capable of fermenting biomass sugars, there is currently a dearth of fermentative microorganisms with the capability to efficiently convert all five biomass sugars in high ethanol yield and productivity under relevant industrial processing conditions. These processes demand robust performance at low pH and high temperature, as well as a high tolerance to ethanol. In most cases, the pentose utilization rate is at least several-fold lower than that of glucose. Therefore, the ethanol yield from pentoses is significantly lower compared to that of glucose [89].

Unlike the starch-based glucose streams, hydrolysates derived from lignocellulosic feedstocks can contain many toxic compounds that inhibit microbial growth and fermentation [88]. Improving our understanding of inhibition mechanisms and microbial physiology during hydrolysate fermentations will require full use of the advanced analytical and "omics" metabolic engineering and modeling tools recently made available. This approach will greatly enhance our capability to develop a new class of robust industrial microorganisms capable of efficiently and productively converting all biomass sugars to ethanol under "dirty" industrial processing conditions. Other important considerations for commercial viable microorganisms are hydrolysate tolerance and media requirements.

Futuristic process scenarios have been proposed that combine key process steps, thus reducing overall process complexity and cost. One notable example is the consolidated biomass processing (CBP) technology proposed by Zhang and Lynd [90] for the *Clostridium thermocellum* case. Their work reminds us that *C. thermocellum* hydrolyzes cellulose by a different mode of action compared to the classical mechanism associated with fungal-derived cellulases, the "cellulosome." Furthermore, for *C. thermocellum*, the bioenergetic benefits specific to growth on cellulose are result from the efficiency of oligosaccharide uptake combined with intracellular phosphorolytic cleavage of β-glucosidic bonds, another pathway not known in fungi. Zhang and Lynd believe that these benefits exceed the bioenergetic cost of cellulase synthesis, supporting the feasibility of anaerobic processing of cellulosic biomass without added saccharolytic enzymes. Another option for CBP is to enable yeast, already ethanologenic, to produce cellulases [91]. In this case, expression of some active and effective cellulases from yeast has proven challenging [92]; however, endoglucanases and beta-glucosidases appear more amenable to yeast processing [93].

Biomass Analysis and Compositional Variability

Introduction

The Importance of Reliable Compositional Methods for Biomass Analysis

The ability to rapidly and inexpensively obtain an accurate chemical composition of complex biomass feedstocks and biomass-derived materials is a key element in enabling commercialization of processes that convert biomass to fuels and valuable chemicals. Robust analytical methods are needed to improve our understanding of and our ability to economically control biomass conversion processes. Additional challenges face these processes because of the heterogeneity that is an inherent property of biomass. The chemical composition of a biomass feedstock varies as a function of many factors, including plant genetics, growth environment, harvesting method, and storage conditions. Many biomass conversion feedstocks are residues of another process. For example, bagasse is a byproduct of sugar production from sugarcane. In these situations, the varying efficiency in the original process can impart an additional source of compositional variance in biomass feedstocks. All of these sources of compositional variance are difficult if not impossible to control. However, the composition of a given feedstock can be measured at any point and that information can be used to adjust process conditions for optimal conversion or steady-state production. The rapid, inexpensive compositional analysis methods described in this chapter are examples of the types of new tools that will be needed for the commercialization of processes that convert biomass into fuels and valuable chemicals.

As illustrated in Fig. 33.14, biomass feedstocks can vary widely in the number of constituents and the concentration of each constituent. In biomass conversion processes, up to 20 constituents may need to be monitored to characterize the conversion of feedstock into a desired product or products. Standard wet chemical methods for the chemical characterization of biomass feedstocks and biomass-derived materials have been validated through the International Energy Agency and are available from the American Society for Testing and Materials (ASTM) [94]. In addition, the National Renewable Energy Laboratory [95] (NREL) has developed and validated a collection of standard laboratory analytical procedures specifically for the compositional analysis of biomass including, but going beyond those of the ASTM. These wet chemical methods of analysis are based on the fractionation of the biomass sample and the isolation of purified fractions that can be quantified using conventional analytical instruments [96]. These methods are primarily used in feedstock-specific portfolios containing analysis methods for each of the relevant constituents.

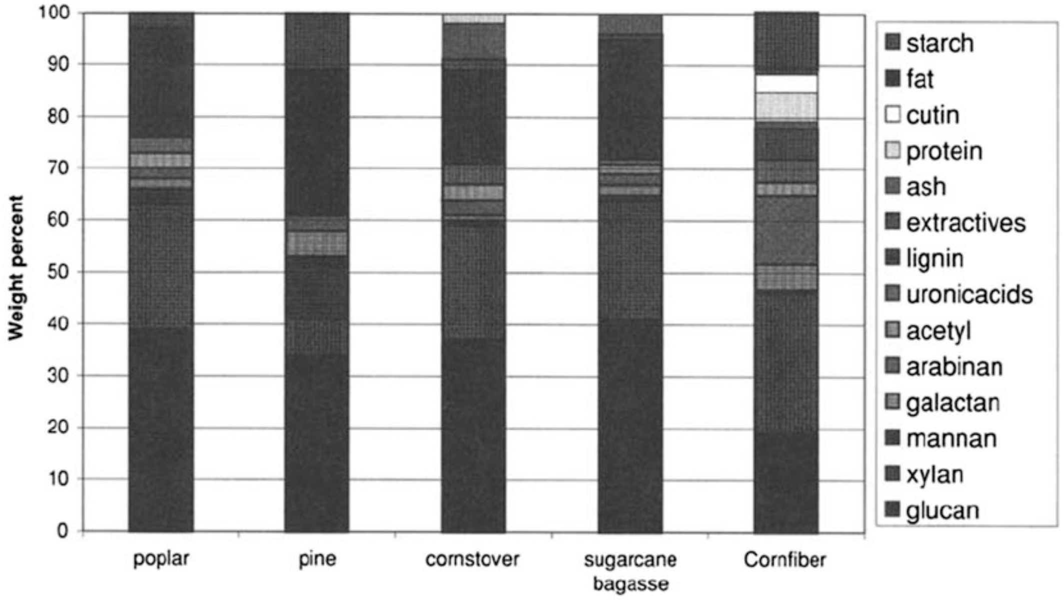


Fig. 33.14 Chemical composition of five commercial biomass feedstocks showing the variance in number and concentration of constituents

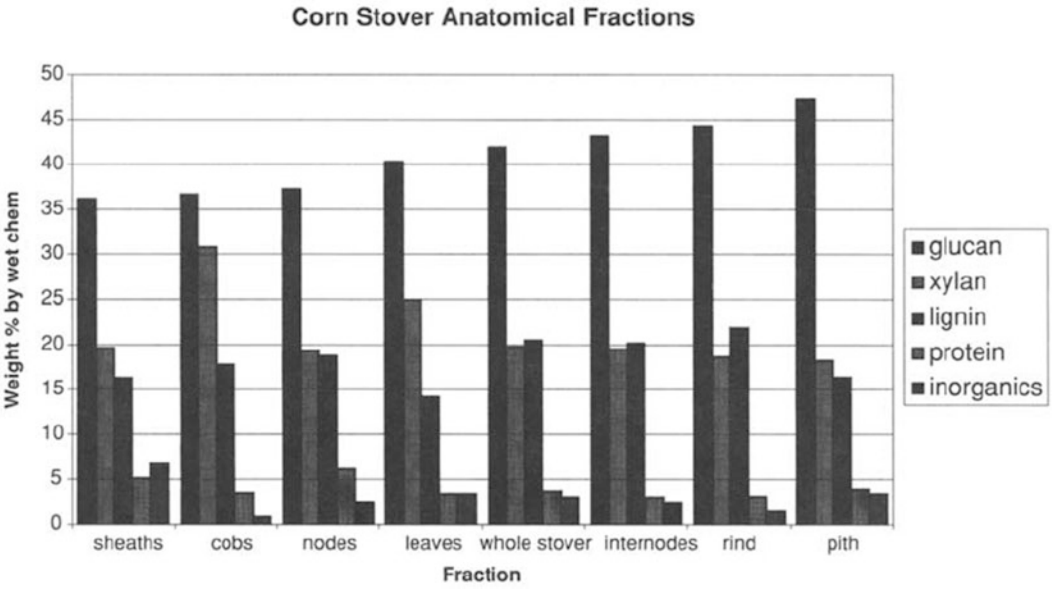


Fig. 33.15 Compositional variability of corn stover tissues from a single hybrid grown in a single field over a single growing season

In most cases, these portfolios enable the identification and quantification of greater than 95% of the dry mass of biomass feedstock and biomass-derived materials.

The Need for Accurate, Real-Time Biomass Analysis Methods

Standard wet chemical methods, although accurate and robust, are not applicable in a commercial setting, because they are very expensive (labor intensive) and cannot provide the analysis information in a timeframe useful for process control. For example, a complete analysis using standard wet chemical methods costs $800–$2,000 per sample and the results are typically not available for days, sometimes weeks. In contrast, new methods are being developed that can perform the same analysis for about $20 per sample and provide results in a timeframe relevant for process control, meaning that the information can be used to make the process adjustments necessary for steady-state production. One approach to reducing the time and cost of compositional analysis is the development of rapid analysis methods that use multivariate statistical analysis software to extract chemical information from easily obtained spectroscopic data. Rapid analysis methods match the precision and accuracy of their calibration methods, so the savings are obtained without loss of precision or accuracy [97]. New techniques, such as rapid analysis, are needed to provide analytical support for large-scale processes that convert biomass to fuels and chemicals.

Heterogeneity and Biomass Analysis

As stated above, plant cell walls are comprised of three-dimensional complexes of natural polymer matrices including cellulose, hemicellulose, lignin, and in some cases protein and silica. These polymer matrices are not uniformly distributed within the plant cell wall, and their relative concentrations change from one morphological region to another, reflecting physiological function. The difference in chemical composition between the tissues within a plant is, in nearly all cases, greater than the variation seen in any one tissue across the entire plant species. An example of this compositional variance is shown in Fig. 33.15 for corn stover from Pioneer 33J56. The concentration of five major constituents is shown for seven tissue types as well as an average composition for the whole plant. The various fractions are according to increasing glucan content. The independent variance of the concentration of various polymer constituents can easily be seen.

Much of the compositional variance in a feedstock can be explained in terms of varying proportions of the tissue types. These proportions can vary by variety, growth environment, harvesting method, and storage conditions. The differences in chemical composition imply structural differences that could also affect conversion efficiency. To understand biomass compositional variability and its implications, it is essential to appreciate the magnitude of variance in available feedstocks and to understand the controlling factors. Appreciating the sources of variance permits the formation of testable strategies for controlling or limiting the sources of variance. This strategy is particularly important in systems where process economics are driven primarily by product yield, as is the case for biomass conversion processes.

Sources of variance. It is well established in the agronomy literature that factors affecting plant characteristics and crop

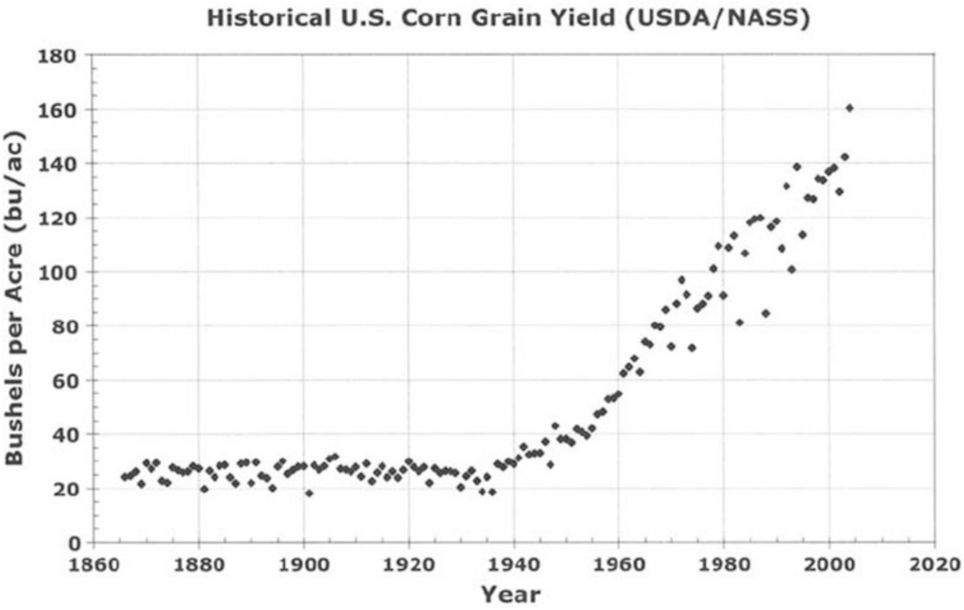


Fig. 33.16 Historical average U.S. corn gain yields in bushels/acre (data from USDA/NASS Web Site)

performance fall into three general categories: genetic, environmental, and interactions between these two classes. In other words, the phenotype (i.e., any measurable characteristic) of an individual plant is the product of its genotype (i.e., the complete set of genes inherited by an individual from its parents) as influenced by the environment in which that individual exists. For example, the same commercial hybrid varieties of corn often have statistically different yields when grown in different locations (i.e., environments). This may be due to differences between locations in weather patterns, soil types, agronomic practice, or other factors. Similarly, genetically distinct commercial hybrid varieties of corn grown in the same environment often have different grain yields (see results from the Wisconsin Corn Hybrid Trials [98] or the Minnesota Corn Variety Trials [99]). An interaction between these two broad classes of variables is said to occur when different genotypes respond in different ways to a change in environmental conditions.

Agricultural systems are notorious for their variability (e.g., harvest yield) from year to year, location to location, and variety to variety. Characteristics other than yield (e.g., plant height, pigmentation patterns, cell wall composition, nutritional quality, resistance to diseases, etc.) are also influenced by some combination of these and other factors. Each phenotype is most likely influenced by different combinations of genetic and environmental factors.

The analytical methods used to determine the composition of biomass materials are also a source of some variance in this kind of analysis. It must be determined whether variance due to sampling and measurement techniques is significant in light of variance from genetic and environmental sources.

Genetic Factors That May Contribute to Cell Wall Compositional Variability

Modern corn breeding dates back to the early 1900s with the work of Shull [100], East [101], and others. Increased grain yield (i.e., productivity) has long been the main objective of corn breeding programs in the United States and elsewhere. Corn breeding programs have produced remarkable results, evident in the historical corn grain yield data for the United States (available online from the USDA National Agricultural Statistics Service Web site [102]) and presented here as Fig. 33.16. It should be realized that the data in Fig. 33.16 reflect both genetic and agronomic improvements made in parallel.

Hallauer et al. [103] present a summary of 13 separate studies that attempt to estimate the fraction of observed corn grain yield improvement due to genetic gains. These estimates range widely (33–89%), but it is clear that genetics is responsible for a significant proportion of yield improvement during the past 100 years. The remaining fraction of yield gains is due to changes in cultural practice and crop management (e.g., fertilizer use, higher density of planting, pesticide availability and use, mechanized harvesting).

Peterson states [104], “Breeders’ efforts with maize have uncovered a highly heterogeneous genotype that can be manipulated in most directions to achieve the desired goal.” In addition to the overall goal of increased yield, this includes traits such as lodging (i.e., ability to stand), resistance to various biotic and abiotic stresses, male sterility, the angle at which leaves are held relative to the stalk, days to grain maturity, nitrogen-use efficiency, seed composition (i.e., starch, protein, and lipid content), and many others.

Studies relating to the genetic manipulation of nutritional quality (i.e., chemical composition) of corn grain are analogous to this discussion regarding manipulation of stover (i.e., cell wall) chemical composition. The following excellent example is used to illustrate what might be accomplished with regard to cell-wall composition using genetic techniques in maize.

The concentration of oil in the kernel varies widely in corn as a species, but corn belt inbred varieties range only from 2.5 to 5.5%. Commercial hybrids have an even narrower range, typically only 4–5%. An assortment of breeding strategies starting with out-crossing populations, in which the gene pool is broader than it is for inbred and hybrid cultivars, has altered the oil content of kernels from about 4% up to more than 21% over a series of selection cycles [105–108].

A similar story can be told regarding corn kernel protein content. Hybrid corn typically contains 8–11% protein, but considerable genetic variability for this trait exists in open-pollinated populations. In an experiment involving serial selection for both high and low protein content, mean protein content has been shifted more than 139% from the starting point mean value (a range of at least 20 standard deviations) [105–107]. At generation 70 (30 years ago!) in the Illinois Long-Term Selection Experiment, the Illinois High-Protein (IHP) line contained 26.1% protein and the Illinois Low-Protein (ILP) line contained 5.8% protein. These same lines were simultaneously the low-starch (44%) and high-starch (74.5%) lines, respectively [109]. Thus, at least in corn, selection for increased protein content in grain seems to occur mainly at the expense of starch.

Three plant genomes have now been completely sequenced, and several more are at various stages in the process. From DNA sequence data, it is estimated that the *Arabidopsis* genome contains around 25,500 genes [110]. The rice genome is estimated to contain somewhere between 32,000 and 55,000 genes [111, 112]. The *Populus trichocarpa* (black cottonwood), genome is estimated to contain about 58,000 genes [113]. To date, sequencing efforts in *Zea mays* (i.e., maize; corn) indicate that its genome also contains about 58,000 genes [114].

Cell wall biogenesis during cell growth and differentiation involves many different enzyme activities and perhaps several thousand genes. It has been estimated that 15% of the *Arabidopsis* genome (i.e., more than 3,800 genes) may be dedicated to cell-wall biogenesis and modification [115]. Although only a very few of these genes have been identified and characterized to date, several efforts are underway to do so (for example refs. [116–118]). Because multiple alleles (i.e., gene variants) probably exist for each of these genes, the possible number of permutations of genes that affect cell walls in a single species is truly enormous.

According to Carpita [115, 116], genes involved in plant cell-wall biogenesis fall into one of six functional categories:

- The synthesis of monomer building blocks of cell-wall polymers, such as nucleotide sugars and monolignols
- The biosynthesis of oligomers and polysaccharides at the plasma membrane and ER-Golgi apparatus (i.e., polysaccharide synthases and glycosyl transferases)
- The targeting and secretion of Golgiderived materials
- The assembly and architectural patterning of polymers (includes glycosyl hydrolases and structural proteins)
- The dynamic rearrangement of wall polymers during cell growth and differentiation (includes lignification)
- Signaling and response mechanisms

Although the metabolic pathways associated with the first category have been pretty well elucidated at this point, the enzymes involved in the second (cellulose synthases, some golgi-based synthases (e.g., mannans, xyloglucans, pectins) and third categories are largely uncharacterized. Likewise, most of the corresponding genes are also unknown. Significant progress has been made in some of these areas in recent years, however [119–123].

Once identified, those genetic factors that exert a large influence on cell-wall composition, cell-wall architecture, or biomass conversion process performance characteristics could become targets of a plant breeding strategy to manipulate these characteristics, while maintaining or increasing traditional crop yield or overall agronomics. To get the most from such a breeding program, it will be essential to monitor the effect of a wide range of environmental variables on genetic potential during the breeding process.

Environmental Factors That May Contribute to Variance

Environmental factors that affect plant phenotypes fall into two categories: namely, those that are at least partially under control of the farmer and those that are not. Factors that are largely not under human control include weather patterns (day length, solar flux, temperature, and precipitation), soil type and pH, and soil mineral content, and are not discussed further. A partial list of factors that can be addressed by the farmer include selection of hybrid variety, planting date, harvest date, tillage practice, irrigation, fertilizer types and amounts, pesticide and herbicide usage, and the strategy employed for harvesting, transporting, and storing crop residues. Soil pH is at least partially under control of the farmer. Producers commonly add lime to soil to make it less acidic. Which of these factors are most influential across the broadest range of situations with regard to stover compositional variability is an open question and will require multiyear, multilocation field trials to determine with confidence.

Most important are likely environmental factors associated with cultivation, crop management, and postharvest processes that can be controlled to some extent and on the other hand also have a large impact on the chemical composition, cell-wall architecture, or conversion processing behavior of plant biomass. These factors could form the basis of a testable strategy to decrease the variance associated with these characteristics.

Table 33.1 Confidence intervals for NREL and ASTM standard wet chemical methods

Constituent	Method variance absolute error 99% confidence
Glucan	1.5
Starch	1.0
Xylan	1.5
Arabinan	1.5
Mannan	1.5
Galactan	1.5
Uronic acids	1.5
Pection	1.5
O-acyl groups	0.5
Structural inorganic	0.5
Sucrose	1.5
Protein	1.0
Lignin	1.0
Ferulic acid esters	1.0
Extractives	1.0
Phytate	0.5
Cutin	1.5

Analytical Variance

A complete analysis characterizing greater than 95% of a biomass sample can require up to 30 independent measurements to report the concentration of 10–15 constituents. Table 33.1 shows the confidence intervals for the standard wet chemical methods used to characterize biomass feedstocks. The confidence interval of 1.5% for glucan means that when a value of 38% is reported for a biomass sample, there is a 98% probability that the "true" glucan concentration lies between 36.5 and 39.5%. Many of these values were determined in an international round robin evaluation of four biomass standard reference materials available from the National Institute of Standards Technology (NIST) [124].

Portfolio Methods

In the past, efforts were made to develop universal methods of analysis that could be used for any type of biomass or biomass-derived material [125]. These methods provide instruction for the various procedural steps, but they rarely discuss why a given step is included in the procedure. The assumption is made that all steps in a standard procedure will be followed as written without omission or amendments. When combined for a complete mass analysis, many analytical methods were compromised to the extent that they were adequate for most samples, but optimal for none. As the biomass conversion industry matures, the need for more accurate and precise methods will increase, as data generated from these methods will be used to evaluate a maturing state of technology for commercialization. In response to this need, newer methods of analysis are being developed in a modular fashion and designed to be incorporated into sample-specific portfolios capable of providing a customized total analysis of any biomass sample. In the portfolio, the individual constituent methods are independently validated relative to standard reference materials. These methods are then validated a second time when the results are combined to reconstruct the entire chemical composition of a biomass sample. Typical mass closures between 96 and 104% confirm that no major components have been overlooked and interferences between methods and double counting of materials are minimized.

Wet Chemical Methods

Within each portfolio, the sequence of application of the various analytical methods is critical, as many methods have been validated with the assumption that all prior steps have been followed. For feedstocks, the sequence is sample preparation, then removal of nonstructural materials through solvent extraction, followed by hydrolysis of the structural polymers to simple forms for chromatographic or spectroscopic analysis. In the past, many methods for biomass analysis have been "behavior based," meaning they define chemical structure based on solubility in certain solvent systems. The newest methods incorporate modern analytical instruments, which allow the various structural components to be measured and tracked based on chemical structure.

Before using biomass analysis methods, analysts are encouraged to review the literature that supports the standard methods [125]. Understanding the science behind the methods helps the analyst understand the ways that the various procedures are connected, the limits within which a step can be altered or modified, and the potential downstream consequences of omitting any given step. With this in mind, the following several paragraphs outline available information concerning the significance, implications, and limitations of each step in a complete characterization of biomass. More detail is available in the open literature and the text of the individual procedures.

Sample preparation. Sample preparation is an important but often overlooked part of biomass analysis. Biomass sample preparation must take several factors into account. The object of sample preparation is to enable the analysis to be done on a small scale while ensuring that the analysis of the small sample provides valid information about the larger bulk sample. Typically, biomass sample preparation includes drying, milling, sieving, and homogenizing. The overarching objective is usually obtaining a representative sample that can be analyzed on a reasonable scale.

Drying. Biomass feedstock samples can contain as much as 60% (w/w) water. If the biomass has been degraded in a manner that disrupts the ultrastructure of the plant (chemical treatment or enzymatic digestion), the moisture of the sample could be even higher. Water in a biomass sample, if high enough, can introduce errors during some of the analysis steps by diluting reagents. Moisture in samples can also cause sample fractionation during milling and sieving. Biomass samples with high moisture contents are still vulnerable to biological degradation and may not remain stable and uniform for the duration of a battery of tests. For these reasons, biomass samples are usually dried until the moisture content is less than 15%. Because some of the constituents are heat sensitive or volatile, certain protocols must be followed when drying a biomass sample for compositional analysis, to avoid sample loss or degradation. Air-drying is preferred if time, space, and ambient humidity will allow the sample to reach a moisture level below 15%. Drying the samples in a convection oven or vacuum oven where the temperature is not allowed to exceed 45°C is an acceptable alternative. When freezing does not cause collapse of the cellular structure of the biomass, samples can also be safely lyophilized, or freeze-dried. When drying biomass, one must also consider the end use of the material. If the material is to be used in enzyme hydrolysis or microbial conversion experiments, then the method of drying becomes doubly important. Complete removal of water or removal of water without structurally stabilizing the pore structure of the cell walls (as in lyophilization) can lead to a collapse of the cellulose microfibrillar network; a process referred to as "hornification". This process makes the cellulose more difficult to digest by cellulases and perhaps more difficult to disrupt by pretreatment catalysts. As it not easy to measure or prevent, the best way to deal with hornification is to dry the materials in as consistent and process-relevant manner as reasonably feasible.

Washing. If the biomass sample is contaminated with soil, the samples should be washed if possible before further processing. Soils vary considerably in chemical content and even a few weight percent of extraneous inorganic compounds can neutralize reagents, catalyze side reactions or affect subsequent analysis in ways that are difficult to observe or control. Washing may be as simple as agitation in clean demineralized water or may require small amounts of mild detergents to be effective. After washing, the sample must be carefully dried before analysis.

Milling. Size reduction is an important step in biomass analysis because standard methods are optimized and validated for materials with a specific particle size. The particle size will affect hydrolysis or digestion rates as well as rates of degradation and should be specified in any standard procedure. As a general rule, the sample to be analyzed should contain at least 100 particles of the biomass sample in order to be representative of the bulk material. For reasons described earlier for drying protocols, care must be taken during the milling process to avoid heating the biomass sample. This is particularly crucial if the biomass must be reduced to a fine powder. Cryo-mills, where the milling apparatus is submerged in liquid nitrogen, or mills that are jacketed for the circulation of cooling solutions are most suitable for milling in these situations. Knife mills such as the standard Thomas-Wiley Mill 4 (Thomas Scientific Model 3375E15 or equivalent) work well for most woody biomass samples. Less dense materials such as corn stalks or straws may be easier to feed into a hammer mill. Most knife and hammer mills will retain and reprocess materials until they pass through a screen of a particular particle size, allowing the user to specify the maximum particle size of the processed sample.

Sieving. All forms of chemical or thermal hydrolysis perform better when the particles are of a uniform size. For this reason, most standard procedures usually specify both a maximum and a minimum particle size. To accomplish this, the milled biomass is sieved through a set of standard screens. To prevent errors associated with incomplete hydrolysis, the oversized material is milled again until it passes through the largest screen. The material that passes through the larger screen but is retained on the fine mesh screen is of uniform particle size and can be used for compositional analysis. The material that passes through the fine mesh screen may be hydrolyzed too quickly leaving the hydrolysis products more susceptible to side reactions or degradation before analysis. Because the fine material cannot be reprocessed to increase particle size, this material must be discarded. The fine fraction should always be removed from the analysis if it has significantly higher ash content than the uniformly sized material. This elevated inorganic content may indicate the presence of non-biomass contaminants (typically soil).

The impact of extraneous inorganic matter on the analysis cannot be predicted. Most chemical analysis methods have been optimized for bark-free wood that has an ash content of less than 2%. These methods may not be applicable to herbaceous materials where structural silica can increase the ash content to more than 15% of the dry weight of the biomass. One potential problem with high ash samples is that inorganic materials, depending on their composition, may neutralize the sulfuric acid solutions used in the hydrolysis steps. The affect of pH changes on the hydrolysis of the carbohydrate polymers is unknown, but could result in incomplete hydrolysis of some carbohydrate structures. It is also possible that soil components such as iron and manganese could catalyze undesired side reactions, which produce products that would not be identified and quantified with the current analytical methods. However, although the impact of soil contamination cannot be easily determined, removal of the −80 mesh fraction can improve the chemical analysis by reducing the ash contents of samples to about the level of structural inorganics expected in that type of biomass. The fine material is weighed, reduced to ash to determine the percentage weight of inorganic material and then discarded. The analysis is performed only on the intermediate, uniformly sized material. The composition of the whole sample can be calculated by assuming that the biomass portion of the fines has the same chemical composition as the fraction that is analyzed. Although this assumption is not completely accurate, the precision gained by removing the fine material more than compensates for the error introduced by this assumption. If the sample has not been washed, removing the very fine material (less than 80 mesh) will remove most of the dirt entrained in the sample.

Making the particle size as uniform as possible also makes it easier for the analyst to select a small but reproducible sample for analysis. In some biomass samples, however, sieving to remove fine material chemically fractionates the sample. This is known to be true with wood that contains a significant amount of bark. The bark tends to granulate during milling and will preferentially be removed with the fines. The sample to be analyzed would then have an artificially low bark content relative to the original sample. A similar problem has been reported with some grasses where the friable pith material may be preferentially removed as fines. These samples should not be sieved for fines removal. Samples should be examined for soil contamination, and special care should be taken to select a representative particle size distribution for analysis.

Extractives Determination

Plants store nonstructural materials in their vessels and water transport channels that can interfere with accurate chemical compositional analysis. These materials are commonly known as extractives, which may give the impression that they are a single substance or at least a few closely related substances. This is not always the case. Extractable materials include gums, resins, pitch, waxes, sterols, flavonoids, tannins, terpenes, quinones, non-structural sugars, chlorophyll, and many other minor building block reserves that vary seasonally and by biomass type. Because extractives vary so much in chemical composition, their behavior during the slate of analyses required for a compositional characterization cannot be predicted. For this reason, non-structural components should always be removed prior to analysis for lignin or carbohydrates. Failure to remove these materials can cause several problems [124]. Some of the extractives may be insoluble in acid, will precipitate, and thus will be falsely counted as lignin. Because of changes in the physical properties of the extractives, the presence of some materials will limit access to the carbohydrate polymers. If extractable ash, starch, or protein is present in the biomass sample, separate measurements must be made to remove the contribution of this material from the extractives values and to report it more accurately in the correct constituent category. Recently, it has become more apparent that starch is not always efficiently or uniformly extracted by the standard extraction procedure. For samples suspected to be high in starch, such as green-harvested herbaceous material or biomass containing seeds, it is advisable to do a separate starch determination or to remove the starch with more aggressive protocols, such as amylase treatment. Care must be taken in this approach, as not all amylase are equal and some contain cellulase and/or hemicellulase activity. Several kits are available commercially to quantify or remove starch from biomass.

Carbohydrate Determination

Most analysis methods for the determination of carbohydrates in biomass incorporate a two-stage acid hydrolysis to separate individual polymers and hydrolyze them to simple compounds that can be readily analyzed by chromatographic or spectroscopic techniques. The first stage subjects the biomass sample to a concentrated acid that disrupts the noncovalent interactions between biomass polymers. A second, more dilute stage follows, which is optimized for complete polymer hydrolysis and minimized degradation of monomeric sugars. Failure to remove non-structural materials may result in incomplete hydrolysis of the more recalcitrant carbohydrate polymers, particularly glucans and galactans. Increasing the hydrolysis severity may result in excessive degradation of the accessible sugars. Hemicellulose sugars, especially xylans and arabinose, are the most susceptible to degradation if hydrolysis severity is increased. If non-structural sugars are present in the biomass, they should be removed and quantified during the extractive procedures. If both starch and cellulose are present, they will both contribute to the measured glucose concentration after hydrolysis. An independent starch measurement allows the two polymers to be accurately quantified. Some degradation of sugars is unavoidable in the two-stage hydrolysis methods. Carbohydrate standards of similar structure and concentration should be run in parallel to assess the magnitude of this degradation and correct the reported values for this loss if necessary. While monomer sugars are often included as “sugar recovery standards”, these do not represent the true course of degradation in the process, as most of the biomass sugars start as polymers, not simple sugars. Inclusion of pure cellulose and hemicellulose of known composition can also serve as a recovery standard. Structures such as uronic acids, ferulic acid esters, and *O*-acetyl groups are substituents on the hemicellulose polymers and should be included in a complete carbohydrate analysis.

Lignin Determination

Most lignin methods are still based on behavioral definitions: lignin is most often defined as the material insoluble in acid after hydrolysis of the carbohydrate fraction. This assumption is invalid for many biomass samples, particularly herbaceous materials where protein will condense with the lignin. Structural inorganics will also partition during acid hydrolysis with an irreproducible portion remaining in the acid-insoluble residue. Additional determinations must be made on the acid-insoluble residue to convert the gravimetric measurement into a valid lignin measurement. Failure to make these corrections results in falsely high mass closures, as protein and silica would be measured in more than one constituent category. In the two-stage hydrolysis, some acid-soluble lignin may be released into the hydrolysis liquor. An accurate lignin value includes a measurement of both acid-soluble and acid-insoluble lignin. Acid-soluble lignin is traditionally measured by UV/VIS spectroscopy. In these determinations, care should be taken to ensure that the measured absorbance is made in a concentration range where detector response is known to be linear and that an appropriate wavelength and absorptivity is selected. The references standard method portfolios include options for several biomass categories.

Protein Determination

To estimate crude protein content of biomass or other materials, the nitrogen content of the material is measured by Kjeldahl or combustion methods and multiplied by a conversion factor where:

$$\text{Protein}\,(\text{wt/wt\%}) = \text{Nitrogen}\,(\text{wt/wt\%}) \\ \times \text{Nitrogen-to-Protein Conversion Factor}$$

A nitrogen-to-protein conversion factor (*N*-factor) of 6.25 is commonly used for animal feeds and other materials. The practice of using 6.25 as an *N*-factor is based on an incorrect assumption that protein in a given material contains 16% nitrogen (100/16 = 6.25) [126].

The correct *N*-factor for protein found in herbaceous biomass will likely be different than 6.25. Yet, determining a perfectly accurate *N*-factor for the complex matrices of biomass feedstock and process samples may not be possible. The challenge is then to determine the most accurate *N*-factor possible for biomass feedstock and process samples. A strategy based on the consensus in the literature has been incorporated into the portfolio methods. These methods calculate the reasonable *N*-factor upper and lower limits for a given material. The limits are calculated using data from an amino acid (AA) analysis with multiple hydrolysis times and an accurate total nitrogen analysis substantially similar to the methods described by Mossé [127].

Other Feedstock Constituents

Validated standard analytical methods are available for a variety of minor constituents or constituents specific to a small class of biomass substrates, examples being pectin, cutin, and phytate. Inclusion of these methods into the analysis portfolio increases mass closure and allows tracking of minor constituents that may be concentrated during the biomass conversion process.

Degradation Products

The liquid portion of biomass-derived process samples may also contain carbohydrate degradation products, such as 5-(hydroxymethyl)-2-furaldehyde (HMF), levulinic acid, and furfural, as well as other components of interest, such as organic acids and sugar alcohols. Portfolio methods are available for the quantitative measurement of these degradation products and byproducts of polymer hydrolysis.

Reconstructing the Composition of the Original Biomass Sample

Following the multistep sequence of the portfolio methods, the constituents' values can be reported in many forms. Because biomass polymers incorporate one molecule of water in the hydrolysis of each monomer molecule, the carbohydrate content of a feedstock may be reported as the theoretical yield of monomeric sugars for the assessment of the conversion efficiency of a biomass process. More commonly, constituent values in solid samples are reported as percent dry weight on either an extractives-free basis or as-received basis. The composition of liquid samples is reported as concentrations in grams per liter for each constituent. Three additional measurements are required to reconstruct the composition of biomass process slurries: total slurry weight, liquor density, and fraction insoluble solids (FIS). The total slurry weight minus the weight of insoluble solids provides the weight of the liquid portion of the sample. Dividing by the density measurement converts this to volume. With the volume measurement, constituent concentrations can be converted to recovered weight for the determination of constituent and process mass closure.

Rapid analytical methods. One approach to industrial biomass analysis is to use published, standard methods to calibrate rapid, inexpensive spectroscopic techniques, which can then be used for feedstock and process analysis. Rapid analysis is a generic term for methods that couple traditional wet chemical methods of analysis with rapid, inexpensive spectroscopic techniques. The methods described here are often further classified as rapid biomass analysis methods. Although many spectroscopic techniques can be incorporated into rapid analysis methods, for biomass analysis, near infrared (NIR) spectroscopy has several advantages over alternate spectroscopic tools. Techniques have been developed for obtaining quality NIR spectra from bulk samples, minimizing the time and expense of sample preparation. Robust NIR instruments are commercially available for process analysis, process control, and field applications. Many of the industrial NIR techniques have been developed for use by the chemical, food processing, and agriculture industries and can be applied directly to the characterization of biomass.

Several steps are involved in rapid analysis method development. These include gathering appropriate calibration samples, chemical characterization of the calibration samples, developing spectroscopic methods for the rapid technique, projection-to-latent-structures (PLS) regression, validation of the PLS algorithm, and the development of QA/QC procedures [128].

Rapid analysis methods based on PLS multivariate modeling require calibration based on robust and accurate methods. The first step in developing a new method is to gather appropriate calibration data. Robust methods usually contain at least 100 well-characterized samples. Collecting and characterizing diverse and a good calibration set costs about $300,000. This is by far the most expensive and time-consuming step in method development. Calibration samples should have compositions similar to the samples to be analyzed. If possible, the calibration set should include samples that represent all known sources of compositional variance. The range of compositional variability within the calibration samples determines the validated calibration range for each constituent.

Quality spectroscopy is the second essential component of method development. The technique selected must contain information about the chemical composition of each sample. The spectroscopic method is the key to cost reduction and speed of analysis. In the next step of rapid analysis method development, multivariate analysis is used to identify spectroscopic patterns that correlate with compositional data. In the methods incorporated into work at the NREL, PLS models use whole NIR spectra collected from 400 to 2,500 nm. Equations are obtained that convert spectroscopic data directly into compositional information. In simplified terms, PLS analysis solves hundreds of equations in thousands of variables to obtain a linear equation that predicts compositional information from spectroscopic data. Multivariate analysis is designed for complex systems such as those found in biomass compositional analysis. These powerful mathematical techniques retain the precision and accuracy of the calibration data. It is important to note that PLS analysis can never be more accurate than the methods used to obtain the calibration data [97]. For this reason, the best and most accurate wet chemical techniques should be used for method calibrations, hence the large expense in setting up a calibration set. Once calibration is complete, compositional analysis becomes as fast and inexpensive as the spectroscopic method.

One of the major limitations of rapid analysis methods is that an answer is always provided and the user must determine the validity of the provided data. Robust QA/QC procedures are needed to assure that the rapid methods are appropriately applied to unknown samples [128].

In addition to significant savings in time and money for routine process samples, rapid analysis methods can be used to provide levels of information that were not previously available. For example, feedstock assessment and genetic studies require the screening of hundreds, sometimes thousands of samples. These studies would have been too costly to pursue without the savings in time and cost provided by rapid analysis methods. For example, approximately 200 samples can be analyzed from a bulk corn stover feedstock in 1 day at a cost of less than $4,000. This type of information has been used to develop protocols for the representative sampling of fields, bales, totes, and bags of feedstock. The ability to accurately sample a bulk feedstock and to analyze hundreds of samples for about $20 each provides a new tool that is being used to assess the compositional variability of corn stover in the United States as a function of variety, geographical location, harvest time, and collection method. Changes in feedstock composition during storage are also being monitored. With these larger data sets, feedstock composition can be more accurately reported as a range of expected normal values.

Application of Methods

Rapid, inexpensive biomass analysis can be useful at many stages of an industrial process. Rapid biomass analysis methods can characterize the feedstock as it enters the reactor. If necessary, rapid analysis can be used to guide feedstock blending. Chemical changes during the processing of biomass can provide feed-forward and feed-backward information that can be used to ensure that the process maintains a steady state in spite of the feedstock variability. Finally, process residues and products can be easily evaluated to assess overall process economics. As more samples are analyzed, information can be obtained about the composition of an “ideal feedstock.” Field-mobile instruments can be calibrated for use as purchasing tools. Buyers can obtain compositional information about a biomass feedstock at the point of purchase. Feedstocks can be assigned values based on quality. Young plants and perhaps even seeds can be evaluated (in breeding programs) and selected for desirable characteristics and production potential.

Biomass Process Monitoring and Improvement

The wet chemical and rapid analytical methods described above are being used on a daily basis in the U.S. Department of Energy’s Office of the Biomass Program. The methods are used primarily to keep track of chemical fractions of plant cell walls during biomass processing. Researchers use this capability to better characterize biomass conversion

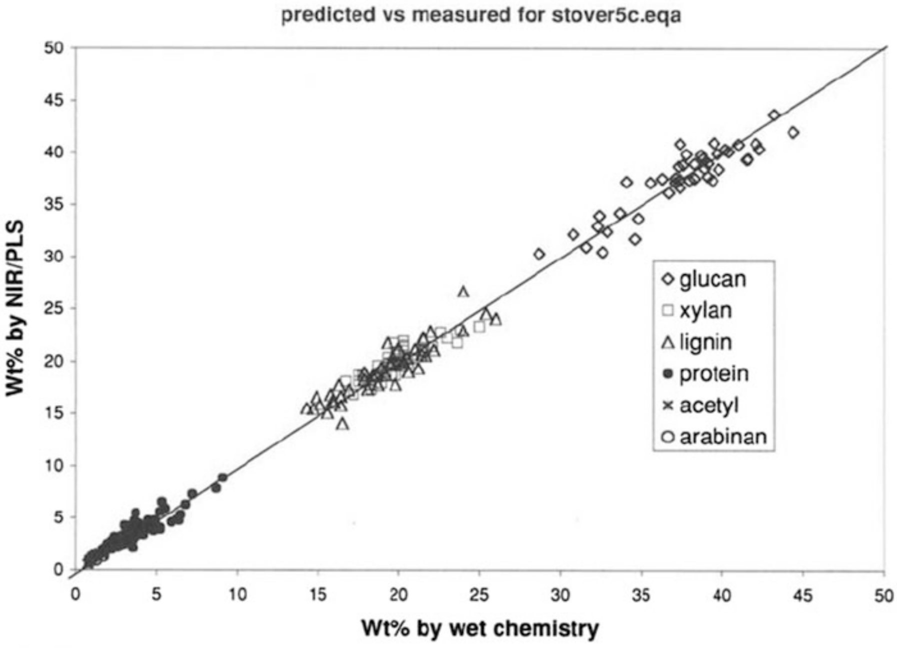


Fig. 33.17 Comparison of corns stover feedstock composition as determined by wet chemical and NIR/PLS method stover 5C

processes and to compare and evaluate alternate approaches to biomass conversion and their impact on overall process economics. Methods have been developed that are specific to and appropriate for determining the composition of raw biomass materials (e.g., corn stover, softwoods, and hardwoods [129]), as well as processing sorghum and other grasses, intermediates (e.g., dilute-acid pretreated hardwoods, softwoods, and corn stover [128]). The high cost of compositional analysis using traditional wet chemical methods was limiting the amount of information available for economic assessments and process optimization. The additional levels of information provided by rapid analysis methods increase the accuracy of technoeconomic models that guide research more efficiently toward the development of commercially viable process options for the conversion of biomass into fuels and valuable chemicals.

Survey of Corn Stover Compositional Diversity

The NREL's corn stover feedstock rapid analysis method has also been used extensively to assess the range of compositional variation present in corn stover from different varieties grown in different locations using a variety of crop management and harvesting strategies. In one study, university agronomy department researchers, agricultural biotechnology and seed companies, and individual farmers all provided stover samples. This was an example of an application that would not and could not have been undertaken using traditional wet chemical methods, because they are both too slow and too expensive. The resource limitations disappeared to a great extent with the development of the rapid analytical methods. This survey has provided indispensable information regarding the range of diversity of corn stover and its impact on biomass conversion process economics.

Using rapid analysis, the NREL has compiled large corn stover composition database. The database includes samples from 53 different locations in ten states (mostly U.S. corn belt) and more than 100 branded hybrids from 22 seed companies. Individual samples typically represent about 10–20 stalks of the same variety collected from the same field and grown during the 2001 season. Stalks were collected by hand after grain harvest was completed, dried to less than 20% moisture at less than 50°C, and shipped to NREL. Stover samples in this survey do not include cobs. Samples were milled using a rotary knife mill (Wiley Mill) to pass a $\frac{1}{4}$-in. screen prior to NIR spectroscopy.

The NIR spectrometer used for method development and sample analysis was a Foss NIR Systems Model 6500 Forage Analyzer with a sample transport module and a standard reflectance detector array. The transport module moves the sample compartment up and down during data collection, thereby allowing a more representative spectrum to be obtained from bulky heterogeneous samples. The reflectance array uses two silicon detectors to monitor visible light from 400 to 850 nm and four lead-sulfide detectors to monitor NIR light from 850 to 2,500 nm. Natural product sample compartment cells in 1/4-cup and 1-cup sizes were used as sample holders in the transport module. This instrument has a maximum resolution of 2 nm.

A key step in the application of rapid analysis methods is the collection of high quality NIR spectra. To minimize the effect of water in the biomass spectra, each sample was air-

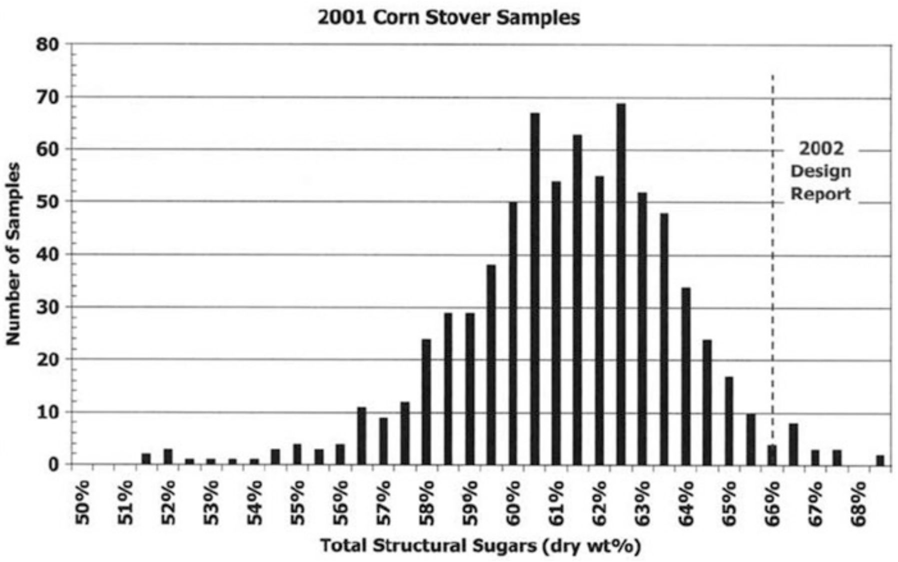


Fig. 33.18 Distribution of total structural sugars content among 738 hybrid corn stover samples collected in 2002

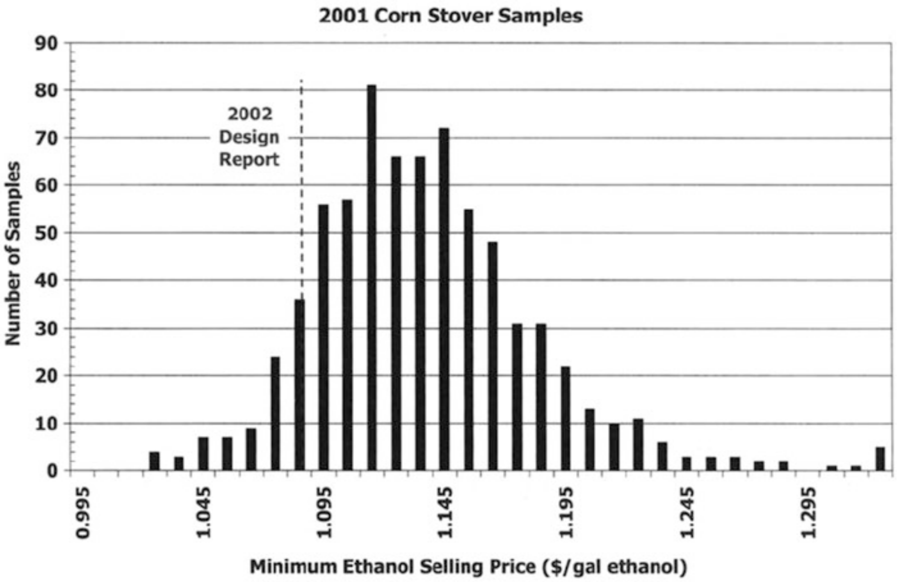


Fig. 33.19 Effect of feedstock composition on minimum ethanol selling price (MESP)

dried to less than 10% moisture prior to NIR spectroscopic analysis. Spectroscopic techniques were used that enable a high quality, reproducible, and representative NIR reflectance spectrum to be obtained. For each sample, a total of 35 spectra were collected and averaged to compensate for sample heterogeneity. Each calibration sample was subsampled three times. Sample cells were emptied and repacked between subsamples. Instrument reproducibility tests demonstrated that the reproducibility limits of the NIR spectrometer contributed less than 0.1% to the absolute prediction errors in the rapid analysis method.

NIR data were converted to compositions using the "stover 5c" rapid analytical method developed at NREL [129]. The ability of the stover 5C methods to accurately measure the composition of corn stover feedstock is shown in Fig. 33.17, where the composition, as determined by NIR/PLS, is compared to measurement of the same samples using standard wet chemical methods.

Improvements in the wet chemical calibration methods and the incorporation of additional calibration samples are continuously improving these rapid biomass analysis methods. In the future, information on these samples of corn stover can be updated using improved methods by reevaluation of their stored electronic spectra.

Figure 33.18 summarizes the results of compositional analysis by showing the distribution of values for total

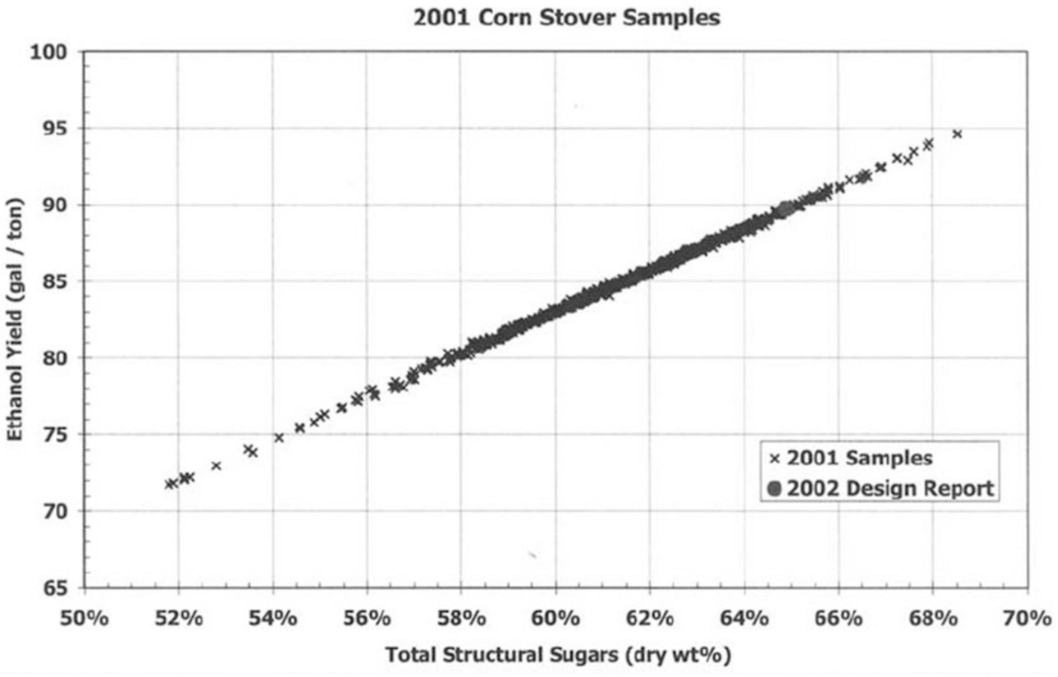


Fig. 33.20 Relationship between carbohydrate content of feedstock and ethanol yield per ton of input material.

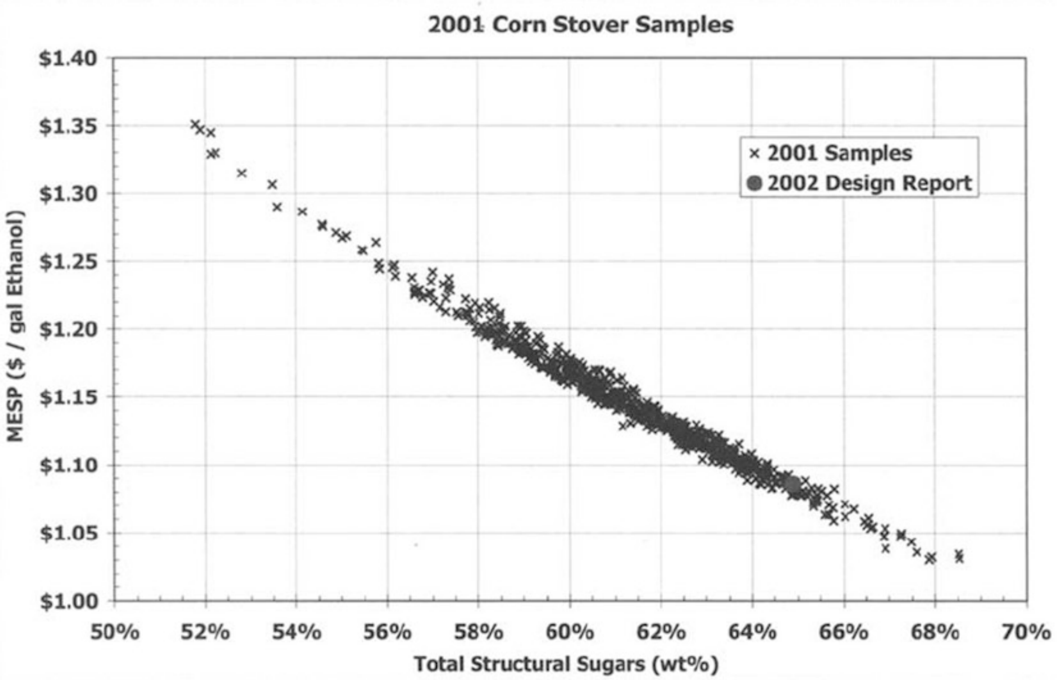


Fig. 33.21 Effect of feedstock composition on minimum ethanol selling price (MESP).

structural sugars (not including uronic acids) as a frequency histogram. Note both the normal distribution and the wide range of values. This same kind of normal distribution with a broad range is echoed by each of the individual major constituents of corn stover (i.e., glucan, xylan, lignin, protein, structural inorganics) across the population of 738 samples (data not shown). The dotted line in Fig. 33.18 indicates the composition of the corn stover used for process modeling efforts in the technoeconomic model by Aden et al. [130]. The fact that the composition assumed for process modeling efforts now appears somewhat optimistic in light of these results has caused subsequent modeling efforts to reconsider the feedstock composition assumptions.

Compositional variability can have a significant impact on biomass conversion process economics. The large effect (i.e., at least $0.30/gal ethanol) of observed compositional diversity on process economics is shown in Fig. 33.19 and is primarily due to the fact that the maximum theoretical product yield is proportional to feedstock carbohydrate content (Fig. 33.20) [131]. Yield is the major economic driver for the technoeconomic model used to assess the economic impact of composition on minimum product selling price [130], as can be seen from the data in Fig. 33.21.

Genetic Screening and Cell-Wall Genomics

The identification and characterization of genes involved in cell wall biogenesis can be approached using either a traditional genetic screening or a reverse genetics strategy. Both approaches are currently being used [116–118].

A forward genetics approach requires the availability of a relatively inexpensive, high-throughput method to assess the phenotype of individuals in a population. In the case of visible phenotypes, this only requires a person to inspect a large number of plants to identify unusual individuals to capture the genetic variants of interest. In the case of chemical phenotypes this is somewhat more difficult, but certainly possible. Chemical screens have been performed previously to find mutations in genes that affect important metabolic pathways and result in accumulation of abnormal branching patterns in starch [132], altered seed protein composition [133], abnormal fatty acids accumulating in storage or membrane lipids [134], or abnormal cell wall composition [135, 136].

Near-infrared spectroscopy has been employed as a high throughput screening tool to collect chemical information from plant tissue in a large genetic screen of a genomewide, transposon mutagenized collection of corn mutants [116, 137]. Recently, direct measurement of genetic, environmental, and other factors effects on biomass recalcitrance have been undertaken by the three DOE funded Bioenergy Research Centers (BRCs). The Bioenergy Science Center has developed a high throughput biomass recalcitrance screening pipeline which screens thousands of biomass samples for differences in pretreatability, enzyme digestibility, and glucan, xylan, and lignin content and composition. As the biomass in this pipeline is treated as near as practical to larger scale processes, the results are directly applicable to current large-scale biomass conversion schemes. In contrast, the Joint Bioenergy Institute uses ionic liquids to completely solubilize the biomass, removing structural and steric considerations from the assay and focusing on understanding the functionality of ionic liquids in large-scale biomass conversion. Similarly, the third BRC, the Great Lakes Bioenergy Research Center, utilizes a high-throughput, small-scale ball mill system to "pretreat" the biomass samples by reducing them to fine powder. This process reduces or eliminates the structural contributions of the cell wall to recalcitrance, but allows screening of the variations in chemistry between large sample numbers.

Future Applications

One of the long-term goals of rapid biomass analysis method development is the generation of robust methods for online and atline process monitoring and control. Because each method is feedstock, process, and sample specific, the achievement of this goal will require a significant long-term effort and considerable financial resources. One step in this direction is the development of automated and robotic methods of wet-chemical analysis that improve sample analysis throughput and decrease the cost of obtaining accurate calibration data.

The development of rapid biomass analysis methods is limited within the DOE program to feedstocks and processes currently used in program demonstration projects and industrial partnerships. Methods for other potential feedstock materials such as wheat straw, switch-grass, sorghum, or alfalfa could be developed if sufficient need and resources are identified.

Summary and Conclusions

Accurate, reliable, precise, and inexpensive methods for determining the composition of biomass at various stages during a conversion process are essential to effectively monitor the progress of the thermochemical and biochemical processes. This information enables the comparison of various process options and enables intelligent decision making in evaluating process options for maximizing product yield while minimizing capital and operating costs. Eventually, with the development of feedback and feedforward loops, a commercial biomass conversion process can be monitored and controlled by a set of automated processes. The availability of more realistic information concerning sources of variability in commercial biomass conversion processes minimizes investment risk and increases the realistic probability of achieving success in the DOE goals in establishing commercial biomass conversion industries in the United States.

Table 33.2 Major cellulase enzyme

IUBMB EC No.	Common name	Activity
Oligomer hydrolase		
EC 3.2.1.21	β-Glucosidase	Hydrolysis of terminal, nonreducing β-D-glucose residues with release of β-D-glucose
Exo-depolymerases		
EC 3.2.1.74	Glucan 1,4-β-glucosidase	Hydrolysis of 1,4-linkages in 1,4-β-D-glucans, to remove successive glucose units
EC 3.2.1.91	Cellulose 1,4-β-cellobiosidase	Hydrolysis of 1,4-β-D-glucosidic linkages in cellulose and celloteraose, releasing cellobiose from the nonreducing ends of the chains
Endo-depolymerase		
EC 3.2.1.4	Cellulose	Endohydrolysis of 1,4-β-D-glucosidic linkages in cellulose, lichenin, and cereal β-D-glucans

Biocatalysts for Biomass Deconstruction

Summary of Plant Cell Wall Structure

Plant cell walls are composed primarily of cellulose, hemicellulose, lignins, and pectins. These give structural rigidity and strength to the plant, deter pathogens, and retain extracellular water. Cellulose, a highly crystalline, insoluble polymer of β-(1,4)-cellobiose, comprises about 50% of the plant biomass. Although cellulose does not degrade easily, it will hydrolyze to glucose by the synergistic action of three distinct classes of enzymes: endoglucanases, exoglucanases, and cellobiases [138, 139]. In contrast with the insoluble linear cellulose homopolymer, hemicelluloses are water- or base-soluble heteropolymers, comprised of a variety of branched and substituted polysaccharides. In addition to providing water retention and structural reinforcement, hemicelluloses act as cross-linking agents.

It is worth noting that although hemicelluloses are considered soluble in their native form in vivo, extraction can result in significant debranching, often rendering them insoluble in water. Recently, Gatenholm and Tenkanen assembled a collection on properties, uses, and examinations of various hemicelluloses, both in industrial and basic science applications [140]. The complex structure of hemicelluloses has dictated an accordingly diverse array of hemicellulases. Generally, each structural feature in hemicellulose has an associated enzyme that can hydrolyze or modify this feature. Pectins are acidic polysaccharides that retain large amounts of water and act as an adhesive between adjacent plant cells, and, along with lignin, comprise much of the middle lamella. Although pectin may be found to some extent in most plant cell walls, it is a minor component in most tissues (fruits being the most notable exception), and its degradation is not covered here. For reviews on biodegradation by and commercial uses of pectinases, see articles by Galante and Formantici, Hoondal et al., Kashyap et al., Lebeda et al., and Naidu and Panda [141–145]. Lignin, by contrast, is a heterogeneous polymer of phenyl propanoid units containing various phenolic derivatives. Lignin is often thought of as the binder that cements the cell wall components together. Enzymatic degradation of lignin is still somewhat controversial, with various hydrolytic and oxidative mechanisms proposed. Many papers have been published on white rot degradation of lignin in recent years, however, the authors are unaware of a unified complete mechanism of lignin degradation, and an overview is not attempted here. Several reviews have, however, covered selected areas including lignin degradation by mycorrhizal fungi, dye decolorization by white rot fungi, lignin biodegradation in compost, applications in pulp and paper and ruminant feed, and the emerging molecular genetics of ligninolytic fungi [146–152].

Enzymatic Hydrolysis of Plant Cell Wall Components

The complex nature and interconnectivity of plant cell wall polymers preclude straightforward enzymatic digestion. There are dozens of enzyme families involved in plant cell wall hydrolysis, including cellulases, hemicellulases, pectinases, and lignin-modifying enzymes. The Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB) has classified cellulases and hemicellulases, like all enzymes, into different classes based on activity. Tables 33.2 and 33.3, compiled from the IUBMB enzyme nomenclature database (http://www.chem.qmul.ac.uk/iubmb/enzyme/), list the IUBMB enzyme classifications for cellulases and hemicellulases [153].

Synergy is a major factor in degradation efficiency, making measurement of these activities difficult, to say the least. The biomass degrading enzyme system, encompassing cellulases, hemicellulases, pectinases, lignin-modifying enzymes, and other accessory activities, is far too complex to be covered in detail in this general overview. The American Chemical Society (ACS) has covered much of this work in the ACS Symposium series including applications to lignocellulosics, fiber processing, pulp and paper, biomass

Table 33.3 Enzyme classes involved in hydrolysis of hemicelluloses

IUBMB EC No.	Common name	Activity
Debranching enzymes		
EC 3.1.1.6	Acetylesterase	An acetic ester + H_2O = an alcohol + acetate
EC 3.1.1.72	Acetylxylan esterase	Deacetylation of xylans and xylo-oligosaccharides
EC 3.1.1.73	Feruloyl esterase	feruloyl-polysaccharide + H_2O = ferulate + polysaccharide
EC 3.2.1.131	Xylan α-1,2-glucuronosidase	Hydrolysis of α-D-1,2-(4-*O*-methyl)glucuronosyl links in the main chain of hardwood xylans
EC 3.2.1.139	α-Glucuronidase	An α-D-glucuronoside + H_2O = an alcohol + D-glucuronate
EC 3.2.1.55	α-*N*-arabinofuranosidase	Hydrolysis of terminal non-reducing α-L-arabinofuranoside residues in α-L-arabinosides
Oligomer hydrolases		
EC 3.2.1.22	α-Galactosidase	Hydrolysis of terminal, nonreducing α-D-galactose residues in α-D-galactosides, including galactose oligosaccharides, galactomannans and galactohydrolase
EC 3.2.1.23	β-Galactosidase	Hydrolysis of terminal nonreducing β-D-galactose residues in β-D-galactosides
EC 3.2.1.25	β-Mannosidase	Hydrolysis of terminal, nonreducing β-D-mannose residues in β-D-mannosides
Exo-depolymerases		
EC 3.2.1.100	Mannan 1,4-mannobiosidase	Hydrolysis of 1,4-β-D-mannoside linkages in 1,4-β-D-mannans, to remove successive mannobiose residues from non-reducing ends
EC 3.2.1.145	Galactan 1,3-β-galactosidase	Hydrolysis of terminal, non-reducing β-D-galactose residues in (1,3)-β-D-galactopyranans
EC 3.2.1.37	Xylan 1,4-β-xylosidase	Hydrolysis of 1,4-β-D-xylans, to remove successive D-xylose residues from the non-reducing termini
EC 3.2.1.58	Glucan 1,3-β-glucosidase	Successive hydrolysis of β-D-glucose units from the nonreducing ends of 1,3-β-D-glucans, releasing α-glucose
EC 3.2.1.72	Xylan 1,3-β-xylosidase	Hydrolysis of successive xylose residues from the nonreducing termini of 1,3-β-D-xylans
Endo-depolymerases		
EC 3.2.1.101	Mannan endo-1,6-α-mannosidase	Random hydrolysis of 1,6-α-D-mannosidic linkages in unbranched 1,6-mannans
EC 3.2.1.136	Glucuronoarabinoxylan endo-1,4-β-xylanase	Endohydrolysis of 1,4-β-D-xylosyl links in some glucuronoarbinoxylans
EC 3.2.1.151	Xyloglucan-specific endo-β-1,4-glucanase	xyloglucan + H_2O = xyloglucan oligosaccharides (endohydrolysis of 1,4-β-D-glucosidic linkages in xyloglucan)
EC 3.2.1.32	Xylan endo-1,3-β-xylosidase	Random hydrolysis of 1,3-β-D-glycosidic linkages in 1,3-β-D-xylans
EC 3.2.1.39	Glucan endo-1,3-β-D-glucosidase	Hydrolysis of 1,3-β-D-glucosidic linkages in 1,3-β-D-glucans
EC 3.2.1.6	Endo-1,3(4)-β-glucanase	Endohydrolysis of 1,3- or 1,4-linkages in β-D-glucans when the glucose residue whose reducing group is involved in the linkage to be hydrolyzed is itself substituted at C-3
EC 3.2.1.71	Glucan endo-1,2-β-glucosidase	Random hydrolysis of 1,2-glucosidic linkages in 1,2-β-D-glucans
EC 3.2.1.73	Lincheninase	Hydrolysis of 1,4-β-D-glucosidic linkages in β-D-glucans containing 1,3- and 1,4-bonds
EC 3.2.1.78	Mannan endo-1,4-β-mannosidase	Random hydrolysis of 1,4-β-D-manosidic linkages in mannas, galactomannans and glucomannans
EC 3.2.1.8	Endo-1,4-β-xylanase	Endohydrolysis of 1,4-β-D-xylosidic linkages in xylans
EC 3.2.1.89	Arabinogalactan endo-1,4-β-galactosidase	Endohydrolysis of 1,4-β-D-galactosidic linkages in arabinogalactans
EC 3.2.1.99	Arabinan endo-1,5-α-L-arabinosidase	Endohydrolysis of 1,5-α-arabinofuranosidic linkages in 1,5-arabinans

for fuels, biomass conversion, and general activities on insoluble carbohydrates [154–160]. Readers are encouraged to find details for specific applications in these proceedings. The burgeoning number of scientific informational databases accessible through the Internet has also helped greatly to clarify (to some extent) the confusing nature of these enzymes. Sites include general enzyme databases, such as ExPASy (http://us.expasy.org/, the server gateway to Swiss-Prot/TrEMBLE (http://us.expasy.org/sprot/) and ENZYME (http://us.expasy.org/enzyme/]) and BRENDA (http://www.brenda.uni-koeln.de), as well as databases devoted exclusively to carbohydrases, such as CAZy (http://afmb.cnrs-mrs.fr/CAZY/acc.html) [161–164].

Current literature describing the assaying of general cellulase activity (or of individual component enzymes) has broadened considerably since Mandels et al. [165] reported reducing sugar release and substrate weight loss as suitable cellulase assay methods. To some extent, and for appropriate substrates, these methods are still considered adequate. However, modern assays based on molecular weight (MW) analysis detected by high-performance liquid chromatography size-exclusion chromatography (HPLC-SEC), coupled enzymes, viscometry, hydrolysis of dyed or derivatized insoluble and soluble polymers, and hydrolysis of derivatized or labeled low-MW substrates, have greatly enhanced the understanding of these complex systems. Cellulose structure and physical disruption of cellulose microfibrils have also been examined using light, electron, and more recently, atomic force microscopy [166–173]. These types of "assay" are more useful when the properties of the cellulose fiber are in question, not when the goal is total hydrolysis. There has also been work using gel-permeation chromatography to characterize changes to cellulose structure by examining the products of cellulase action on wood fiber [174]. As with much of biotechnology today, high-throughput methods have also been developed to increase the speed and accuracy of cellulase assay [175]. Despite all of the assays, or perhaps because of them, workers in the field are reminded that only assays designed to measure the conversion of cellulose from the actual biomass substrates in question are ultimately valid performance measures.

Most enzymes have very specific requirements for tight substrate binding and precise transition-state formation, which usually leads to high catalytic turnover rates. However, even ideal catalytic sites must be "carried to the substrate" by the macromolecule within which it is housed. Enzymes are also large compared to the polysaccharide oligomers under attack, especially as the particular site of action may be buried in a heterocrystalline structure of mixed polysaccharides. The complexity of biomass, specifically in regard to hemicellulose structure, requires a high degree of coordination between degradative enzymes. Studies show correlations between the enzymatic digestibility of cellulose and the removal of hemicellulosic sugars and lignin, supporting the notion of close spatial relationships [35, 176]. Of further complication is that the actions of glycosyl hydrolases often change the chemical environment of the partially degraded substrate, which in turn affects the actions of other glycosyl hydrolases. For example, partly because of the substituents attached to the main chain, most hemicelluloses are quite water soluble in their native state. These side chains disrupt the water structure and help to solubilize the hemicellulose. Debranching enzymes that remove these substituents generally decrease substrate solubility, and in turn lower the polysaccharide's susceptibility to endo-acting hydrolases [177]. Thus, a xylan that has been subjected to acetyl xylan esterase is less susceptible to enzymatic degradation than a xylan subjected to a mixture of branching and debranching enzymes [178]. As the substituents are removed, xylan can become less soluble, forming aggregates that sterically hinder and finally block further degradation [179]. The endoxylanases, for example, cleave the main chain linkages and are often quite particular about the type of linkage, type of sugar, and presence or absence of nearby substituents [180]. An endoxylanase that cleaves β-(1,4) linkages will usually have no effect on β-(1,3) linkages. Also, an endoxylanase that cleaves main-chain linkages near an *O*-2 linked arabinose may have no effect on an open-chain xylan [181].

Cellulases

Ask an industrial biochemist about cellulases and *Trichoderma reesei* will invariably come up. A mesophilic filamentous fungus typically found growing on cellulosic biomass, *T. reesei* produces arguably the most studied cellulase system in the world. Early research began when fungus-laden U.S. Army accoutrements were sent from the jungles of south-east Asia to the U.S. Army Natick Research and Development Laboratories during the 1940s [182]. This research program, led by Drs. Elwyn Reese and Mary Mandels for 40 years, succeeded not only in providing the U.S. Army and allies with chemical agents useful for protecting cotton goods, but also demonstrated the biological nature of their decay. The work at Natick demonstrated that many genera and species of fungi produce cellulose-degrading proteins, and that *T. reesei* was especially effective at degrading cotton. The Natick labs went on to learn how to grow *Trichoderma*, induce cellulase production, and determine operational ranges of pH and temperature, as well as tolerance to cations, anions, metals, solvents, and biomass-derived toxic compounds [183]. Furthermore, internship opportunities at Natick during the 1950s and early 1960s served to train the leaders of the next generation of cellulase scientists. In the end, this was perhaps one of the most important contributions of the program.

Despite cellulose's simple linear glucose polymer structure, its hydrolysis requires the coordinated action of at least three different types of β-1,4-glucanases (Table 33.2).

Endoglucanases

Endo-β-1,4-glucanases (EC 3.2.1.4) cleave the cellulose strand randomly along its length, generating new chain-end sugars, one reducing and one nonreducing. Usually, these chains remain insoluble and are thought to remain embedded in the cellulose crystal. The new ends, however, when exposed to the bulk water/cellulose interface, become new recognition sites for exocellulases.

Exoglucanases

Exo-β-1,4-glucanases (EC 3.2.1.91) cleave cellobiosyl (β-1,4-glucose dimers) or glucosyl units endwise from cellulose, chains. Most fungal, and some bacterial cellulase, systems display two types of exoglucanases: reducing-end and nonreducing-end cleavers. Exocellulases are thought to act processively on cellulose chains, starting at one end and

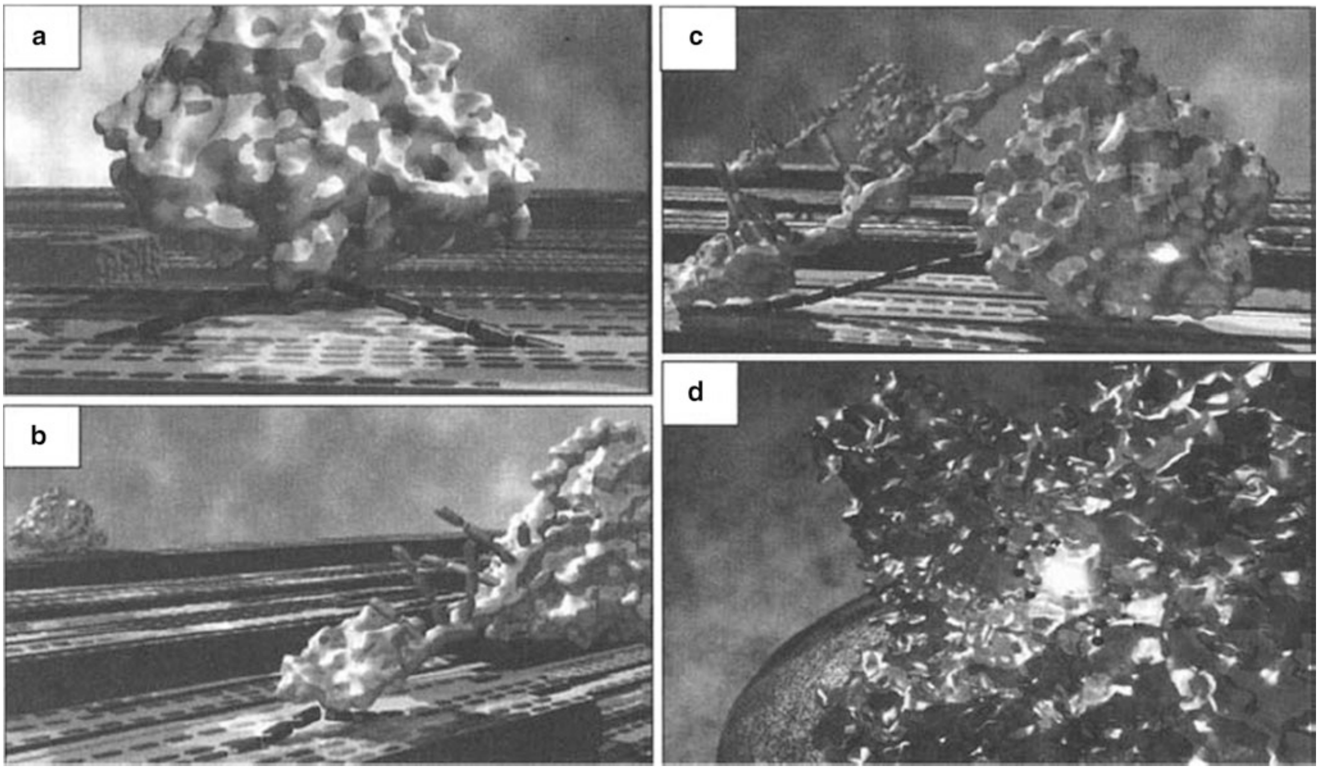


Fig. 33.22 Three enzyme activities in the degradation of cellulose: (**a**) endocellulase cel5A from *Acidothermus cellulolyticus*; (**b**) exocellulose cel7A from *Trichoderma reesei* finding an open reducing end; (**c**) cel7A processing along cellulose chain releasing cellobiose; and (**d**) bglA from *Bacillus polymyxa* hydrolyzing cellobiose

proceeding down the chain. The products are either cellobiose or glucose, and along with the reducing/nonreducing end-specificity, are specific to the enzyme in question.

Endoglucanases and β-Glucosidases

When endo-activity occurs close to a chain end in an amorphous region of the cellulose, the new cellulose oligomer may become dissociated from the larger structure. The degree of polymerization of the chain determines its solubility, ranging from highly soluble glucose dimer cellobiose to nearly insoluble cellohexaose. New soluble oligomers are usually acted upon by a third class of cellulases, the β-glucosidases (EC 3.2.1.21). This class acts primarily, if not exclusively, on soluble cello-oligomers, with decreasing specific activity as chain length increases. As their preferred substrate is cellobiose, they have also been termed cellobiases.

All three of these general activities are required for efficient and complete hydrolysis of cellulose to glucose. A conceptualized view of each typical cellulase action is shown in Fig. 33.22 [184]. The process is initiated by an endocellulase (in this case cel5A from *Acidothermus cellulolyticus*), which nicks the cellulose chain (Fig. 33.22a), exposing a new reducing and nonreducing chain end. Exocellulases (in this case the reducing-end specific cel7A from *T. reesei*) find these nicks (Fig. 33.22b) and feed the ends through their active sites, cleaving off products in a processive manner (Fig. 33.22c). If cellobiose is the product, as in the case of cel7A, then β-glucosidase hydrolyzes it to two glucose molecules (Fig. 33.22d).

Total Cellulase Activity Measurements

Each cellulase class has had its share of assay development. Assays that specifically measure each class are difficult to establish, as cross-reactivity between classes is greater than zero. However, there are assays that can be used to determine the relative ratios of these activities.

When using cellulases to degrade lignocellulosic material, total activity is usually the primary interest. The complexity of lignocellulose substrates and the huge heterogeneity in cellulase systems, however, create a quagmire for the cellulase enzymologist asked to determine the "activity" units of a particular cellulase. There are numerous methods for assaying cellulase activity, but most fall into one of three

Table 33.4 Specific activities of various *Trichoderma reesei* cellulase preparations

Source	Mix of component enzymes	Specific activity[a] (FPU/g protein)	References
Multifect GC[b]	Native	431	[675]
Ultra-low microbial cellulase 100[c]	Unknown	478	[675]
Mutant strain RUT-C30	Probably nonnative	495	This study[d]
Spezyme lot #41-59034-004CL121	Native	565	[675]
Cellulase TRL[e]	Native	569	[675]
Spezyme lot #GC310 S1.2E2Z1P1	Native	571	[675]
Mutant strain L27	Non-wild-type [676]	581	This study[d]
Mix CBH I:CBH II:EG 1/60:20:10	Native mix [677] reconstituted	711	This study[f]
Mix CBH I:CBH II:EG 1/40:50:10	Mix giving max activity	761	This study[f]

[a] Using the method of Ghose [186] (in order to facilitate comparison, all assays reported here were performed in a consistent manner)
[b] Spezyme and Multifect from Genencor International, South San Francisco, CA
[c] Iogen Corporation, Ottawa, ON
[d] Grown on 1% Sigmacel 50 + 2% lactose in Mandels media [678]
[e] Cellulase TRL form Solvay Enzymes, Elkhart, IN
[f] Reconstituted mixtures from purified cellulase enzymes

categories: sugar release (measured as reducing sugar), chain length reduction (measured by viscometry, light scattering, or size exclusion chromatography), and artificial substrate hydrolysis (measured by dye, chromophore, or fluorophore release).

The IUPAC Filter Paper Assay

In 1987, after significant effort, an international committee of cellulase researchers and the International Union of Pure and Applied Chemists (IUPAC) produced a procedure seeking to standardize cellulase activity measurements. This procedure uses microcrystalline cellulose and the dinitrosalicylic (DNS) acid method of Miller [185] to measure reducing sugars in the context of a highly specific assay protocol [186]. Because cellulose hydrolysis is nonlinear, the units were not defined as specific activity units per se (i.e., μmoles product/mg protein seconds), but rather as 4% hydrolysis of a 50-mg coupon of Whatman #1 filter paper, although many cellulase researchers have noted that after accounting for the addition of water to the glycosidic bond, the actual level of hydrolysis measured is 3.6% of the substrate. Regardless, the "filter paper unit" (FPU) was decreed the amount of enzyme that released 2.0 mg glucose from a 1.0 × 6.0 cm, 50-mg Whatman #1 test coupon after 60 min at 50°C. The dilution of enzyme required is converted, through a somewhat indirect procedure, to FPU per mL. For example, an undiluted cellulase preparation that yields exactly 2.0 mg glucose has 0.37 FPU/mL. This fractional unit is the lowest cellulase activity measurable with the IUPAC assay.

The IUPAC cellulase assay has significant limitations and merely serves as the best-defined method. Extrapolation of required glucose release from highly diluted or concentrated enzyme solutions is not permitted; assays must be conducted with enzyme dilutions that closely bracket the actual value. Cellulase solutions too dilute to release 2.0 mg glucose must either be concentrated to an appropriate level, or determined incapable of assay by the IUPAC method. For highest accuracy, every working solution made from an enzyme stock must be reanalyzed for activity, a condition that complicates most analytical procedures.

Cellulase activities of *T. reesei* broths are normally reported to lie between 400 and 600 FPU/g total protein [187]. Our work to assess the specific activities of *T. reesei* cellulase preparations has led us to the direct comparison of commercial cellulase products, typically highly selected *T. reesei* mutants, and reconstituted, purified cellulase enzymes (Table 33.4). Although the range of specific activities found from this internally consistent study generally agrees with the literature, our estimation of the maximal specific activity of reconstituted *T. reesei* cellulases is somewhat higher, 760 FPU/g protein. As some of the protein in commercial cellulases are comprised on non-active protein, this result is in line with an assay based on activity per gram of protein.

In fact, the activity of the native mixture of *T. reesei* cellulases (i.e., 10:20:60 for endoglucanase:cellobiohydrolase II:cellobiohydrolase I [EG I:CBH II:CBH I]) using reconstituted enzymes is about 10% less than that obtained with a mix higher in CBH II (Table 33.4). One should be aware that it is possible to find reports in the literature of *T. reesei* cellulases with specific activities greater than about 600 FPU/g protein. Considering these findings in view of recent studies of the magnitudes of possible inter- and intralaboratory errors in filter paper

[188] and protein assays [189], assay errors or inconsistencies are most likely responsible for most discrepancies in specific activity reported [190].

Variations on the filter paper assay have been developed. Irwin et al. have established a method for differentially measuring soluble and insoluble reducing sugars released from filter paper [191]. Ghose and the IUPAC committee also proposed a similar approach for measuring hemicellulases in 1987 [192]. This method relies on meeting a standard level of conversion of the xylan fraction in oat spelt xylan to xylose in a specified period of time under standard conditions. This latter method; however, is not often cited in the literature.

General Non-IUPAC Cellulases Assays

Many cellulase enzyme preparations are simply not concentrated enough to cause the required release of 2 mg glucose from the 50-mg filter paper sample in 60 min. If these samples cannot be concentrated accurately (which is often the case), traditional FPU cannot be measured. In such cases, however, the IUPAC committee recommends that the reducing sugar release per unit time be accepted as a "provisional" measure of enzyme activity. This is similar to the pseudo-initial rate approach often used in the decade previous to the IUPAC report to measure cellulase activity from a wide variety of substrates. These substrates may include filter paper [193], Avicel [194], de-waxed cotton [195], or phosphoric-acid-swollen cellulose (PSC) [196]. Methods based on the use of antibiotic disks [197] and turbidity development [198] also predated the IUPAC study. Because all three cellulase types generate new reducing ends, the most direct measurement of cellulase activity is the detection of new reducing ends in the cellulose chain. These methods are traditionally based on initial rate measurements introduced by Wood and McCrae [194].

The reducing sugars are typically measured by the Somogyi [199] and Nelson [200] procedures, which measure reduction of Cu^{2+} to Cu^{+} in alkaline solution, or by the dinitrosalisylic acid (DNS) assay, in which reducing sugars reduce DNS to 3-amino, 5-nitrosalicylic acid under alkaline conditions. Disodium 2,2′-bicinchoninate (BCA) can also be used to measure reducing sugars [201]. This reagent was found to be the best choice in a recent comparison of methods for the determination of endoglucanase activity [202]. The BCA method was highly sensitive, simple to perform, and directly gave the number of bonds broken, thus allowing for expression of endoglucanase activity in international units (μmol of β-1,4-glucosidic bonds hydrolyzed in 1 min during the initial period of hydrolysis).

Other methods used to measure cellulase activity include an automated filter paper assay for high throughput evaluation of cellulases [175], a diafiltration saccharification assay for minimization of product inhibition and high degree of digestion [203], measurement of cellulose size reduction by multiangle laser light scattering [202], use of dye release from various dyed substrates, and differential staining of cellulose digestion products [204–207].

Mathematical Modeling

A mathematical model has also been proposed for evaluating cellulase preparations. Sattler et al. [208] describe a relationship between hydrolysis extent, reaction time, and enzyme concentration. This procedure permits the effectiveness of different enzymes and of different pretreatment methods to be ranked. This method examines cellulose hydrolysis data collected from hyperbolic functions of substrate concentration vs. cellulase enzyme concentration at various timed incubations. The model is based on a double reciprocal plot of the relationship

$$(Y/C_o)^{-1} = (KC_o/Y_{\max})[E]^{-1} + (Y_{\max}/C_o)^{-1}$$

Y/C_o is the fraction of substrate hydrolyzed, $[E]$ is given in FPU/g substrate initially added, and $Y_{\max}/C_o$ is the fraction of substrate that could be maximally hydrolyzed at an infinite enzyme concentration. The y-axis intercept in the double reciprocal plot, $(Y_{\max}/C_o)^{-1}$, may be used to quantify the quality of the enzyme preparation. Ideally, an enzyme should have a high $Y_{\max}$ and a low value for $KC_o/Y_{\max}$. Adney et al. used this general method successfully to model the action of commercial *T. reesei* cellulase preparations on Sigmacell-50 [209]. Results from double reciprocal plots of enzyme activity, (percent conversion)$^{-1}$, vs. loading, (FPU/g cellulose)$^{-1}$, enabled extrapolation to infinite enzyme loading or maximal digestibility.

Endocellulase Activity Measurement

The "endo-1,4-β-glucanases" or 1,4-β-D-glucan 4-glucanohydrolases (EC 3.2.1.4), which act randomly on soluble and insoluble 1,4-β-glucan substrates, are commonly measured by detecting the decrease in viscosity or increase in reducing groups using carboxymethylcellulose (CMC) [202, 210, 211]. It is thought that the methyl groups limit activity of the exoglucanases while still permitting the endo-types to hydrolyze the chain interior, perhaps in regions with low side-chain density. Endoglucanases are also capable of hydrolyzing various polymeric, substituted substrates, such as Ostazin brilliant-red hydroxyethylcellulose (OBR-HEC) and azo-dyed and cross-linked HEC (AZCL-HEC), as well as the low-MW fluorogenic substrates, such as 4-methylumbelliferyl-cellobiose (MUC), 4-methylumbelliferyl-lactopyranoside (MUL), or 4-methylumbelliferyl-cellotriose (MU-G3). Although these substrates make it easy to measure activity, care should be taken in relating the measured activities to performance on biomass, as the two substrates are not equivalent.

Viscometric Assays

Viscometric approaches to cellulase measurement activities are important because other methods measure only the number of glycosidic bonds cleaved in a polymeric substrate. They do not provide any information about location in the substrate of the bonds cleaved. Viscometric methods measure a substantial change in a physical property of the substrate polymer, a very sensitive function of both the number and the location of the bonds cleaved [212–214]. For this reason, even though the recommended international units of carboxymethylcellulase are given in terms of glycosidic bonds cleaved, the most careful workers measure both bond cleavage (most often by measurement of sugar-reducing groups) and the change in solution viscosity as enzymatic hydrolysis proceeds. Vlasenko et al. [202] found the viscometric method to be simple to perform and highly sensitive for the internal bonds cleaved. However, this method does not account for the hydrolysis of CMC near the chain end and thus only allows for expression of endoglucanase activity in arbitrary viscometric units.

Activity Measurements

The "exo-1,4-β-D-glucanases" include both the 1,4-β-D-glucan glucohydrolases (EC 3.2.1.74), which liberate D-glucose from 1,4-β-D-glucans and hydrolyze D-cellobiose slowly, and 1,4-β-D-glucan cellobiohydrolase (EC 3.2.1.91), which liberates D-cellobiose from 1,4-β-glucans. Differentiation of these enzyme classes requires analytical techniques to distinguish glucose and cellobiose and is usually carried out by high-performance liquid chromatography (HPLC) or gas chromatography (GC). These enzymes can be further distinguished by their ability to liberate free sugars from either the reducing or nonreducing end of the cellulose chain [215, 216]. Synergy studies with enzymes of known orientation are usually used to determine chain-end preference given enzyme [191, 217, 218].

The process of detecting and verifying exoglucanases has long been controversial. If purified proteins are available, careful comparisons of reducing sugar yields and fluidity values from CMC hydrolysis as a function of enzyme concentration can be used to judge whether an enzyme is more EG-like or CBH-like. Of course, purified enzymes can also be further verified by analyzing the product from hydrolysis of a series of derivatized (i.e., radiolabeled, chromophoric, or fluorophoric) cello-oligomers [219, 220]. One class of these derivatives, cellobiosyl fluorides, has been reported to distinguish between CBH I and CBH II from *T. reesei* based on cleavage activity on the alpha and beta conformations of the cellobiosyl fluorides [221]. Claeyssens et al. also reported this type of rigorous analysis for fungal CBH I and CBH II [222]. Further specificities can be determined for proposed exoglucanases from analytical product evaluation by HPLC [223, 224]. This is a much more definitive method of distinguishing endo-from exo-acting cellulases.

In general, exoglucanases such as CBH I can be expected to hydrolyze the aryl-substrates MUC and MUL at the agluconic bond, but not the substituted soluble celluloses such as AZCL-HEC, OBR-HEC, and CMC [225]. Phosphoric acid swollen cellulose is also used as a substrate for exoglucanases; however, some endoglucanases will hydrolyze this cellulose form as well [224]. Analysis of activity on higher oligomeric derivatives has proven to be challenging, especially with EGs and CBH I [225]. Furthermore, because endoglucanases are highly synergistic with exo-acting glucanases, the presence of endoglucanases significantly complicates efforts to quantify exoglucanase activity and can be compensated for only by the separate purification and kinetic characterization of the endoglucanase. Attempts persist to link this synergy effect to the cellulose-binding module (CBM). There is some evidence that the CBM alone can have a synergistic effect on the activity of fungal cellulases, both exo- and endo-acting types [226–228].

β-Glucosidase Activity Measurements

The "β-D-glucosidases" or β-D-glucoside glucohydrolases (EC 3.2.1.21) act to release D-glucose units from cellobiose, soluble cellodextrins, and an array of glycosides. For cellobiose or cello-oligomers, this activity is usually measured and analyzed by HPLC or GC, or by direct spectrophotometric or fluorometric analysis of various chromogenic and fluorogenic analogues of cellobiose and cello-oligomers. β-D-glucosidase/cellobiase activities are usually determined according to the method of Wood [229] as aryl-β-glucosidase activity by the hydrolysis of *p*-nitrophenyl-β-D-glucopyranoside. The concentration of *p*-nitrophenol is determined from the absorbance at 410 nm under alkaline conditions induced by the addition of 1-2 volumes M Na_2CO_3. One unit of activity is defined as the amount of enzyme that catalyzes the cleavage of 1.0 μmol substrate per minute at 99°F (37°C). The temperature selection here seems somewhat arbitrary, as most cellulases have optimal activity at somewhat higher temperatures. If necessary, β-D-glucosidases can be distinguished from other cellulases by the relative differences in the initial rates for aryl-β-D-glucosides and cellobiose. Also, the unique and acute sensitivity of β-D-glucosidase to inhibition by gluconolactone provides a method to assess exoglucanase activity in mixed systems of these two enzymes. This approach is necessary because β-D-glucosidase cleaves the agluconic, as well as the holosidic bond of aryl-glucosides [230]. Note that a similar approach is often used to assay other aryl-glycosidases [231–233]. This practice is made possible by the availability of many *o*- and *p*-linked aryl-glycosides including (but not limited to) β-xylosides, β-mannosides, β-galactosides, and L-arabinofuranosides.

Hemicellulases

In contrast to cellulases, the hemicellulases encompass a much broader suite of activities. In addition to analogues versions of endo-, exo-, and glycosidase cellulase activities, multiple debranching activities are needed to handle the high complexity of the heterogeneous hemicelluloses (Table 33.3). The varied backbone composition of hemicelluloses also adds complexity. Xylans, xyloglucans, man-nans, and numerous other minor polysaccharide chains form the backbone for different hemicelluloses.

1. The endo-enzymes attack polysaccharide chains internally, with decreasing activity as chain length rises. Endoxylanases (EC 3.2.1.8) are specific for β-(1 → 4)-xylopyranose polymers, (i.e., the backbone of xylan) and others are specific for other hemicellulose polymers, such as mannan (endo-(1 → 4)-β-mannosidases, EC 3.2.1.78) or β-glucanases (endo-(1 → 3)-β-D-glucosidase, EC 3.2.1.39). As with endocellulases, these activities can be measured by viscometry or production of reducing sugar end groups on the appropriate hemicellulosic polymer.
2. The exo-enzymes act processively from either the reducing or nonreducing termini and again are specific to the type and length of the polymer. Some exo-acting enzymes have preferences for short chain substrates (DP 2–4), acting as cellobiase analogues, whereas others prefer larger substrates (DP > 4). Xylan (1 → 4)-β-xylosidase (EC 3.2.1.37), glucan (1 → 3)-β-glucosidase (EC 3.2.1.58), and mannan (1 → 4)-mannobiosidase (EC 3.2.1.100) are exo-acting enzymes specific for xylan, β-(1 → 3)-glucan, and mannan, respectively.
3. So-called "accessory" enzymes are also required for hydrolysis of hemicellulose in native plant tissue. These enzymes include a variety of acetyl xylan esterases (EC 3.1.1.72), acetyl esterases (EC 3.1.1.6), and esterases, such as feruloyl acid esterase (EC 3.1.1.73) [177], that hydrolyze lignin glycoside bonds. They also include classes for cleavage of specific hemicellulose side chains, such as α-L-arabinofuranose, glucuronic acid, and 4-*O*-methyl-glucuronic acid groups.

General Hemicellulase Assays

Xylan is the most predominant hemicellulose in most plant cell walls, comprising about $\frac{1}{4}$ to $\frac{1}{3}$ of total plant biomass [234]. Xylans function primarily by forming crosslinks between the other cell wall components, such as cellulose, lignin, other hemicelluloses, and pectin. The term "xylan" is a catch-all for polysaccharides that have β-(1 → 4)-D-xylopyranose backbones with a variety of side chains usually attached at the *O*-2 and *O*-3 positions, and include glucuronic acid, 4-*O*-methylglucuronic acid, L-arabinofuranose, xylose, and acetyl groups. The types and levels of side chains are dependent on the particular plant, with hardwoods having high acetyl and glucuronic acid moieties (glucuronoxylans) and grasses having mainly arabinofuranose and acetyl groups (arabinoxylan). Because of the complex nature of xylans, their enzymatic hydrolysis is intrinsically more complicated than that of most other plant polysaccharides. Typically, the polymer is debranched, either prior to, or in conjunction with, depolymerization of the backbone. Different debranching enzymes are required depending on the specific type of xylan being hydrolyzed. These include arabinofuranosidases, ferulic and coumaric acid esterases, acetyl and acetyl xylan esterases, glucuronidases, and xylosidases. Removal of these side chains synergistically enhances the rate of degradation by endoxylanase enzymes [177, 235–241].

Hemicellulose Debranching Enzymes

Debranching of the xylan backbone produces a wide variety of soluble low-molecular-weight compounds. Typically, these products are measured either by HPLC or GC. The difficulty in assaying these products is not so much in the detection as in obtaining the correct substrate for the enzyme. Most commercial xylan products are extracted by alkaline treatment, essentially hydrolyzing any ester linkages by saponification; that is, any acetyl-, coumaroyl-, or feruloyl esters are destroyed. Glycosidic side chains, such as arabinose or glucuronic acid are left intact, however, the polymer is typically insoluble. Enzyme studies using these substrates must be interpreted with caution, as the native esterified xylan is soluble. Extraction by DMSO or steam has been used to prepare "native" xylan, in which the esters are still intact and the polymer is soluble in water [242].

Arabinofuranosidases. α-L-arabinofuranosidases (EC 3.2.1.55) cleave α-L-arabinofuranosides from the arabinoxylan xylose backbone and have been shown to enhance the release of ferulic and coumaric acid from arabinoxylan, presumably through a preference for hydrolyzing phenolic acid substituted arabinose side chains [243]. In the context of hemicellulose hydrolysis, the activity most often reported is hydrolysis of the α-(1 → 2)-glycosidic linkage of the arabinofuranoside to the xylan backbone. Some of these enzymes have been shown, however, to cleave linear or branched α-(1 → 5)-linked arabinan side chains found in some pectins, resulting in some confusion regarding the specificity of this enzyme class [244–247]. Although most assays are carried out on extracted arabinoxylan, *p*-nitrophenyl-arabinofuranoside has also been used as a substrate [248–250].

Esterases. Acetyl esterase (EC 3.1.1.6) removes acetyl esters from acetylated xylose and short-chain xylo-oligomers. It's polymer-acting counterpart, acetyl xylan esterase (EC 3.1.1.72), has a similar activity, but prefers polymeric xylan [242]. In addition to acetate-specific enzyme detection kits, HPLC or GC analysis of acetate release from native extracted xylan and chemically acetylated xylan, colorimetric substrates, such as *p*-nitrophenol acetate and β-napthyl acetate, or the fluorometric substrate, 4-methylumbelliferyl acetate are also used to assay acetyl esterases [242, 251]. The third esterase, ferulic acid esterase (EC 3.1.1.73), hydrolyzes the ester bond between ferulic acid or coumaric acid and the arabinose side chain of arabinoxylan. Assays for this activity are usually carried out using starch-free wheat bran or cellulase-treated gramineous biomass as a substrate and monitoring ferulic or coumaric acid released by HPLC or TLC. When preparing enzyme-treated substrates, care must be taken to employ phenolic-acid-esterase-free cellulases [242]. Other substrates include methyl and ethyl esters of the phenolic acids, as well as finely ground plant biomass [238, 252, 253].

Glucuronidases. In hardwood xylans, xylan α-1,2-glucuronosidase (EC 3.2.1.131) and α-glucuronidase (EC 3.2.1.139) are involved in debranching the xylan backbone through removal of α-(1 → 2)-linked glucurono- and 4-*O*-methyl-glucuronosides [254–257]. Although relatively little work on these enzymes has been carried out, Tenkanen and Siika-aho reported synergy with endoxylanase utilizing deacetylated birch glucuronoxylan. The same report also demonstrated that acetylation interferes with glucuronidase activity and that higher activity was observed on soluble softwood 4-*O*-methylglucuronoxylan [255]. Such synergy has also been reported by others [240, 255, 258]. *Para*-nitrophenyl-α-D-glucuronide is used as a substrate for α-glucuronidase [259], whereas xylan α-1,2-glucuronosidase is specific for an α-(1 → 2)-linked glucuronoside. Some glucuronidases, including membrane-bound enzymes, have been found to prefer glucuronoxylo-oligomers as substrates [258, 260–262]. One report demonstrated the specific requirement for the 4-*O*-methyl group for efficient binding and positioning of the side chain in the enzyme active site [256].

Hemicellulose Depolymerization Enzymes

As noted for cellulases, hemicellulose depolymerizing enzymes are divided into three classes; endo-acting enzymes, exo-acting enzymes, and oligomer-hydrolyzing enzymes. Although mechanisms of hemicellulose hydrolysis have been steadily studied over the years, they have not received the attention given to cellulose hydrolysis. Despite this, a general pattern of degradation is beginning to emerge. Although there are specific examples of endo-acting enzymes requiring side chains for maximal activity [263], the majority of the endo-acting hemicellulose hydrolases tend to be more active on debranched or partially debranched hemicellulose, especially in the case of xylanases. The limitation on this increased activity is probably due to solubility analogues of the polysaccharides, which tend to become more insoluble as the debranching process continues. Decreasing chain length from the activity of endo-hemicellulases mollifies this, allowing the shorter, less substituted fragments to remain soluble. Exo-acting enzymes, which probably fall into reducing- and nonreducing-end-specific groups and oligomer-hydrolyzing enzymes, also require debranching as a precursor to maximal activity. Overall, a balance must be met among removing the branching side chains from the polysaccharide backbone, decreasing the average chain length, and hydrolyzing the oligomers into free monomers, all while maintaining enough solubility of the fragments to allow enzyme access. The concerted action of the various hemicellulase enzyme classes probably accounts for the high synergy observed when the enzymes are used in concert [240].

Xylanases. Depolymerization of the xylan backbone is mediated by endoxylanases with oligomers hydrolyzed by β-xylosidases. Structurally, the endoxylanases are divided into glycosyl hydrolase families 10 and 11. These enzyme families are similar in that they both depolymerize xylan via the Koshland-type, two-step catalysis that leaves products with retained stereochemistry of anomeric configuration. Family-10 enzymes typically yield lower molecular weight products (tetramers) than do family-11 enzymes (pentamers) [264]. This is likely a function of the difference in binding sites, with family-10 enzymes having a binding site that recognizes shorter oligosaccharides than do family-11 enzymes [257, 265]. Although these enzymes are active on native branched xylan, debranching may increase their activity [258, 266].

Assays for endoxylanases follow the same general patterns as endocellulase assays. Viscosity reduction, reducing sugar production, dye-release, solubilization, zymogram analysis, and colorimetric/fluorometric analogues are all used to determine endoxylanase activity [267–275]. DNS detection of reducing sugars from xylan is the most cited method. Endoxylanases tend to have a preference for polymers of a certain degree of polymerization (DP). Bray and Clarke reported a *Schizophyllum commune* endoxylanase with a preferred substrate DP of seven [276], whereas others reported enzymes exhibiting true endo-type activity, with decreasing activity at lower DPs [263, 277, 278].

There are numerous reports of β-xylosidases that cleave short chain xylo-oligomers to xylose. In these cases, product detection was typically carried out by direct HPLC analysis or hydrolysis of *p*-nitrophenyl-β-D-xylopyranoside [279–282].

Xyloglucanases. Xyloglucans are polysaccharide polymers composed of a linear backbone of β-(1 → 4) linked glucopyranose moieties with some monomers substituted with xylopyra-nose in an α-(1 → 6) linkage. The xylose side chains can in turn be substituted with one or more of the disaccharides α-(1 → 2)-L-fucosylpyranose-β-(1 → 2)-D-galactopyranose and α-(1 → 2)-L-galactopyranose-β-(1 → 2)-D-galactopyranose, with the fucose residues being found mainly in primary cell walls [283–288]. α-(1 → 2)-L-arabinofuranose has also been shown to be substituted onto either the main glucose chain or onto the xylose side groups [283, 289]. It has been shown that xyloglucans are acetylated through *O*-linkages to the arabinose or galactosyl side chains [283, 287, 289, 290]. Despite this side chain substitution, a specific acetylxyloglucan esterase has not been discovered.

There is an exo-acting enzyme that acts on the nonreducing end of xyloglucan oligomers. Oligoxyloglucan beta-glycosidase (EC 3.2.1.120) is produced by *Aspergillus oryzae* and removes an α-xylo-β-(1 → 6)-D-glucosyl dimer (isoprimverose) from the nonreducing end.

β-*Glucanases*. β-Glucan is a glucopyranose polymer containing either β-(1 → 3) or mixed β-(1 → 3), β-(1 → 4) linkages. The ratio of (1 → 4) to (1 → 3) linkages varies by species and gives specific properties to individual β-glucan polymers. Because of the differences in the linkages, different enzymes are required to cleave the two forms of β-glucan [291–296]. We now know that β-glucan can be degraded by glycosyl hydrolase family-12 enzymes (EC 3.2.1.4). Although these endo-acting enzymes are active on β-(1 → 4) glycosidic linkages, they are differentiated from other β-(1 → 4)-acting enzymes by the distinction of being able to hydrolyze the β-(1 → 4) linkages in mixed β-(1 → 3, 1 → 4)-linked polysaccharides. Glucan endo-1,3-beta-D-glucosidase (β-(1 → 3) glucanase) (EC 3.2.1.39) is an endo-acting glycosyl hydrolase that acts on β-(1 → 3) glucan, but has very limited activity on the mixed linkage β-glucan. Endo-1,3(4)-beta-glucanase) β-(1 → 3,1 → 4) glucanase) (EC 3.2.1.6), is also an endo-acting glycosyl hydrolase. There is an exo-acting glycosyl hydrolase that is active on β-(1 → 3) glucan. Glucan 1,3-beta-glucosidase (EC 3.2.1.58) acts by processively releasing glucose from β-(1 → 3) glucan from the nonreducing end. One interesting report utilized an enzyme-linked sorbant assay in microtiter plates coated with biotinylated β-glucan to determine activity [297].

Mannanases, *gluco- and galactomannanases*. Whereas mannan is characteristically described as a linear β-(1 → 4) mannopyranose polymer, galactomannan is composed of a polymeric β-(1 → 4) mannopyranosyl backbone highly substituted with β-(1 → 6) linked galactopyranose residues [298, 299]. The degree of substitution varies with source. Glucomannan, found mainly in the root of the konjac plant (*Amorphophallous konjac*) consists of a β-(1 → 4) linked mannopyranose and glucopyranose backbone in a ratio of 1.6:1 [300]. The backbone residues are substituted in a β-(1 → 3) linkage with several sugars and short oligosaccharides, as well as with *O*-linked acetyl groups about every 15 residues [301–303]. The structure, and hence the degradation, of mannan is very analogous to cellulose, both being linear β-(1 → 4) linked monosaccharide polymers [304]. Mannan, however, is found in only a few particular plants, notably in the endosperm of the ivory-nut from the Tagua palm (*Phytelephas macrocarpa*) and a few other plants [304]. In glucomannan and galactomannan, as with xylan, degradation requires both debranching and depolymerizing enzymes, which work in synergy [240]. There are two major types of galactomannan, differentiated mainly by the source and degrees of galactose substitution, locust bean gum, derived from the carob tree (*Ceratonia siliqua*), contains an average of 2,000 sugar residues, with a galactose about every 3.5-mannosyl residues. The other major commercial source is guar gum, from the seed of the leguminous shrub *Cyamopsis tetragonoloba*. Guar gum contains more galactose residues than locust bean gum, having a galactose every 1.5–2 mannose units, and is longer, with residue counts of around 10,000.

The enzymes involved in depolymerization of the mannans consist of β-mannanase (EC 3.2.1.78), the endo-acting enzyme, and β-mannosidase (EC 3.2.1.25), which produces mannose from the nonreducing end of the mannose chain [305–307]. Debranching of galactomannan is primarily carried out by α-D-galactosidase (EC 3.1.2.22) [308, 309]. Tenkanen and co-workers have also reported an acetyl glucomannan esterase active on the acetyl side chains in glucomannan [236, 310, 311]. There are little other data on specific debranching enzymes involved in degradation of glucomannan. Assays for mannan hydrolysis have been carried out using extracted polysaccharides as substrates, colorimetric analogues, and dyed polysaccharides. Dyed polysaccharides have been utilized to determine activities of cultures on various polysaccharides, both as activity screens and as quantitative measures. These substrates include both soluble and insoluble forms (dependent mainly on the properties of the native polysaccharides) and include azurine (azo-), Remazol Brilliant Blue (RBB), and Ostazin Brilliant Red (OBR), among others [312–315]. For azo-, RBB-, and OBR-linked substrates, clearing zones on Petri plate or acrylamide gel agar overlays indicate active colonies or protein bands [316–320]. The crosslinked version (AZCL-polysaccharides, Megazyme, Inc. Bray, Ireland) has also been used to screen for activity of various glycosyl hydrolases [321–323]. In the case of AZCL substrates, the result of activity is a blue halo surrounding active colonies or dye release into microtiter plate wells. The authors have used this technique extensively to screen both environmental samples and recombinant libraries for glycosyl hydrolase activities (Fig. 33.23).

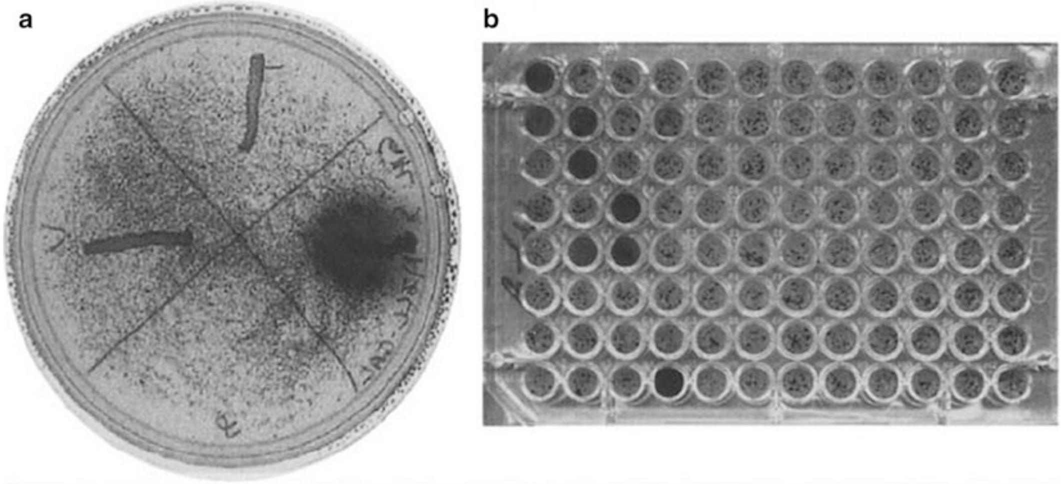


Fig. 33.23 AZCL-polysaccharide hydrolysis in (**a**) Petri plates and (**b**) microtiter plates. The dark particulates are the AZCL-β-Glucan (**a**) and AZCL-galactan (**b**). Soluble blue dye is released upon hydrolysis

Pectinases

In addition to cellulose and the hemicelluloses, pectins are a third class of polysaccharides found in the cell wall matrix of plant cells. Further information and good structural diagrams can be found in the recent review by Ridley et al. [324]. Found mainly between adjacent cell walls in the middle lamellae, pectins fall into three classes differentiated by their backbone structure and branching patterns [325, 326]. Homogalacturonan (xylogalacturonan) is comprised of α-(1 → 4)-linked galacturonic acid chains containing xylose side chains and makes up the smooth region of pectin [327]. Homogalacturonan is methylated through ester-linkages to the galacturonic acid residues. Once in place, pectin methyl esterases (EC 3.1.1.11) remove these side chains and allow formation of the gel matrix [328]. The rhamnogalacturonans make up the "hairy" region of pectin. In rhamnogalacturonan I, (RG I) the backbone chain is comprised of the disaccharide (→4)-α-D-galacturonic acid-α-(1 → 2)-α-L-rhamnopyranose-(1→). The rhamnose is typically substituted at the C-4 position with a branched chain of sugars made up of either galactose or arabinose or a combination of both and other sugars. The arabinose residues can be derivatized with ferulic acid. The galacturonic acid residues in the backbone are usually *O*-2 or *O*-3 acetylated and *O*-6 methylated. The structure and substitution patterns of RGI vary widely across plant species. Where the majority of the side chain is comprised of arabinose, the side chains are referred to as arabinans. These arabinans are predominantly α-(1 → 5) linked arabinofuranosyl residues substituted at either or both the *O*-2 and *O*-3 positions [326]. Side chains comprised of galactose residues are referred to as galactans. When these galactans are further substituted with arabinan chains, they are referred to as arabinogalactans [329]. In contrast to RG I, the structure of Rhamnogalacturonan II (RG II) is highly conserved across the plant kingdom [330]. It is comprised of 28 glycosyl residues, of which seven are found in the galacturonic acid backbone. The backbone is specifically branched at four points (designated A–D chains) with some unusual sugars such as 2-*O*-L-fucose and 3-deoxy-D-manno-2-octulosonic acid, aceric acid, and apiose [330]. The structures of each branch are known, although the exact point of attachment of branch D is still unclear.

As is apparent from the complex structure, there is a diverse enzyme suite required to hydrolyze pectin. As in hemicellulase systems, there are depolymerizing and debranching enzymes, mainly esterases, that act synergistically [315, 331]. The depolymerizing enzymes include both glycosyl hydrolases, which cleave glycosidic bonds by an acid–base catalysis mechanism, and polysaccharide lyases, which hydrolyze the glycosidic bond through a β-elimination mechanism, resulting in a double bond between the C-4 and C-5 of the new nonreducing end [315]. Reviews by Kashyap et al. and by Kennedy and Methacanon outline the pectinase enzymes in detail [143, 332]. Assay techniques involve the usual assortment of reducing sugar production, viscosity reduction, HPLC analysis, and dye release [315, 331, 333–338]. Ruthenium red staining in plates and zymograms have also been used [331]. Because of its solubility, pectin incorporated into plates can be detected by precipitation with hexadecyl-trimethyl-ammonium bromide, resulting in clear halos of hydrolysis around active colonies [339]. As with other polysaccharide degradation studies, structural determination of products has been carried out by nuclear magnetic resonance spectroscopy [333].

Chemical Catalysis for Biomass Deconstruction

Introduction

Deconstruction is a somewhat misleading term, in that it suggests processes exhibiting low selectivity or harsh conditions. To the contrary, *selective* biomass deconstruction can be a powerful tool for use within the biorefinery, as it provides both monomeric and polymeric materials for eventual conversion into final refinery outputs. Biomass deconstruction is a separation process highly analogous to petrochemical refining. In any process designed to convert a source of raw materials into products, the raw material almost always requires some form of preliminary upgrading in order to separate a complex mix of materials into individual streams of initial building blocks. The petrochemical industry "deconstructs" crude oil, using distillation to make mixed hydrocarbon streams for various fuel uses, or cracking (primarily using chemical catalysis) to prepare the industry's primary building blocks of ethylene, propylene, BTX, butenes, and methane for chemical uses [340].

The biorefinery is no different. Renewable raw materials are complex mixtures of different biopolymers, and chemical deconstruction processes (as differentiated from biochemical processes described earlier in this chapter) may be used to separate these biopolymers from one another. However, deconstruction processes optimized for the biorefinery offer more flexibility than those used in the petrochemical industry. The petrochemical industry tends to focus on refining processes that give low molecular weight (very roughly) (Fig. 33.24).

The wide range of different biomass materials can be somewhat daunting. However, biomass deconstruction processes greatly simplify this initial slate of raw material sources. Biomass provides three primary feedstocks: sugars, in the form of starch, cellulose, and hemicellulose; aromatics, in the form of lignin; and hydrocarbons, in the form of plant triglycerides. Further deconstruction of these polymers and oligomers leads to monomeric materials including individual sugars and related polyols such as glycerol from triglycerides, low molecular weight aromatics from lignin, and fatty acid hydrocarbons from triglycerides [341, 342]. Under proper conditions, deconstruction of the polymeric materials can also lead to structural modification of the monomeric units, leading to materials such as furfural, levulinic acid, and levoglucosan [343].

Chemical catalysis can play a key role in each of these deconstruction processes, promoting conversions of raw materials into individual polymers, and individual polymers into low-molecular-weight compounds. This section reviews examples of each type of process, showing how catalysis leads to core building blocks for the biorefinery, analogous to the initial building blocks manufactured by the petrochemical industry. Although highly important to overall biorefinery operation, further catalytic transformations of this initial set of building blocks to chemical products is not included in this discussion.

The selective deconstruction of biomass for biorefinery use is of particular importance when viewed in the context of biomass availability. Biomass is exceedingly abundant. A vast amount of renewable carbon is produced in the biosphere, as much as 170 × 10^9 metric tons is fixed annually. This amount could supply almost all domestic organic chemical needs, currently about 7–8% of our total nonrenewable energy and materials consumption [344–347]. When measured in energy terms, the amount of carbon synthesized is nearly ten times the world consumption [348]. Cellulose, the most abundant organic chemical on earth, has an annual production of about 90 × 10^9 metric tons [349–351]. Lignin production by the pulp and paper

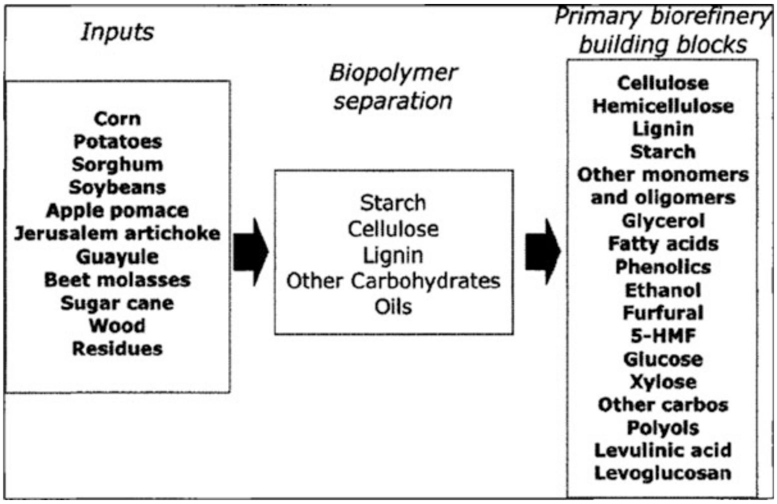


Fig. 33.24 Biomass deconstruction is a key step in its conversion to chemical products and fuels

industry is 30–50 × 10^6 metric tons/year [352]. The yearly availability of renewables is most important, inasmuch as—in contrast to petrochemicals—it makes this resource almost unlimited, if used in a sustainable manner. Some projections have world oil production peaking in the next 5–10 years [353, 354]. It is unlikely that demand will decrease in line with production. In the United States, energy consumption has increased by more than 28%—about 21 ekajoules (EJ)—during the last 25 years, with more than half of this energy growth—about 11 EJ—occurring during the last 6 years, indicating that other feedstock sources will be needed [355]. Chemical catalysis has been a key contributor to the success of the petrochemical industry, with more than 90% of all processes using some form of catalysis [356]. It is reasonable to assume that catalysis tailored for the specific features of biomass raw materials will play a similar role within the biorefinery, both to effectively and sustainably use this resource and to help address the growing needs of the world.

Scope of the Review

Chemical catalysis for processes leading to cellulose, hemicellulose, lignin, and their derivatives is an extremely broad topic. It is also somewhat hard to define for biomass. A number of different processes exist for separation of biomass into its constituent biopolymers. The corn wet mill produces carbohydrates as starch. The kraft process for pulping of wood produces carbohydrates as cellulose, and under certain conditions, isolable lignin. The kraft process itself is stoichiometric, but extensive research has been done on the use of catalytic amounts of additives to promote this and related separations [357]. Many other separation processes have been categorized under the term “pretreatment” [358–361] and are linked closely to the generation of fermentable sugars for fuel ethanol production, but can also provide polymeric materials for chemical use. Some pretreatments are solvolyses, which are technically noncatalytic. Yet the active component of the solvent (e.g., a proton) is not consumed within the molecular level mechanism, and serves to promote the reaction in a catalytic sense. Other systems are autocatalyzed through generation of a catalytically active species over the course of the separation. More recently, a new class of pretreatment has been getting more scrutiny. Ionic liquids, i.e. salts that are molten at "reasonable" temperatures, have shown the capacity for completely solubilizing biomass. Ionic liquids are being studied for several applications in biomass conversion including complete solubilization, swelling of fibers for increased digestibility, as a solvent system for increasing enzymatic digestion, and as a co-pretreatment in conjunction with other "standard" pretreatments.

The subsequent deconstruction of biopolymers into single, low-molecular-weight chemicals, allows the identification and definition of catalyst to become clearer. Chemical catalysis has been used to improve many of these processes, and several examples of these conversions are described below.

The options for practical operation of a biorefinery are also complex. Very broadly, the biorefinery will have the capability of deconstructing biomass into several categories of outputs including:

- Constituent polymers of biomass
- Oligomeric fractions of these polymers
- Monomers comprising these polymers or oligomers
- Individual chemicals from structural reorganization of other fractions

Ideally, biorefinery operation will be selective and flexible enough to switch among these categories depending on the needs of the market. Moreover, these categories could, in principle, be combined to provide diverse product profiles [362, 363]. In one instance, the biorefinery would provide cellulose, hemicellulose, and lignin. However, modification of conditions could further transform the initially produced carbohydrate polymers in situ, for example, hydrolyzing the hemicellulose to xylose, and then converting it to furfural. The Acetosolv process for the separation of wood is operated in this manner [364]. Pulping of wood with the Formacell process (pulping with HOAc/HCOOH mixtures) can produce cellulose, lignin, furfural, and 5-hydroxymethylfurfural (5-HMF) as primary products [365]. Thus the exact choice of operating mode and resulting product profile can be strictly dependent on the needs of the biorefinery operators.

Even with this potential variability in operation types, it is anticipated that the biorefinery industry will migrate toward a small group of easily produced primary building blocks from which a much larger group of final products will be derived. This approach would directly parallel the petrochemical industry, which uses seven primary building blocks for the eventual production of tens of thousands of marketplace products. The analogous group of primary building blocks for the biorefinery has not yet been defined, making the potential matrix of deconstruction processes and initial outputs complicated. Using carbohydrate polymers again as an example, hemicellulose, xylose, or furfural could each be a biorefinery primary building block as they are all linked by chemically catalyzed deconstruction processes.

In this section, we seek to identify materials that are the reasonable first structures to arise from biomass deconstruction, and to describe how chemically catalyzed processes are being developed for their production. For that reason, commercially practiced processes that use catalysis, such as the reduction of glucose to sorbitol, are mentioned only briefly or not at all. Chemical catalysis will certainly play an additional role in the further conversion of these initial building blocks into secondary intermediates or final marketplace products (e.g., oxidative conversion of levulinic acid into succinic acid), but such multistep possibilities are outside the scope of this discussion.

Separation of Biopolymers from Biomass Raw Materials

Biomass raw materials are made up primarily of the three biopolymers, cellulose, hemicellulose and lignin, and each of these materials finds use as a chemical feedstock. A wide variety of methods has been described for their separation and isolation. The paper industry describes these processes as pulping, whereas the biorefinery industry more normally refers to them as pretreatments. The goal in each case is the same: production of separate biopolymer streams for conversion to final product outputs. The uses of chemical grade cellulose are well known in industry and the literature and include cellophane and various cellulose esters and ethers [213]. Certain forms of lignin are also items of commerce [366]. Chemical catalysis of the following pretreatment/pulping processes has been examined to improve the initial separation of lignocellulosics into their constituent polymers.

Dilute Acid Pretreatments

Within the context of the biorefinery, the use of dilute acid is probably the most widely studied technology for the separation of the starting biomass raw material into its individual components. This process has been and continues to be extensively investigated, and the reader is referred to the many reviews available in the literature [88, 358–361, 367–373].

Solvent Separation Processes

Biomass separation based on the use of organic solvents ("organosolv" processes) offers more selective access to the biopolymeric constituents of renewable raw materials, as such operations normally give separate and high-quality process streams of cellulose, hemicellulose, and lignin. Several reviews on organosolv technology and the mechanism of separation are available [374–380]. Most examples of organosolv processes use volatile alcohols, such as methanol and ethanol, however, a huge variety of additional solvents has been employed, including 1,3-BDO, 1,4-BDO [381], ethylene glycol [382], glycerol [383], phenol [384], and organic acids under the names of Formacell, Milox, and Acetocell [385–390]. Attempts to commercialize organosolv technologies have also appeared as the Alcell, Organocell, or ASAM processes [380, 391–397]. Because these processes were primarily focused on new methods of producing cellulose pulp for the paper industry, their expense precluded commercialization. However, the selectivity of organosolv technology makes it well suited for the production of chemical building blocks from renewable raw materials.

In almost all organosolv processes, chemical catalysis plays a necessary role, as solvents alone do not function effectively for the separation of biomass [398]. The most commonly employed processes, based on the treatment of biomass with aqueous alcohols at elevated temperatures, are autocatalyzed. Acetic acid is generated during the separation process through hydrolysis of acetate groups present on the hemicellulose polymer [392, 399–401]. Alternatively, acid can be added to the separation medium prior to the process. Adding acid catalyst normally allows lower separation temperatures and milder conditions. Chemical catalysis has proven to be of particular importance for the organosolv separation of softwoods. Normally, softwoods are poor substrates in solvent-based processes. However, the addition of alkaline earth metals to alcohol separation processes has resulted in effective softwood separation [398, 402–404].

An example of an acid-catalyzed organosolv biomass separation is the clean fractionation (CF) process, developed at the NREL (Fig. 33.25) [405]. The process treats the biomass with a ternary mixture of methyl isobutyl ketone (MIBK), ethanol (EtOH), and water (H_2O) in the presence of sulfuric acid. The separation is carried out using a 16/34/50 ratio of MIBK/EtOH/H_2O for 56 min at 140°C in the presence of 0.025–0.2 M H_2SO_4 as standard conditions. The solvent mixture selectively dissolves the lignin and hemicellulose components, leaving the cellulose as an undissolved material that can be washed, fiberized, and further purified. The soluble fraction containing the lignin and hemicellulose is treated with water, causing a phase separation yielding an organic phase containing the lignin and an aqueous phase containing the hemicellulose. Solvent removal gives CF lignin in high yield as a shiny, brown, free-flowing powder exhibiting properties consistent with other organosolv lignins, and a generally lower-molecular-weight profile and greater solubility in polar organic solvents than other lignins. CF appears to be a very general process, and can be used for fractionation of a wide range of biomass feedstocks including corn stover, newsprint, bagasse, corn fiber, and rice straw.

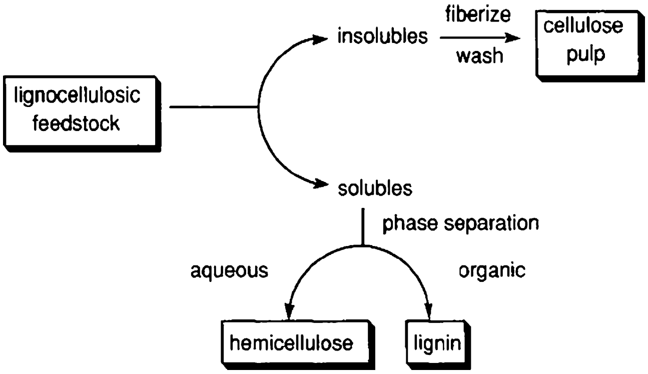


Fig. 33.25 The NREL clean fractionation process

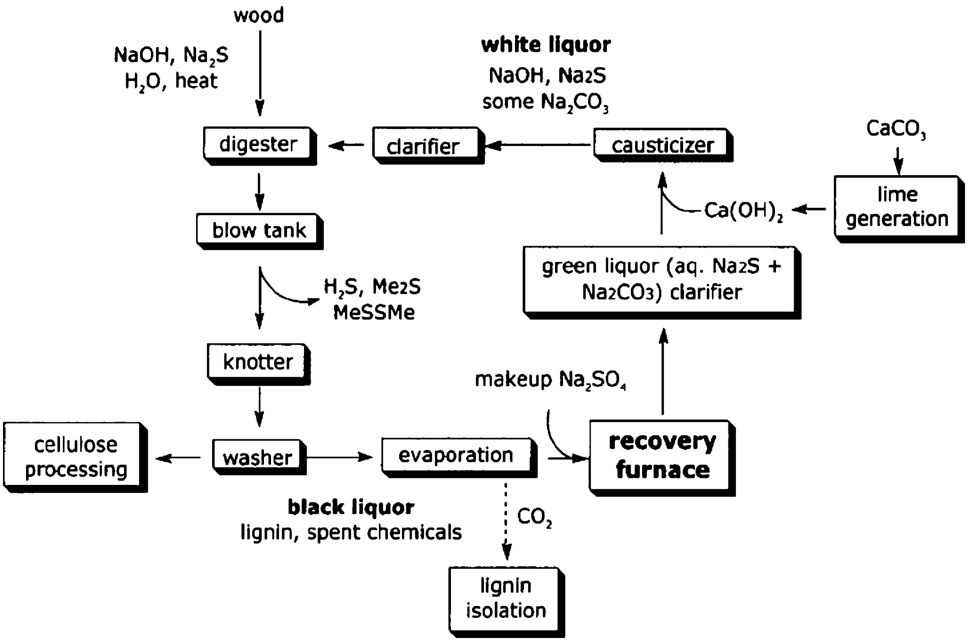


Fig. 33.26 Overview of the kraft process

Steam Explosion and Related Processes

The steam explosion process is a rapid and decades-old treatment process for lignocellulose that releases individual biomass components through steam impregnation under pressure, followed by a rapid pressure release. Steam explosion is carried out using high-pressure steam and short contact times [406–409]. This process has generally been used as a method for preparing cellulose pulp but can also serve as source of lignin. Alkaline extraction after explosion recovers more than 90% of the lignin when wood is used as the feedstock [410–414].

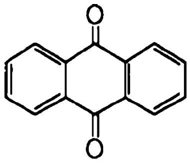
Catalysis of steam explosion processes has been carried out, primarily through the addition of acid [371, 373, 415, 416] or SO_2 to the separation. Added SO_2 has been shown to exhibit benefits in conversion yields and separation conditions [417, 418]. The characteristics of the oligosaccharides and lignin from these processes have been examined [410–414, 419].

The Kraft Process

Kraft pulping is the most widely used process in the world for separating cellulose from wood, but is not normally considered as a potential operating unit for biorefineries. Certainly for new construction, use of the kraft process would face considerable challenges, including the enormous capital cost of a new plant. Nonetheless, the existing infrastructure of the pulp and paper industry consumes more than 100 × 10^6 metric tons/year of wood, the majority of this with kraft technology, offering a possible opportunity for any underused capacity [103].

The process is conducted at the high pH in the presence of sodium sulfide, sodium hydrosulfide, and sodium hydroxide at 150–180°C for about 2 h (Fig. 33.26). Along with the primary cellulose product, the process generates a solution containing dissolved lignin, most of the wood's hemicelluloses, and spent pulping chemicals ("black liquor"). This solution is used in the kraft recovery boiler as a source of power for the mill and fuel for chemical recovery operations crucial to economic operation. The chemistry of pulping reactions has been the subject of much study, and summaries are available the literature [420–422].

The kraft process itself is stoichiometric, however, the process can be improved through the addition of catalytic amounts of anthraquinone (AQ, 1). The use of AQ in alkaline pulping processes has been widely examined by the industry [357, 423, 424].



1

As a catalyst, AQ offers a number of benefits, including an increase in pulping rate, a decrease in the amount of pulping chemicals used, a decrease in degradation of the cellulose during pulping, and improvement of cellulose bleaching processes. In the context of the biorefinery, these improvements lead to faster production and higher yield of a more easily purified cellulose biopolymer. AQ is typically added at levels of 0.05–0.1% based on the weight of raw wood feedstock used. Considerable mechanistic study has been performed to determine how AQ exerts these effects [425, 426].

Deconstruction of Biorefinery Raw Materials and Biopolymers to Monomeric Products

Initial separation of biomass raw materials can yield separate streams of biopolymers, each of which has potential utility as a product within the biorefinery. Further selective deconstruction processes can convert these biopolymers into their individual monomeric units, or to structurally rearranged materials. The following are examples of processes that have been improved through the use of chemical catalysis.

Polymeric Carbohydrate Hydrolysis Processes

The hydrolysis of polysaccharides into monomeric sugars is a well-studied process; its mechanism is still believed to be accurately described by the work of Saeman [427]. Its significance is linked to the production of fermentable sugars for ethanol manufacture. Many of these processes are based on acid catalysis and overlap closely the pretreatment methods described earlier [358, 359, 368]. Treatments with both concentrated and dilute acids are well known as methods for the hydrolysis of cellulose and hemicellulose, as is the use of organic dicarboxylic acids as alternative catalysts to mineral acids [428].

However, further cellulose deconstruction is considerably more difficult after acid hydrolysis. Commercial processes for acid hydrolysis of cellulose were known in the early part of the twentieth century [429]. The use of either low- or high-acid concentrations has disadvantages. At high concentrations, corrosion-resistant reactors and an effective acid recovery process are needed, raising the cost of the intermediate glucose. Dilute acid treatments minimize these problems, but a number of kinetic models indicate that the maximum conversion of cellulose to glucose under these conditions is 65–70% because subsequent degradation reactions of the glucose to HMF and levulinic acid take place. The modern biorefinery is learning to exploit this reaction manifold, because these “decomposition” products can be manufactured as the primary product of polysaccharide hydrolysis (see below).

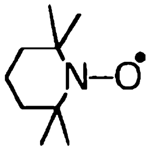
Carbohydrate Pyrolysis

Pyrolysis of C6 mono- and polysaccharides deconstructs the polymer chain forming either levoglucosan or levoglucosenone. In the presence of an acid catalyst (normally H_3PO_4), levoglucosenone is also formed in low yields (Fig. 33.27) [430–435]. The transformations are straightforward, but separation of these potential biorefinery building blocks is complicated by the formation of considerable amounts of tar. Nonetheless, attempts have been made to develop levoglucosan as a commercial product [343].

The formation of levoglucosan and levoglucosenone is found to be strongly dependent on the presence of catalytic amounts of various cations during pyrolysis. However, the literature is inconsistent regarding the effect of these cations. By varying the amount of acid catalyst used, Faix and coworkers were able to control the relative amount of the two products [436, 437]. Transition metal catalysts were also found to influence the course of the pyrolysis. Addition of Fe or Cu increased the yield of levoglucosan formation from wood or newsprint [438]. The presence of Ca ions also increased the yield of these products, whereas catalytic amounts of Na ions were found to decrease the pyrolysis yields [439]. In contrast, earlier work indicated that Li, K, and Ca ions all had a negative effect on the pyrolysis process [440].

Carbohydrate Oxidation

A considerable amount of recent work has focused on the oxidation of polymeric and monomeric carbohydrates in aqueous media. In the context of the biorefinery, these processes could be used for the preparation of oxidized carbohydrates as primary outputs of biomass deconstruction. Of particular interest are processes catalyzed with stable oxygencentered radicals such as the nitroxyl radical TEMPO (2) (2,2,6,6-tetramethylpiperidinoxyl) and using bleach as the stoichiometric oxidant.



TEMPO, **2**

The reaction appears to be well suited for selective conversion of biomass carbohydrates into their corresponding oxidized derivatives. This system serves to oxidize several polymeric carbohydrates including starches and pullulan [441–444]. More recent work has described the TEMPO-catalyzed introduction of carboxyl groups in native cellulose and its different morphological forms [445–447], cellulose derivatives [448], and the surface of cellulose nanocrystals [449]. The related biopolymer, chitin, also is oxidized under these conditions [450].

Under proper conditions, biomass deconstruction will generate streams of monomeric or oligomeric carbohydrates, with glucose and xylose (from cellulose/starch or hemicellulose, respectively) being the most likely initial feedstocks. Glucose oxidation in the presence of TEMPO and bleach has been used for the production of glucaric acids (see below) by selective, high-yield oxidation of the initial and terminal carbons of the starting sugar [451–454].

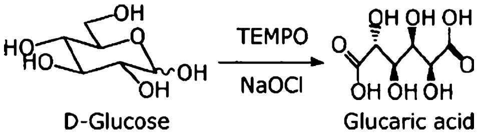
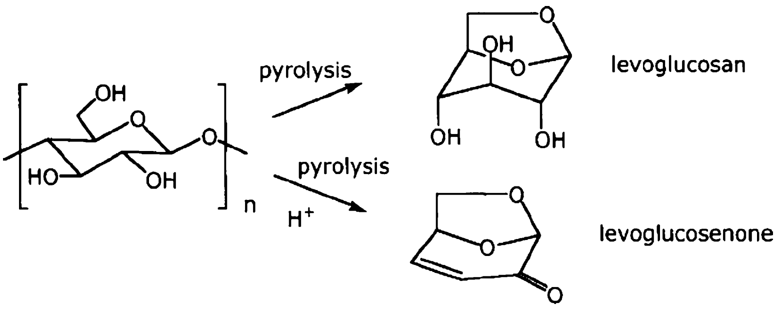


Fig. 33.27 Production of levoglucosan and levoglucosenone



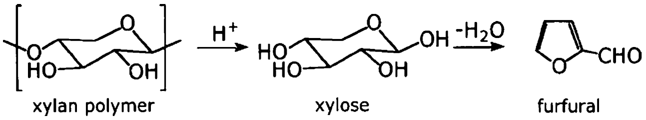
These materials can serve as novel polymer components in the production of new polyhydroxylated polyamides [455] and as starting points for the production of other biorefinery intermediates. Conventional processes for the conversion of sugars to aldaric acids use HNO_3 as the oxidizing agent. However, yields of aldaric acids using these processes are only fair, and the use of HNO_3 as the oxidant is potentially hazardous and environmentally unfriendly (NO_x emissions). Development of single-step biorefinery processes could make aldaric acids a useful primary building block of the biorefinery.

Other carbohydrates can also undergo selective, high-yield oxidation. For example, sucrose (the primary product of the sugar industry) converts into the corresponding tricarboxylic acid [456].

Carbohydrate Dehydration

The removal of water from initially formed biomass sugars is an important process for the production of primary biorefinery building blocks. Of particular interest are 5-HMF and levulinic acid (from the dehydration of glucose or other C6 sugars) and furfural (from xylose dehydration). Recent research has led to new catalytic processes for the production of each of these materials.

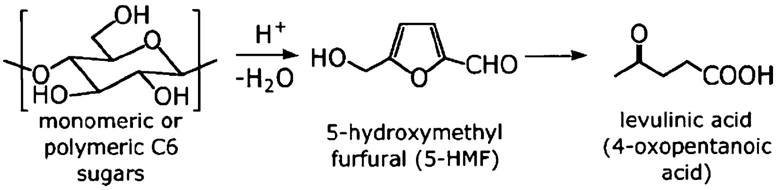
Furfural. Dehydration of xylose, available from biorefinery hemicellulose, leads to the production of furfural [457–459].



Although the process for furfural production has been known for decades and is applicable to most any C5-sugar-containing feedstock [460–463], it is not highly efficient, with commercial yields generally around 50% [464]. Current research is directed at improving the production of furfural from xylose or hemicellulose with new catalytic transformations. Improvement of process conditions for the standard H_2SO_4-catalyzed furfural production from the xylose in corn cobs was realized by use of HCl, leading to an 85% furfural yield [465]. A related process charges dry feedstock with catalytic gaseous HCl, minimizing decomposition reactions through a more rapid removal of furfural from the reactor [466].

The need to recover homogeneous acid catalysts in the production of furfural has spurred the investigation of acidic heterogeneous catalysts to promote the dehydration of C5 sugars. A relatively high conversion of xylose to furfural was reported using heterogeneous sulfonated mesoporous silicas [467]. Rapid removal of furfural from the reaction mixture using supercritical CO_2 has been examined as a method to improve furfural yields [468, 469]. Combining supercritical CO_2 extraction with xylose dehydration in the presence of sulfated titania and zirconia catalysts, led to high furfural yields at high xylose conversions [470]. Faujasite and mordenite catalysts have been examined as acid catalysts [471]. The presence of metal oxides has also been reported to improve furfural yield [461].

Levulinic acid. Dehydration of glucose or other monomeric and polymeric C6 sugars leads to the direct formation of levulinic acid (LA) as a potential primary building block for the biorefinery, and several reviews have described its potential commercial utility [472, 473]. The preparation of levulinic acid is not difficult, although the mechanism of its formation from carbohydrates is complex, and offers several alternative decomposition pathways (33.3) [474].



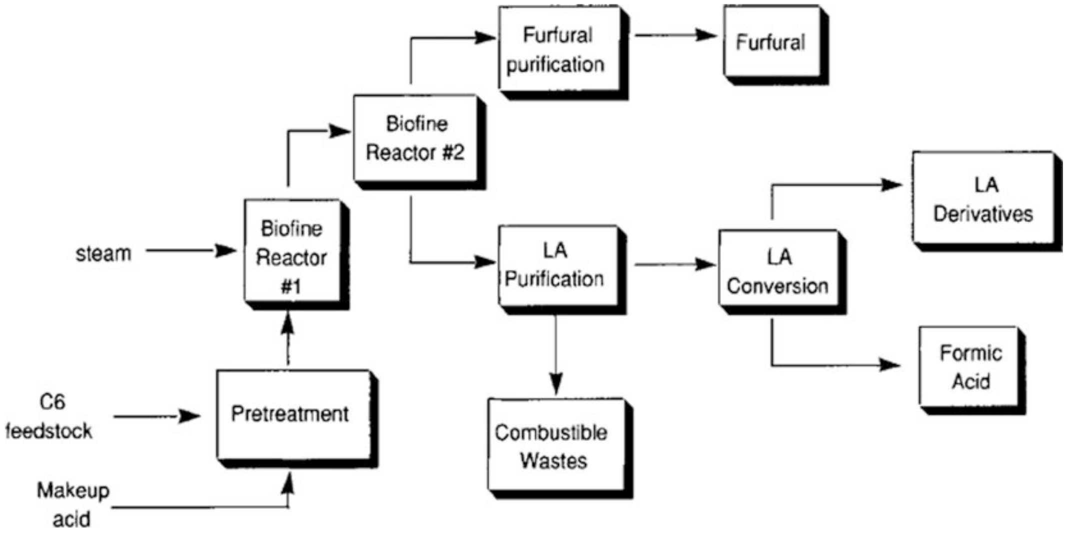


Fig. 33.28 The biofine process

Dehydration of glucose in the presence of microporous pillared clays has also been explored but, the selectivity to levulinic acid was low [475].

The most widely used approach for producing levulinic acid is the direct hydrolysis of biomass or carbohydrates, catalyzed by aqueous acid [473, 475–478]. However, this forms large amounts of side products and intractable materials, requiring an expensive separation that drives up its cost. Several years ago, a new process developed by the Biofine Corporation eliminated many of the existing production problems. The Biofine process also uses acid hydrolysis of carbohydrate sources (cellulose, starch, paper mill sludge, urban waste paper, agricultural residues, or paper fines) to produce LA, but in a novel, two-reactor system that minimizes sideproduct formation and associated separation problems (Fig. 33.28) [479, 480].

The process supplies carbohydrate-containing materials to a first reactor where they are briefly hydrolyzed at 210–230°C in the presence of mineral acid. This initial hydrolysis produces hydroxymethylfurfural, which is removed continuously and supplied to a second reactor. In the second reactor, the hydroxymethylfurfural is hydrolyzed further at 195–215°C for 15–30 min to produce levulinic acid, which is again continuously removed. The yield of LA is 75% or greater, based on the C6 content of the carbohydrate containing starting material, one of the highest reported. Two real strengths of this process are its relative insensitivity to the type of feedstock employed and the wide availability of feedstock sources. Any material containing C6 sugars will serve as a good starting material. The result is a cost-effective process, making it suitable as a starting material for a wide variety of products. Economic projections indicate that the LA production cost using this process could be as low as $0.08–$0.12/lb depending on the scale of the operation. This cost is consistent with the use of levulinic acid as a primary biorefinery building block.

5-HMF. Dehydration of C6 sugars also produces 5-HMF (shown above), which is an intermediate within the dehydration mechanism leading to levulinic acid. 5-HMF has been suggested as a potential primary building block for the biorefinery [481–483], but its relatively high instability and associated difficulties in its isolation have hindered its development.

Polymeric carbohydrates available from the biorefinery can serve as starting materials for 5-HMF. Recently, $LaCl_3$ has been used to catalyze the conversion of cellulose to 5-HMF (along with glucose, levulinic acid, and cellobiose) at elevated temperatures in water [484].

A primary reason for low yields of 5-HMF is its rapid conversion to levulinic acid in aqueous media. However, catalysis of the transformation with lanthanides has led to dramatic increases in the yield of 5-HMF. The effectiveness of different lanthanide cations has been surveyed [485, 486]. In nonaqueous solvents, such as DMSO, almost quantitative yields of 5-HMF have been reported [487]. Performing the dehydration in the presence of activated carbon (to adsorb the generated 5-HMF) has also been reported as effective [488].

Recent work has investigated the selective catalytic dehydration of fructose (available from corn starch) and related sugars such as glucose, sucrose, and inulin as a source of 5-HMF. Standard preparation of 5-HMF is carried out using aqueous acid catalysts [489–491]. The process can be improved by dehydrating the fructose with an Amberlyst catalyst in an ionic liquid as the reaction medium. Other solid acid catalysts have been reported [492–494]. Sub- and supercritical acetone has also been used as an alternative medium to improve 5-HMF production [495]. 5-HMF yields of 80% were reported under these conditions in comparison to control reaction yields of only 50% [496]. A number of heterogeneous catalysts have been used to promote this transformation [478]. The use of vanadyl phosphate doped with trivalent metal cations has been reported as effective for the production of 5-HMF in high yield without the formation of polymeric byproducts [497]. Niobium phosphate catalysts have also been investigated [498, 499], as well as acidic mordenites [500]. A mechanism for the catalytic reaction over zeolites has also been suggested [501].

Lignin Deconstruction

Deconstruction of lignin has been carried out for over a century, but in a largely nonselective manner. There is no doubt that lignin is widely available. It contributes as much as 30% to the weight of lignocellulosic biomass, making it second only to cellulose as a source of renewable carbon. Because lignocellulosic will be a major raw material source available to the biorefinery, isolation of sugars for use in one operating unit of the biorefinery guarantees a parallel lignin process stream. Moreover, the separation processes described earlier will provide separate carbohydrate and lignin components. Lignin's native structure (a complex network of oxygenated aromatic rings) certainly suggests that it could play a central role as a new chemical feedstock. But today's reality is that lignin is routinely relegated to very low value uses when considered as a biorefinery process stream. In the pulp and paper industry, the bulk of the removed lignin is used as a boiler fuel for chemical recovery. In other parts of the process, elegant bleaching sequences have been developed for the sole purpose of nonselectively converting residual lignin remaining with the cellulose into small, low-molecular-weight materials that are easily washed out [502]. Catalysis is playing a role in these processes [503–513]. Nonetheless, more selective catalytic lignin conversion processes continue to be examined [514].

General lignin oxidation processes. A vast amount of work has been done on lignin oxidation, primarily to understand and develop new pulp bleaching processes. More recently, the use of catalytic processes has been examined. If made selective, these types of processes could yield a broader family of lignin-derived primary building blocks. Because of the complexity of lignin's structure, much of the actual research has been carried out using lignin models. The oxidation of lignin models with HOOH in the presence of Co catalysts has been studied [515, 516].

Several more selective oxidation processes have been examined for the conversion of lignin to biorefinery building blocks. Oxidation of lignin residue from dilute acid pretreatment leads to a stream of mixed aromatic aldehydes [517]. Cu-catalyzed lignin oxidation with oxygen leads to the production of monomeric phenols [518]. Improved separation technology can allow effective use of these mixed streams within the biorefinery [519, 520]. Modeling has shown lignin conversion into quinones using cobalt and heteropolyacid catalysts [521–523]. and aromatic aldehydes with catalytic NO_2 in the presence of oxygen [524].

Phenol hydrogenation. In principle, appropriate lignin deconstruction processes will provide a stream of mixed phenols. Reduction of these phenols will leads to a new source of cyclic aliphatic alcohols of potential use in the manufacture of adipic acid derivatives. Several catalytic processes for these types of reductions have appeared for phenol and should be applicable to lignin-derived mixed phenols. Phenol itself is reduced to cyclohexanol in the presence of various heterogeneous catalysts based on Pd [525–530].

Vanillin. Catalytic oxidation of lignin has long been used as a source of vanillin. For the most part, yields in these processes tend to be very low, generally less than 10%. However, vanillin production can be used to provide a low-volume, high-value revenue stream within certain lignin-generating operations [531, 532]. The mechanism of the transformation has been reported [533–537].

Conclusions

Deconstruction of biomass raw materials offers the biorefinery access to a large number of both polymeric and monomeric primary building blocks. Biomass offers significant flexibility as to the product slate that various separation processes may realize. These materials, which are the biorefinery analogues to ethylene, propylene, BTX, and other building block chemicals of the petrochemical refinery, will serve as the platforms from which much larger families of products can be produced. Importantly, these processes also offer access to a wide range of new structures that can complement those currently available from the petrochemical industry. The challenge in using these materials will be development of appropriate technology tailored to accommodate these structures and facilitate their conversion into biorefinery intermediates and final products.

Thermochemical Biomass Conversion

Introduction

Biomass combustion has been a source of heat and light for thousands of years and is one of the oldest examples of thermochemical biomass conversion. More recently, biomass thermochemical conversion processes and technologies are receiving renewed attention as concerns about the sustainability of energy resources increase and mounting evidence of global climate change brought about

Table 33.5 Proximate and ultimate analyses of selected biomass feedstocks

	Eucalyptus	Poplar	Willow	Switchgrass	Rice straw	Wheat straw	Corn stover	Almond shells	Alfalfa stems	Black Thunder coal	Pittsburgh #8 coal
HHV, MJ/ kg, dry	19.18	19.46	19.34	18.36	14.71	17.64	18.06	19.01	18.63	29.75	31.80
Proximate wt.% as received											
Ash	0.48	1.16	0.85	4.22	17.30	6.55	4.75	2.80	4.78	5.08	7.90
Volatile	78.52	80.99	76.52	72.73	65.62	75.54	75.96	70.13	71.59	42.70	36.80
Fixed carbon	11.66	13.05	12.40	14.89	14.23	16.22	13.23	19.22	14.34	30.91	1.14
Moisture	9.34	4.80	10.23	8.16	7.35	6.58	6.06	7.85	9.29	21.30	54.16
Ultimate wt.% as received											
Moisture	9.34	4.80	10.23	8.16	7.35	6.58	6.06	7.85	9.29	21.30	1.14
C	44.89	47.05	44.07	43.04	35.42	41.96	43.98	46.20	42.79	55.01	78.02
O[a]	39.92	41.01	39.21	38.58	34.14	39.28	39.10	36.94	35.09	0.74	1.36
N	0.13	0.22	0.32	0.53	0.81	0.40	0.62	0.68	2.43	0.38	2.78
S	0.03	0.05	0.03	0.10	0.17	0.14	0.10	0.03	0.18	13.45	3.93
Cl[b]	0.05	<0.01	<0.01	0.46	0.54	0.21	0.25	0.01	0.45	0.06	0.09

by fossil fuel consumption continues to be revealed. Advances in crop production and harvesting, collection, and use of biomass from thinning operations for improved forest health, and utilization of biomass residues can provide lower-cost biomass feedstocks. Integrating lower-cost feedstocks with advances in higher-efficiency biomass thermochemical conversion processes for the production of fuels, chemicals, and combined heat and power, provides a number of realizable social, economic, energy, and infrastructure security benefits for the future [11, 538, 539].

As defined above, biomass materials can be quite varied with a wide range of physical, fuel, and chemical properties. This variability in the inherent properties of biomass resources determines conversion technology options that are appropriate for specific applications.

Thermochemical conversion applications are generally fuel-flexible when it comes to variations in biomass composition. The main biomass fuel properties that affect the long-term technical and economic success of a thermochemical conversion process are moisture content, fixed carbon and volatiles content, impurity (S, N, Cl) concentrations, and ash content. Collectively, these properties affect the energy density of biomass fuels. The higher heating value of biomass on a dry and ash-free basis is typically around 8,000–9,000 Btu/lb (18.5–20 MJ/kg). High moisture and ash contents reduce the usable energy of biomass fuels proportionally. Therefore, from an energy perspective, maximum system efficiencies are possible with dry, low-ash biomass fuels. However, drying biomass can be costly and low-ash biomass resources are generally considered premium fuels that tend to be more expensive.

The bulk density of biomass feedstocks is also quite low compared to traditional fossil fuels. Lower bulk densities and lower energy densities translate into higher costs for feedstock preparation, handling, and transportation, putting biomass conversion technologies at an economic disadvantage in current energy markets. Consequently, biomass utilization is most advantageous when the feedstock is either a high-volume waste product of an existing system or is readily available in close proximity to the conversion plant. Examples of this can be found in the pulp and paper and wood products industries. In fact, the pulp and paper industry is the largest generator and consumer of biomass energy in the form of process heat and electricity produced from black liquor recovery boilers and bark/hog fuel boilers. Agricultural practices for grain production also provide the potential to use localized residues such as straws and stover in biomass thermochemical processes. In the future, dedicated energy crops such as switch-grass, poplar, and willow could potentially provide a renewable sustainable source of fuel for thermochemical conversion processes.

Important biomass fuel properties for thermochemical conversion processes are reported as proximate and ultimate analyses. The proximate and ultimate analyses for selected biomass feedstocks are presented in Table 33.5. For comparison, the analyses from two selected coal samples are also presented. Biomass generally has a lower energy density than coal, oils, and natural gas; it also has higher oxygen content. The higher volatiles and oxygen content of biomass translate into a higher reactivity compared to traditional fossil fuels. In terms of thermochemical conversions, this means that less severe process conditions (lower temperature and shorter residence time) are required for biomass-fueled systems. The chemical composition of biomass ash and the concentration of S, N, and Cl in the biomass can have detrimental impacts on long-term system operability and environmental performance.

Biomass can be used in thermochemical processes to produce heat and electricity, as with other renewable energy technologies. However, biomass resources are uniquely

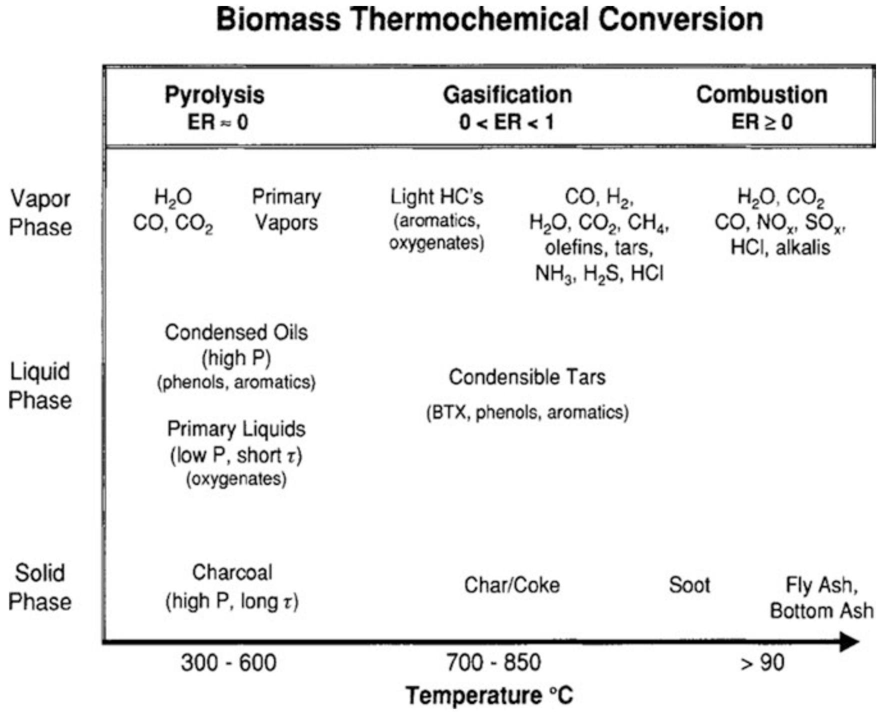


Fig. 33.29 General product composition as a function of process severity for biomass thermochemical processes

suited for conversion into liquid transportation fuels, chemicals, and materials. Thermochemical conversion of biomass into heat and power, transportation fuels, and chemical feedstocks is achieved by using one of three processes: pyrolysis, gasification, or combustion. All of these processes effectively use thermal energy to dehydrate, devolatilize, depolymerize, and oxidize, partially or completely, lignocellulosic materials to varying degrees. These thermochemical conversion processes are arranged in Fig. 33.29 in terms of process severity that can be considered a function of temperature and residence time. Increased temperature also correlates with increasing oxidation, especially for autothermal processes. Pressure does not necessarily affect process severity but can affect the product composition of thermochemical conversion.

Biomass pyrolysis is the thermal depolymerization of biomass at modest temperatures in the absence of added oxygen. The slate of products from biomass pyrolysis depends on the process temperature, pressure, and residence time of the liberated pyrolysis vapors [540–543]. Charcoal yields of up to 35% [544] can be achieved for slow pyrolysis at low temperature, high pressure, and long residence time. Flash pyrolysis is used to optimize the liquid products in an oil known as bio-crude or bio-oil [542]. High heating rates and short residence times enable rapid biomass pyrolysis while minimizing vapor cracking to optimize liquid product yields with up to 80% efficiency. Hydrothermal upgrading is a related process that produces a bio-oil product that converts wet biomass at high pressures into oxygenated hydrocarbons. Biomass liquefaction converts wet biomass feedstocks into liquid hydrocarbons at low temperatures in a high-pressure hydrogen environment.

Biomass gasification consists of several elementary steps that occur during the partial oxidation of biomass by a gasifying agent, generally air, oxygen, or steam. Initial heating of the biomass leads to evaporation of water. A further increase in temperature initiates biomass pyrolysis followed by the partial oxidation of pyrolysis vapors. The char remaining after a biomass particle is devolatilized is also gasified. The biomass gasification product gas is a low- to medium-energy content gas (depending on the gasifying agent), known as synthesis gas or syngas, that consists mainly of CO, H_2, CO_2, H_2O, N_2, CH_4, and other hydrocarbons. Minor components of the product gas include tars, sulfur and nitrogen oxides, alkali metals, and particulates. These minor components of the product gas potentially threaten the successful application of downstream syngas utilization processes. Gas composition and quality are dependent on a wide range of factors including feedstock composition, type of gasification reactor, gasification agents, stoichiometry, temperature, pressure, and the presence or lack of catalysts.

Biomass combustion systems range from very small-scale home heating applications (wood and pellet stoves) to small-scale commercial heating applications (furnaces and boilers) to modest-size biomass power plants (20–50 MW_e) to large-scale recovery boilers (400 MW_{th}) used in the pulp and paper industry. Direct biomass combustion systems can be used to generate process heat and electricity through traditional Rankine steam cycles at relatively

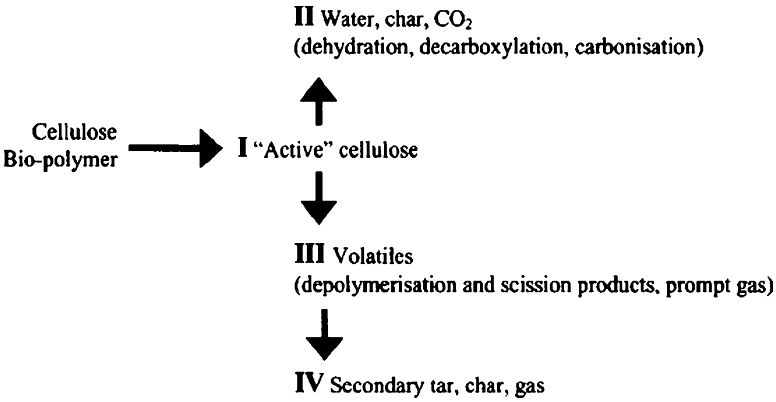


Fig. 33.30 Broido and Shafizadeh model for pyrolysis of cellulose

low (17–25%) thermal conversion efficiencies. Biomass can also be co-combusted with coal to take advantage of the high efficiency of large-scale pulverized-coal-fired power plants. Biomass combustion processes are generally feasible only if feed-stocks contain less than ~50% moisture.

Each of these thermochemical conversion processes has unique technical barriers and challenges that have been overcome with varying degrees of technical and economic success. The applications of these thermochemical conversion technologies are also in various stages of commercial application. For example, biomass combustion and biomass pyrolysis for production of a specific food additive are commercially practiced. Other technologies such as biomass gasification and use in integrated combined cycles for power production are in the demonstration stage. The technical details and challenges facing present and future application of these thermochemical biomass conversion processes are discussed in the following sections.

Pyrolysis

Pyrolysis as defined is a process of thermal decomposition occurring in the absence of oxygen. Pyrolysis of biomass is a complicated multistage reaction for which many pathways and mechanisms have been proposed [545–552]. The best known is the model developed by Broido and Shafizadeh [553, 554] for pyrolysis of cellulose that can be also applied, at least qualitatively, to whole biomass (Fig. 33.30).

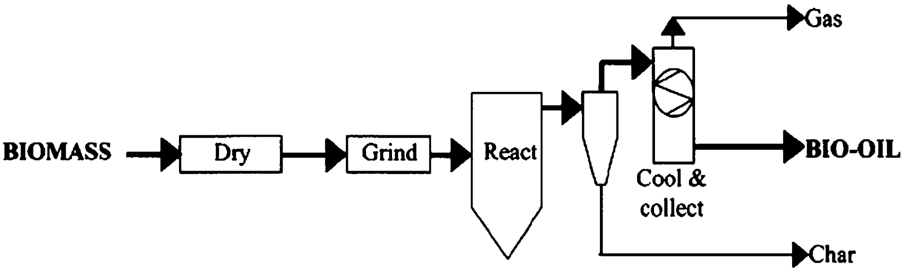
As shown in this model, pyrolysis of cellulose always results in solid, liquid, and gaseous products. However, the proportions of the product yields can change depending on the process conditions. The knowledge of thermodynamics and kinetics of the reaction pathways allows us to adjust the conditions to maximize the yield of desired products. Dehydration of cellulose is exothermic, whereas depolymerization and secondary vapor cracking are endothermic and have higher activation energy than dehydration. Therefore, lower process temperature and longer vapor residence times will favor the production of charcoal. High temperature and longer residence time will increase the biomass conversion to gas, and moderate temperature and short vapor residence time, necessary to minimize secondary cracking, are optimum for producing liquids. Table 33.6 provides data on the product distribution obtained by different modes of pyrolysis process.

Table 33.6 Modes of pyrolysis and typical products of wood pyrolysis

	Liquid (%)	Char (%)	Gas (%)
Fast pyrolysis Moderate temperature and short residence time, particularly vapor	75	12	13
Carbonization Low temperature and long residence time	30	35	35
Gasification High temperature and long residence times	5	10	85

Pyrolysis is always the first step in biomass thermochemical processes in which it is followed by total or partial oxidation of the primary products. Pyrolysis of biomass has been practiced for many thousand years to produce charcoal. In 1996, the world annual charcoal production was estimated at 100 million tons [544, 555]. Charcoal is a very important fuel in developing countries and is still a highly desired reductant in the metallurgical industry because of its low sulfur and mercury content. Large amounts of charcoal are also used to produce activated carbons extensively used for cleaning water and air. Charcoal is produced by a slow pyrolysis process occurring at temperature 350–450°C with low biomass heating rates (1–10°C/min). In most processes, part of the raw material is oxidized, using a limited air access, to provide heat for the process. Although charcoal formation reactions are slightly exothermic, heat is needed for drying biomass, which always contains moisture. Existing charcoal plants range from earth mound kilns to retorts operating in batch mode (cycles from 20 days to 24 h)

Fig. 33.31 Model for bio-oil production from biomass



to continuous shaft reactors having throughput up to 30 t/day. Depending on the production method, the yields of charcoal can vary from 10 to 35% based on dry wood. In the more developed processes (Degussa, Lambiotte, Lurgi) byproducts such as acetic acid, methanol, and other chemicals are also recovered. The technologies for producing charcoal have been reviewed in many textbooks [556] and a comprehensive review of the fundamentals of slow pyrolysis processes was recently published [544].

Fast Pyrolysis

Fast pyrolysis is one of the most recent renewable energy processes developed to maximize the production of the liquid product, bio-oil. It has achieved commercial status for production of chemicals and is very close to commercialization for the production of liquid fuels. Although related to the traditional slow pyrolysis processes for making charcoal, fast pyrolysis is carried out at a higher temperature (450–550°C) and with much higher biomass heating rates (100–1,000°C/s). As a result, biomass decomposes to generate mostly vapors, gases, and aerosols, and less charcoal. After cooling and condensation of the volatiles, a dark brown liquid is formed that has a heating value about half that of conventional fuel oil.

The essential features of a fast pyrolysis process are:

- Very high heating and heat transfer rates, usually requiring a finely ground biomass feed
- Carefully controlled pyrolysis reaction temperature of around 500°C in the vapor phase, with short vapor residence times of typically less than 2 s
- Rapid cooling of the pyrolysis vapors to give the bio-oil product

The main product, bio-oil, is obtained in yields of up to 75% wt on dry feed basis, together with byproduct char and gas, which are used within the process, so there are no waste streams. Several reviews on fast pyrolysis have been published in recent years that cover, the production, utilization, storage, and upgrading of biomass pyrolysis oils [540, 541, 543, 557–564].

A wide range of reactors such as bubbling and circulating fluidized beds, ablative, entrainment, rotating cone, auger, and vacuum have been operated. Fluidized beds were the most often used because of their ease of operation and ready scaleup. A typical bubbling fluidized-bed configuration presented in Fig. 33.31 uses the byproduct gas and char to provide the process heat. It also includes the necessary steps of drying the feed to less than 10% water to minimize the water in the product liquid oil, and grinding the feed to around 2 mm to supply sufficiently small particles to ensure rapid reaction.

Table 33.7 Typical properties of wood derived bio-oil

Physical property	Typical value
Moisture content	15–30%
pH	2.5
Specific gravity	1.20
Elemental analysis	
C	56.4%
H	6.2%
O	37.3%
N	0.1%
Ash	0.1%
HHV as produced (depends on moisture)	16–19 MJ/kg
Viscosity (at 40 °C and 25% water)	40–100 cp
Solids (char)	1%
Vacuum distillation residue	Up to 50%

Bio-oil from rapid pyrolysis is usually a dark brown, free-flowing liquid having a distinctive smoky odor. It has significantly different physical and chemical properties compared to the liquid from slow pyrolysis processes, which is more like a tar. Bio-oils are multicomponent mixtures comprised of different size molecules derived primarily from depolymerization and fragmentation reactions of the three key biomass building blocks: cellulose, hemicellulose, and lignin. Therefore, the elemental composition of bio-oil resembles that of biomass rather than that of petroleum oils. Basic properties of bio-oils are shown in Table 33.7. More detail on fuel-related characteristics is provided in the literature [565].

The single most abundant bio-oil component is water. The other major groups of compounds identified are hydroxyaldehydes, hydroxyketones, sugars, carboxylic acids, and phenolics. Most of the phenolic compounds are present as oligomers having a molecular weight ranging from 900 to 2,500. The presence of oxygen in many bio-oil components is the primary reason for differences in properties and behavior from those of fossil hydrocarbon fuels. Because of its high oxygen, bio-oil has an energy density (heating value) that is less than 50% of that for conventional fuel oils and it is also immiscible with hydrocarbon fuels. An even more important consequence of the organic oxygen in bio-oil is its limited stability.

Liquid bio-oil can be easily transported and stored. Czernik and Bridgwater reviewed the research on use of bio-oils for heat and power generation, showing that it is possible and usually only requires minor modifications of existing equipment [544]. Bio-oil has been successfully used as boiler fuel and also showed promise for diesel engines and gas turbines [543, 559, 565–578]. Upgrading bio-oil to a quality for liquid transportation fuels still poses several technical challenges and is not currently economically attractive [543, 558, 577, 579–582].

Slow Pyrolysis

For many centuries, wood slow-pyrolysis liquids were a major source of chemicals such as methanol, acetic acid, turpentine, and tars. At present, most of these compounds can be produced at a lower cost from other feedstocks derived from natural gas, crude oil, or coal. Though more than 300 compounds have been identified in wood fast pyrolysis oil, their amounts are small. Isolation of specific single compounds is seldom practical or economical, because it usually requires complex separation techniques. Some chemicals produced from the whole bio-oil or by its fractionation are already commercial products. Liquid smoke food flavoring is one example. A few others such as pyrolytic lignin as phenol replacement in resins and bio-oil-based slow-release fertilizer have a chance for short-term commercialization, especially in the context of a biorefinery based on a fast pyrolysis process. Commercialization of specialty chemicals such as glycolaldehyde or levoglucosan [583] from bio-oil requires more work to develop reliable low-cost separation procedures.

Gasification

Biomass gasification can effectively convert a very heterogeneous material into a consistent gaseous fuel intermediate for heating, industrial process applications, electricity generation, and liquid fuels production. Biomass gasification is a complex thermochemical process that consists of a number of elementary chemical reactions, beginning with the thermal decomposition of a lignocellulosic fuel followed by partial oxidation of the fuel with a gasifying agent, usually air, oxygen, or steam [584]. Volatile matter released as the biomass fuel is heated partially oxidizes to yield the combustion products H_2O and CO_2, plus heat to continue the endothermic gasification process. Water vaporizes and biomass pyrolysis continues as the fuel is heated. Thermal decomposition and partial oxidation of the pyrolysis vapors occur at higher temperatures, and yield a product gas composed of CO, CO_2, H_2O, H_2, CH_4, other gaseous hydrocarbons (including oxygenated hydrocarbons from some processes), tars, char, volatile inorganic constituents, and ash. A generalized reaction describing biomass gasification is as follows.

$$\begin{aligned}&\text{biomass} + O_2(\text{or } H_2O) \rightarrow CO, CO_2, H_2O, H_2, CH_4\\&\quad + \text{other hydrocarbons} \rightarrow \text{tar} + \text{char} + \text{ash}\\&\quad \rightarrow HCN + NH_3 + HCl + H_2S\\&\qquad + \text{other sulfur gases}\end{aligned} \tag{33.1}$$

The actual amount of CO, CO_2, H_2O, H_2, tars, and hydrocarbons depends on the partial oxidation of the volatile products, as shown in (33.2).

$$\begin{aligned}&C_nH_m + (n/2 + m/4)O_2\\&\quad \rightarrow nCO + (m/2)H_2O\end{aligned} \tag{33.2}$$

The char yield in a gasification process can be optimized to maximize carbon conversion or the char can be combusted to provide heat for the process. Char is partially oxidized or gasified according to the following reactions.

$$C + \frac{1}{2}O_2 \rightarrow CO \tag{33.3}$$

$$C + H_2O \rightarrow CO + H_2 \tag{33.4}$$

$$C + CO_2 \rightarrow 2CO\ (\text{Boudouard reaction}) \tag{33.5}$$

The gasification product gas composition, particularly the H_2/CO ratio, can be further adjusted by reforming and shift chemistry. Additional hydrogen is formed when CO reacts with excess water vapor according to the water–gas shift reaction.

$$CO + H_2O \leftrightarrow CO_2 + H_2 \tag{33.6}$$

Reforming the light hydrocarbons and tars formed during biomass gasification also produces hydrogen. Steam reforming and so-called dry or CO_2 reforming occur according to the following reactions and are usually promoted by the use of catalysts.

$$C_nH_m + nH_2O \rightarrow nCO + (n + m/2)H_2 \tag{33.7}$$

$$C_nH_m + nCO_2 \rightarrow (2n)CO + (m/2)H_2. \tag{33.8}$$

The actual biomass gasification product gas composition depends heavily on the gasification process, the gasifying agent, and the feedstock composition [585, 586].

Gasifier Reactor Designs

Four primary types of biomass gasification systems have been developed. These include fixed-bed reactors, bubbling fluidized-bed reactors, circulating fluidized-bed reactors, and entrained-flow reactors. Fixed-bed gasifiers can be classified primarily as updraft and downdraft [587]. Updraft gasifiers represent the oldest and simplest gasifiers. The updraft gasifier is a counterflow system where fuel is introduced into the top of the gasifier and flows downward and the gasifying medium is introduced below the grate and flows upwards through the reactor. Feed particle size needs to be controlled to maintain a uniform bed.

Char burns in the combustion zone at the bottom of the reactor, forming CO_2 and steam, which flow upward through the bed. Residual ash is removed from the bottom of the reactor at the grate. Operational problems can be avoided for feeds with low-melting-point ash by carefully controlling the temperature in the combustion zone or with fuel blending or additives to alter the ash chemistry on the grate.

The exothermic combustion reactions supply the energy to drive gasification, pyrolysis, and drying. Above the combustion zone is the reduction zone where CO_2 and H_2O are partially reduced to CO and H_2 according to reactions (33.4) and (33.5). The next area of the reactor is the pyrolysis zone where these gases are hot enough to devolatilize the biomass to produce pyrolysis products and residual char. At the top of the reactor the product gases and pyrolytic vapors dry the wet biomass. Product gases exit the gasifier at temperatures (typically 80–100°C) higher than the dew point of the product gas. A wide range of condensable tars and oils is produced in the pyrolysis zone. These can condense in transfer lines at the output of the gasifier. For this reason, updraft gasifiers are usually operated in a close-coupled mode to a furnace or boiler to produce steam or hot water.

In downdraft gasification, the gasifying medium and biomass fuel both flow in the same direction as the solid bed. This design tends to minimize tar and oil production. The fuel and pyrolytic gases and vapors move co-currently downward through the bed. The pyrolysis products pass through hot char (about 15% of the original feed) into a combustion zone where they mix with air and are thermally cracked and partially oxidized. The oxidizer concentration diminishes downstream of the combustion zone forming a reduction zone. The remaining hot char in the reduction zone reduces some of the combustion products, CO_2 and H_2O, to CO and H_2. The product gas exits the gasifier at fairly high temperature, around 700°C. Downdraft gasification systems typically have low overall thermal efficiency because of heat losses from the product gases, however, tar conversion is high, greater than 99%, and is a function of temperature, combustion efficiency, and channeling.

Downdraft and updraft gasifiers have the same general constraints on feed properties. The feed needs to have a fairly uniform particle size with few fines to maintain bed physical properties and minimize channeling. A low-ash feed with a high fusion temperature is also preferred to prevent slagging. Feed moisture content also needs to be less than about 20% to maintain the high temperatures required for tar cracking. A variation on the downdraft gasifier is the crossflow gasifier in which air is introduced tangentially at the bottom of the reactor. The principle of operation of the crossflow gasifier is the same as the downdraft gasifier.

Alternative gasifier designs are based on more complex reactor systems to maximize heat transfer and gas–solid interactions by introducing biomass into a moving bed of high-temperature solids [588]. In a gas–solid fluidized bed, a stream of gas passes upwards through a bed of selected free-flowing granular materials such as silica, olivine, limestone, dolomite, or alumina. A gas distribution manifold or series of sparge tubes [589–591] is used to maintain gas velocities high enough to freely circulate the widely separated solid particles so that the bed resembles a boiling liquid and has the physical properties of a fluid.

Above the bed in a fluidized-bed gasifier, the cross-sectional area of the reactor is increased to produce a disengaging zone where the superficial gas velocity is below fluidization velocity [592]. This allows the entrained sand particles to fall back down and maintain the bed inventory over time. This larger cross-sectional area zone, or freeboard, can be extended to obtain the total desired gas-phase residence time for complete devolatilization. Eroded bed material or fine char and ash particles that escape the reactor can be collected in a cyclone and either returned to the bed or removed from the system.

In a fluidized-bed gasifier, the oxidizer (air, oxygen, or steam) typically doubles as the fluidizing medium. Biomass can be introduced either on top of the bed or through an auger into the bed. In-bed feeding improves the conversion of fine particles that would otherwise be entrained in the fluidizing gas if they were introduced on top of the bed. As biomass is introduced into the bed, the organic pyrolysis vapors are released and partially oxidized in the bed. The exothermic combustion provides the heat to maintain the bed temperature around 800°C and to volatilize additional biomass. Bed temperature is governed by the desire to obtain complete devolatilization (95–99% carbon conversion) vs. the need to maintain the bed temperature below the ash-fusion temperature of the biomass ash.

As the gas flow in a bubbling fluidized bed is increased, a turbulent fluidized bed is formed. The gas bubbles become larger, larger voids are formed in the bed, and more solids are entrained in the exit gas flow. A circulating fluidized-bed system results if these solids are collected, separated from the gas, and returned to the bed through a continuous solids circulation loop. A circulating fluidized bed differs from a bubbling fluidized bed in that there is no distinct separation between the dense solids zone and the dilute solids zone. Circulating fluidized-bed densities are about 560 kg/m^3 compared to a bubbling-bed density of 720 kg/m [593]. To achieve the lower bed density, gas rates are increased from the 1.5–3.7 m^3/s of bubbling beds to about 9.1 m^3/s. The residence time of the solids in a circulating fluidized bed is determined by the solids circulation rate, the solids attrition rate, and the collection efficiency of the solids separation device.

Biomass-Gasification Product Gas

Various gasification technologies have been investigated for converting biomass into a gaseous fuel. The energy content of the gasification product gas ranges from 5 to 15 MJ/Nm^3 and is considered a low-to-medium energy-content gas compared to natural gas (35 MJ/Nm^3) [594, 595]. The relative amount of CO, CO_2, H_2O, H_2, and hydrocarbons depends on the stoichiometry of the gasification process and the selected gasification medium. The air/fuel ratio in a gasification process generally ranges from 0.2 to 0.35 and if steam is the gasifying agent, the steam/biomass ratio is around 1. If air is used as the gasifying agent, then roughly half of the product gas is N_2 [596].

Air-blown, or directly heated gasifiers, use the exothermic reaction between oxygen and organics to provide the heat necessary to devolatilize biomass and to convert residual carbon-rich chars. In these directly heated gasifiers, the heat to drive the process is generated internally within the gasifier. Thus, when air is used, the resulting product gas is diluted with nitrogen and typically has a dry-basis calorific value of about 5–6 MJ/Nm^3. The dry-basis calorific value of the product gas can be increased to 13–14 MJ/Nm^3 by using oxygen instead of air. Oxygen production is expensive, however, and its use has only been proposed for direct-heating gasification applications involving the production of synthesis gas where nitrogen is not permitted in downstream synthesis conversion operations. Oxygen typically costs \$40–\$60 per Mg and typically is used at the rate of 0.25–0.3 Mg/Mg of biomass, a cost equivalent to \$10–\$20 per Mg of biomass [597]. Indirectly heated gasifiers accomplish biomass heating and gasification through heat transfer from a hot solid or a heat transfer surface. Because air is not introduced into the gasifier, little nitrogen diluent is present and a medium calorific gas is produced; dry basis values of 18–20 MJ/Nm^3 are typical.

Gas phase impurities in syngas include NH_3, HCN, other nitrogen-containing gases, H_2S, other sulfur gases, HC1, alkali metals, organic hydrocarbons (tar), and particulates. The concentration of these non-syngas components strongly depends on the feedstock composition. Gasification of biomass containing high nitrogen and sulfur contents yields high levels of NH_3 and H_2S in the syngas stream. HCl concentration in biomass-derived syngas directly correlates with the chlorine content of the feedstock. Alkali metal, mostly potassium, in syngas is related to the alkali content in the biomass ash. Ash particles entrained in syngas affect the alkali metal content in syngas. The concentration of alkali vapors or aerosols in syngas depends on the ash chemistry of the selected biomass feedstock and the temperature of the gasification process.

The organic impurities in syngas range from low-molecular-weight hydrocarbons to high-molecular-weight polynuclear aromatic hydrocarbons. The lower-molecular-weight hydrocarbons can be used as fuel in gas turbines or engines, but are undesirable products in fuel cell applications and methanol synthesis. The higher-molecular-weight hydrocarbons are collectively known as “tar.” Tar yields in biomass-derived syngas can range from 0.1% (downdraft) to 20% (updraft) or greater (pyrolysis) in the product gases.

One issue associated with biomass gasifier tars is how they are defined. More often than not, tar is given an operational definition by those conducting biomass gasification research. An excellent report by Milne et al. [598] describes in detail the operational definitions of biomass gasification tars as published in the literature and provides a comprehensive survey of tar formation and conversion. For the most part, “tars” are considered to be the condensable fraction of the organic gasification products and are largely aromatic hydrocarbons, including benzene. The diversity in the operational definitions of “tars” usually comes from the variable product gas compositions required for a particular end-use application and how the “tars” are collected and analyzed. Tar sampling protocols are being developed [599–601] to help standardize the way tars are collected; however, these methods are not yet widely established.

Syngas Cleanup and Conditioning

Gas conditioning is a general term for removing the unwanted impurities from biomass-gasification product gas and generally involves an integrated, multistep approach that depends on the end use of the product gas [602–605]. For close-coupled gasifier-combustor systems, no gas cleanup is required. For gas-turbine power generation, the gas has to be free of particulates, tars, sulfur, and chlorine compounds, and of alkali metals to ensure the integrity of the turbine hot section [606–609]. Particulate removal to protect the turbine blades from erosion requires filtration technology, and the removal of tars ensures an even and less luminous combustion process (to avoid radiative heat transfer problems at the turbine). Alkali metal removal avoids deposition and corrosion of the turbine blade materials. Potassium and sodium levels must be reduced to less than 1 ppm to ensure long turbine blade and hot-section lifetimes. The tolerance to alkali metals is a function of the temperature of operation of the turbine inlet section. At very high temperatures (greater than 1,350°C) the level has to be less than about 25 ppb. The deposition and corrosion problems can be also addressed with turbine blade coatings, and it is likely that both coating and strengthening of components will be undertaken as well, as biomass power technology develops. For internal combustion use, it is necessary to cool the gas to ensure that a sufficient charge of energy can be put into each cylinder. Particular attention has to be given to both tar and particulate contents, to ensure that valves and cylinders are protected. Fuel-cell applications would require the gas to be mainly hydrogen without any significant sulfur or chloride contamination, to protect the electrodes. For synthesis operations such as methanol and hydrogen production, removal of particulates and contaminants such as H_2S is required to prevent poisoning of downstream catalysts.

Tar removal, conversion, or destruction is seen as one of the greatest technical challenges for the successful development of commercial advanced-gasification technologies [598]. Tars can condense in exit pipes and on particulate filters, leading to blockages and clogged filters. Tars also have varied impacts on other downstream processes. Tars can clog fuel lines and injectors in internal combustion engines. Luminous combustion and erosion from soot formation can occur in pressurized combined-cycle systems where the product gases are burned in a gas turbine. The product gas from an atmospheric-pressure gasification process needs to be compressed before it is burned in a gas turbine and tars can condense in the compressor or in the transfer lines as the product gas cools.

If the end use of the gas requires cooling to near ambient temperatures, it is possible to use wet scrubbing and filtration or other physical removal methods, to remove tars. Wet scrubbing is an effective gas-conditioning process that condenses the tars out of the product gas. This technology is available and can be optimized for tar removal. A disadvantage of wet scrubbing for product gas conditioning is the formation and accumulation of wastewater. This technique does not eliminate tars but merely transfers the problem from the gas phase to the condensed phase. Wastewater minimization and treatment are important considerations when wet scrubbing is used for tar removal. Also, when tar is removed from the product gas stream, its fuel value is lost and the overall efficiency of the integrated gasification process is reduced.

If the end use requires that the product gas remain at high temperature, at or slightly below the gasifier exit temperature, then some method of hot gas cleaning will be needed for tar elimination. Wet scrubbing is still an option, however, there will be a severe thermodynamic penalty from cooling and reheating the conditioned product gas, reducing the overall efficiency of the process. Hot gas conditioning eliminates tars by converting them into desired product gas components, thus retaining their chemical energy in the product gas and avoiding treatment of an additional waste stream. Thermal cracking is a hot gas conditioning option but it requires temperatures higher than typical gasifier exit temperatures (greater than 1,100°C) to achieve high conversion efficiencies. Increased temperatures for thermally cracking tars can come from adding oxygen to the process and consuming some of the product gas to provide additional heat. Thermal destruction of tars may also produce soot that is an unwanted impurity in the product gas stream.

An attractive hot gas conditioning method for tar destruction is catalytic steam reforming [610–615]. This technique offers several advantages: (1) catalyst reactor temperatures can be thermally integrated with the gasifier exit temperature, (2) the composition of the product gas can be catalytically adjusted, and (3) steam can be added to the catalyst reactor to ensure complete reforming of tars. Catalytic tar destruction has been studied for several decades [616–618] and a number of reviews have been written on biomass gasification hot gas cleanup [598, 615, 619]. Numerous catalysts have been tested for tar destruction activity at a broad range of scales. Novel catalyst formulations have been sought to increase the activity and lifetime of tar reforming catalysts. Different approaches for integrating catalytic tar destruction into biomass gasification systems have also been investigated.

Calcined dolomites are the most widely used nonmetallic catalysts for tar conversion in biomass gasification processes [620–625]. They are relatively inexpensive and are considered disposable. However, they are not very robust and quickly undergo attrition in fluidized-bed reactors. Consequently, dolomites find most use in fixed-bed catalytic reactors. Tar conversion efficiency is high when calcined dolomites are operated at high temperatures (900°C) with steam. Olivine, another naturally occurring mineral, has also demonstrated tar conversion activity similar to that of calcined dolomite. Olivine is a much more robust material than calcined dolomite and has been applied as a primary catalyst to reduce the output tar levels from fluidized-bed biomass gasifiers.

Commercial Ni catalysts are designed for use in fixed-bed applications and are not robust enough for fluidized-bed applications, and therefore are not useful as primary in-bed catalysts. These catalysts, however, have been extensively used for biomass-gasification tar conversion as secondary catalysts in separate fixed-bed reactors operated independently to optimize performance [626–628]. They have high tar-destruction activity with the added advantages of completely reforming methane and of promoting water–gas shift activity that allows the H_2:CO ratio of the product gas to be adjusted. Some studies have also shown that nickel catalyzes the reverse ammonia reaction, thus reducing the amount of NH_3 in gasification product gas.

A limitation of nickel catalyst use for hot gas conditioning of biomass gasification product gases is rapid deactivation, limiting catalyst lifetimes. Ni catalyst deactivation is caused by several factors. Sulfur, chlorine, and alkali metals, that may be present in gasification product gases, act as catalyst poisons. Coke formation on the catalyst surface can be substantial when tar levels in product gases are high. Coke can be removed by regenerating the catalyst, however, repeated high-temperature processing of nickel catalysts can lead to sintering, phase transformations, and volatilization of the nickel. Continued disposal of spent toxic Ni catalysts is not economical and poses an environmental hazard.

Using fixed dolomite guard beds to lower the input tar concentration can extend Ni catalyst lifetimes. Adding various promoters and support modifiers has been demonstrated to improve catalyst lifetime by reducing catalyst deactivation by coke formation, sulfur and chlorine poisoning, and sintering. Several novel, Ni-based catalyst formulations have been developed that show excellent tar reforming activity, improved mechanical properties for fluidized-bed applications, and enhanced lifetimes.

Hot gas conditioning using current or future commercially available catalysts offers the best solution for mitigating biomass gasification tars. Tars are eliminated, methane can be reformed if desired, and the H_2:CO ratio can be adjusted in a single step. The best currently available tar reforming process consists of a calcined-dolomite guard bed followed by a fixed-bed Ni-catalyst reforming reactor operating at about 800°C. Selection of the ideal Ni catalyst is somewhat premature. Commercially available steam reforming catalysts have been demonstrated, but several of the novel research catalysts appear to have the potential of longer lifetimes. This dual-bed hot gas-conditioning concept has been demonstrated and can be used to condition the product gas from any developing gasification process. A proprietary Ni-monolith catalyst has also shown considerable promise for biomass gasification tar destruction [614].

Biomass-Derived Syngas Utilization

Biomass gasification is a promising thermochemical conversion technique for producing electricity and liquid fuels for power generation because of the high efficiencies projected for integrated-gasification combined-cycle systems. Solid or gasified biomass can be burned as boiler fuel to generate electricity via the steam cycle. Biomass syngas can, however, also be used in diesel, internal combustion, or Stirling engines, in high-efficiency gas turbines, or in high-efficiency fuel cells, all with higher efficiency potential. All of these are suitable for even higher efficiency use with combined heat and power operation. Gasification also enables production of a wide range of fuels and chemicals, such as syngas, hydrogen, and liquid fuels or methanol, ethanol, and Fischer–Tropsch liquids using commercial processes developed in the petrochemical industry for the conversion of methane.

Integrated Gasification Combined Cycle

The integrated gasification combined cycle (IGCC) system has two main components, a high-efficiency gas turbine and a heat recovery steam generation system. This system uses heat from the turbine exhaust to raise steam and produce electricity in a traditional steam turbine section. Integrated biomass gasification combined cycle systems for electricity production have a number of potential advantages [606, 607, 629]. Projected process efficiencies are much higher than the direct combustion systems in commercial use today. In a co-generation application, overall system efficiencies can approach 85%, thus maximizing the conversion of biomass, a relatively low-energy-density fuel, to heat and power [630–633]. These process efficiencies are comparable to high efficiency large-scale pulverized coal systems, but can be achieved at a smaller scale of operation.

The increased efficiency in IGCC systems translates into environmental benefits because emissions per unit energy produced are lower. This includes CO, NO_x, SO_x, CO_2, and particulates. NO_x and SO_x emissions are inherently lower in biomass IGCC systems because biomass fuels tend to have low N and S contents to start with, and gas cleanup and conditioning in biomass gasification systems removes these impurities before the syngas is combusted in the gas turbine. The same is true for particulates. CO_2 emissions are effectively zero, because biomass is a renewable fuel.

Fuel Cell Applications

Fuel cell systems have the potential to substantially reduce air and water emissions associated with electricity production. In all fuel cell systems, hydrogen is consumed at the anode and water is produced at the cathode. The higher system efficiencies for fuel cells translate into enhanced fuel utilization and therefore reduced CO_2 emissions compared to lower efficiency systems. Fuel cell power plants will be capable of exceeding stringent present and future environmental regulations for particulates, NO_x, and SO_x emissions.

Without considering batteries and other chemical storage devices, there are effectively six types of primary or direct fuel cell technologies currently being developed: alkaline fuel cells (AFC), polymer electrolyte fuel cells (PEFCs), a specialized PEFC using methanol as the fuel called the direct methanol fuel cell, phosphoric acid fuel cells (PAFC), molten carbonate fuel cells (MCFCs), and solid oxide fuel cells (SOFC) [634–638]. These technologies are at various stages of commercialization and individual fuel cell systems have different applications [637]. The high cost of fuel cell systems ($1,000–$20,000/kW) is one of the more significant barriers to commercialization [635].

Integrated biomass gasification fuel cell systems have potential use for distributed and centralized power production.

Greater fuel flexibility and greater resistance to impurities makes developing high-temperature MCFC and SOFC systems the most likely technologies for future biomass applications. These systems are currently in various stages of development and demonstration, although MCFC systems are nearer to commercialization [639, 640]. Unlike PAFC systems, MCFCs can tolerate the high concentrations of CO that are produced in biomass gasification [634, 641, 642]. This eliminates the need for water–gas shifting and selective CO oxidation that would otherwise increase system costs and reduce overall efficiency. The higher operating temperatures of MCFC systems also provide an opportunity for co-generation or waste heat utilization that could improve overall system efficiency to about 85% (lower heating value) [643].

Syngas to Liquid Fuels

In its simplest from, syngas is composed of two diatomic molecules, CO and H_2 that provide the building blocks upon which an entire field of fuel science and technology is based [644–649]. Sabatier and Sanderens, who produced methane by passing CO and H_2 over Ni, Fe, and Co catalysts, discovered the synthesis of hydrocarbons from CO hydrogenation in 1902. At about the same time, the first commercial hydrogen from syngas produced from steam methane reforming was commercialized. Haber and Bosch discovered the synthesis of ammonia from H_2 and N_2 in 1910 and the first industrial ammonia synthesis plant was commissioned in 1913. The production of liquid hydrocarbons and oxygenates from syngas conversion over iron catalysts was discovered in 1923 by Fischer and Tropsch. Variations on this synthesis pathway were soon to follow for the selective production of methanol, mixed alcohols, and isosynthesis products. Another outgrowth of Fischer–Tropsch Synthesis (FTS) was the hydroformylation of olefins discovered in 1938. Many of the syngas conversion processes were developed in Germany during the First and Second World War at a time when natural resources were becoming scarce and alternative routes for hydrogen production, ammonia synthesis, and transportation fuels were a necessity.

Syngas composition, most importantly the H_2/CO ratio, varies as a function of production technology and feedstock. Steam methane reforming yields H_2/CO ratios of three to one whereas coal and biomass gasification yields ratios closer to unity or lower. Conversely, the required properties of the syngas are a function of the synthesis process. Fewer moles of product almost always occur when H_2 and CO are converted to fuels and chemicals. Consequently, syngas conversion processes are more thermodynamically favorable at higher H_2 and CO partial pressures. The optimum pressures depend on the specific synthesis process.

Catalytic syngas conversion processes are exothermic reactions generating large excesses of heat. This highlights the specific need for removing this heat of reaction to carefully control reaction temperatures to maintain optimized process conditions. Maximizing product yields, minimizing side or competing reactions, and maintaining catalyst integrity dictate optimum synthesis reaction temperatures.

Appropriate catalysts are necessary for all fuel and chemical synthesis. The basic concept of a catalytic reaction is that reactants adsorb onto the catalyst surface and rearrange and combine into products that desorb from the surface. One of the fundamental functional differences between various syngas synthesis catalysts is whether the adsorbed CO molecule dissociates on the catalyst surface. For FTS and higher alcohol synthesis, CO dissociation is a necessary reaction condition. For methanol synthesis the CO bond remains intact. Hydrogen has two roles in catalytic syngas synthesis reactions. In addition to being a reactant needed for CO hydrogenation, it is usually used to reduce the metalized synthesis catalysts and activate the metal surface.

Since the genesis of syngas conversion to fuels and chemicals, a tremendous amount of research and development has been devoted to optimizing product yields and process efficiencies. This includes the discovery of catalysts with optimized formulations containing the most active metals in combination with appropriate additives to improve activity and selectivity in a given process. Mechanistic studies have been conducted to interpret the fundamentals of specific conversion processes and measure the kinetic rates of key chemical reactions. Reactor design and engineering is another active research and development area of syngas conversion technology. Temperature control and stability in conversion reactors is a critical process parameter because of the large excess heat of reaction. To optimize commercial synthesis processes, detailed process engineering and integration are used with respect to heat integration and to syngas recycling to improve conversion efficiencies. Given the rich history of syngas conversion and the extensive research and development efforts devoted to this field of study, it is not surprising that a vast amount of literature is available that tracks the scientific and technological advances in syngas chemistry (see the Fischer–Tropsh Archive at www.fischer-tropsch.org). A summary of various catalytic syngas processes, including types of catalyst and reaction conditions, is shown in the “star” diagram in Fig. 33.32 [649]. This is by no means a comprehensive list of potential products, but identifies the available processes for utilizing syngas for producing fuels and chemicals.

Combustion

Biomass combustion is the most common and historically oldest method of extracting energy from biomass (other than food) either directly, in the form of heat and light from fire, or indirectly through use of this heat to produce steam that

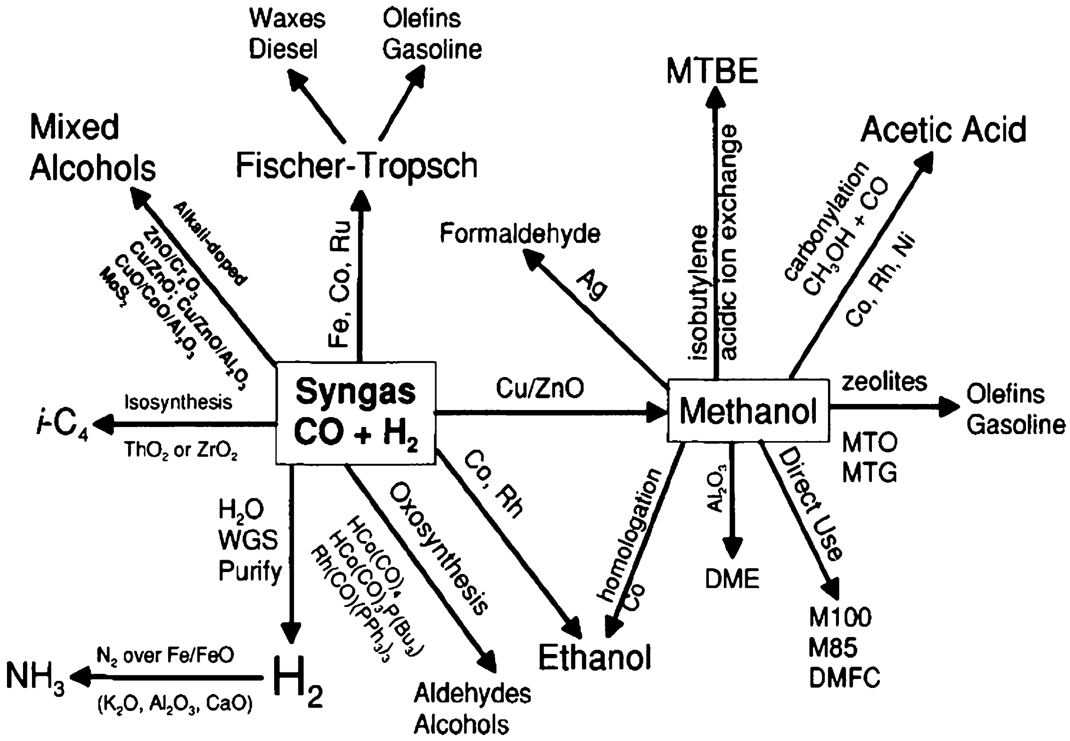


Fig. 33.32 Star diagram showing various high-pressure catalytic processes for converting syngas to fuels and chemicals

drives electricity-generating turbines. Biomass is the third largest global source of energy, accounting for 14% of annual global energy consumption [650]. The scale of biomass combustion systems ranges from residential heating and cooking applications to small-scale distributed energy production to large-scale centralized power production. Biomass use is very prominent in developing countries, particularly for distributed heat and power and for domestic heat, cooking, and light. Large-scale biomass power plants and industrial biomass utilization for heat and power supplant the residential and commercial biomass heating applications in developed countries.

The impetus for displacing fossil fuels with biomass in heat and power combustion systems is typically to reduce emissions of CO_2, and of NO_x, SO_x, and other air toxics, or to improve utilization of biomass residues and wastes. The threat of increased global warming has subjected the use of fossil fuels to increasing scrutiny in terms of greenhouse gas and pollutant emissions. Renewable and sustainable energy resources such as biomass can be combusted instead of coal to produce heat and power and help reduce the accumulation of greenhouse gases as carbon dioxide is consumed during plant growth. Also, on average, the sulfur content of biomass is lower than coal, so increasing the use of biomass power can reduce sulfur emissions from power production. The agricultural and wood products industries generate large quantities of biomass residues that could provide fuel for electricity production. Increasing the use of these waste biomass fuels could alleviate the burdens of waste disposal in the agricultural and wood products industries.

A number of biomass combustion technologies exist, including grate-fired and fluidized-bed processes [651]. The simplest biomass combustion systems are wood stoves and furnaces. These devices generate hot air that is used for heating residential and commercial buildings. Biomass boiler systems are more fuel flexible than the smaller furnaces and can be scaled to larger size and used to generate hot water or steam for heat and power applications. Larger-scale biomass boilers include pile burners in which biomass accumulates in a furnace and combustion air is fed from above and below the pile. This relatively simple design is fuel flexible but has low efficiency and poor combustion control, leading to relatively high emissions.

In grate-fired boilers, the biomass is fed in a thin layer, so it is evenly distributed over a sloped stationary, traveling, or vibrating grate. Improved control over the combustion process can be achieved with better carbon conversion. Fluidized-bed combustors are more complex systems, but offer much better control of combustion temperature, improved carbon conversion, and fuel flexibility. Using a boiler to produce both heat and electricity (co-generation) can improve the overall system efficiency to as much as 85%. Boiler efficiencies are affected by fuel moisture content, air–fuel ratio, excess air, combustion temperature, and biomass ash content.

Direct combustion to raise steam is used in all of the existing biomass generation plants in the United States today. Biopower is a commercially proven electricity-generating option in the United States and the following statistics for the use of biomass for electricity production can be found in the 2004 Energy Information Agency publications (www.eia.doe.gov). Renewable energy consumption in the United States in 2003 accounted for 6.2% of the nation's 98-quad total energy supply. The 2.9 quads of energy from biomass accounted for 47% of the renewable energy consumption. Biomass surpassed hydropower as the single largest renewable energy resource for the last several years. Of the total, nearly 90% of the biomass energy was consumed in the pulp and paper and forest products industry, three quarters of which was for process heat. Sixty percent of the electricity from biomass is produced from wood or wood wastes with the remaining 40% produced from landfill gas, municipal solid waste, and other biomass. Electricity production from biomass is being used and is expected to continue to be used as base load power in the existing electrical distribution system.

Today's biopower capacity is based on mature, direct-combustion boiler/steam turbine technology. The average size of biopower plants is 20 MW (the largest approach 100 MW) and the average efficiency from steam-turbine generators is 17–25%. The small plant sizes lead to higher capital cost per kilowatt-hour of power produced and the low electrical conversion efficiencies increase sensitivity to fluctuations in feedstock price [652].

All biomass energy systems suffer from the economic barrier associated with the energy cost of producing, transporting, and preparing the biomass feedstock. Significant progress has been made in this area, but to be truly economically competitive, new feedstocks and methods for their harvest and preparation must be developed. Harvesting, preparation, transportation, and feeding of a variety of biomass feedstocks suitable for power production must be demonstrated, and new methods developed for reducing costs and energy requirements must be verified. This will reduce the delivered cost of feedstock to the energy facility to a level more competitive with fossil fuels.

Technical barriers for biopower applications arise from the variability of biomass fuel compositions. The trace elements in biomass—N, Cl, K, Na, and ash—contribute to operational problems and pollutant emissions in biomass combustion systems [653–656]. The nitrogen content of biomass can vary considerably. Wood and wood wastes tend to have low nitrogen contents because hardwood and softwood trees are relatively slow growing and unfertilized. Agricultural residues and grasses tend to be faster growing and are fertilized and harvested on a yearly basis. Some agricultural residues such as alfalfa and soybeans have value as an animal feed because of their high protein (nitrogen) content. Variability in fuel-bound nitrogen (protein) in biomass leads to varying NO_x emissions in biomass combustion systems.

The chlorine content of biomass is generally a function of the soil conditions it is grown in and the amount of fertilizer applied during growth. Consequently, agricultural residues such as various stovers and straws can have high Cl levels. Subsequent HCl emissions during combustion can catalyze high- and low-temperature corrosion.

The major technical barrier associated with biomass combustion systems is the formation of tenacious deposits on heat transfer surfaces caused by the unique high-temperature chemistry of biomass ash. Many types of biomass used in combustion systems contain alkali metal species: sodium, potassium, and calcium. The ash content of woody biomass is quite low and not a problem. The ash content of agricultural residues such as wheat straw, rice straw, corn stover, and alfalfa stems can be quite high, on the order of 5% or up to 20% for rice straw and rice husks. The presence of alkali metals in conjunction with the high silica content of some biomass ashes can lead to molten ash at combustion temperatures [657–662]. In some cases, the K, Si, and Al contents are such that very low melting-point eutectic mixtures can form. The materials can be fluid at combustion temperatures, but form glasslike deposits on colder downstream surfaces such as heat exchanger tubes.

Ash deposition in biomass combustion systems has been the focus of numerous research efforts [653, 663]. The basic mechanism for deposit formation in biomass combustion systems starts with the vaporization of alkali metals, usually chlorides, in the combustor. Fly ash particles, which are predominantly silica, impact and stick to boiler tube surfaces. As the flue cools the alkali metal vapors and aerosols quench on the tube surfaces. When the ash chemistry approaches equilibrium on the surface and the deposit becomes molten, the likelihood increases that additional fly ash particles will stick, and deposits grow rapidly. Ash deposits can also accelerate the corrosion or erosion of the heat transfer surfaces. This greatly increases the maintenance requirements of the power plant often causing unscheduled plant interruptions and shutdown.

Co-firing

One solution to increasing biopower is to build dedicated biomass power plants. An alternative, lower capital cost option for increasing the use of biomass to produce electricity is to co-fire biomass and coal in existing coal-fired power plants. Coal-fired power plants are used to produce the majority of the electricity in the United States. If biomass were co-fired at low percentages in even a small number of coal-fired power plants, the use of biomass for power production would dramatically increase. Co-firing biomass and coal increases the use of sustainable fuels without the need for large capital investments while taking advantage of the high efficiencies obtainable in existing coal-fired power plants. Fuel diversity is another advantage of bio-mass/coal co-firing. Co-firing reduces the need for a constant supply of biomass that would be required in a dedicated biomass power plant. Co-firing biomass and coal is also a viable way to manage the increasing emissions of greenhouse gases and other pollutants from power-generating facilities.

Biomass and coal have fundamentally different fuel properties that can lead to benefits or deterrents to co-firing. For instance, biomass is a more volatile fuel than coal and has higher oxygen content. Coal, on the other hand, has more fixed carbon than biomass. Wood fuels tend to contain very little ash (on the order of 1% ash or less) and consequently increasing the ratio of wood in biomass/coal blends can reduce the amount of ash that needs to be disposed. A negative aspect of biomass is that it can contain more chlorine than coal. This is particularly true for some grasses, straws, and other agricultural residues.

Better environmental performance at a modest cost is one of the drivers for biomass/coal co-firing in utility boilers [664–667]. Biomass usually has lower sulfur content than coal, so co-firing reduces SO_x emissions because of displacement of sulfur in the fuel blend. Similar reductions are also observed for NO_x emissions, because the nitrogen content of the co-fired biomass fuels is generally lower than the nitrogen content of the coal. Initially, any change in NO_x emissions as a result of blending the biomass and coal can be attributed to changing the amount of nitrogen in the fuel blend.

The addition of biomass has been shown to reduce NO_x emissions in most commercial facilities, usually beyond the reductions expected because of a lower overall fuel-bound nitrogen content. The high volatiles content of biomass can effectively establish a fuel-rich zone early in the burner flame that can reduce NO_x emissions similar to fuel injection in flue gases (reburning) for NO_x reduction. Adding biomass can also reduce flame temperatures, leading to lower levels of thermal NO_x. The high moisture content of some biomass may also be effective for NO_x reduction at full-scale.

When biomass is co-fired with coal (even in small percentages), the alkali metals in biomass ash can alter the properties of the resulting mixed ash. This could have a significant impact on the coal plant's operating and maintenance costs or even operability. The addition of biomass to a coal-fired power plant can also nullify ash sales contracts for coal flyash. Biomass ash components in feedstocks may also reduce the long-term efficiency and effectiveness of certain (selective catalytic reduction, SCR) systems for the SCR of NO_x.

Fuel preparation and whether to premix the biomass and coal or introduce the two fuels separately into the boiler is another important issue that needs to be addressed for successful implementation of co-firing. Fuel handling of biomass in co-firing systems will need to be demonstrated with a variety of biomass feedstocks—such as switchgrass, willow, and energy cane—to take advantage of lower cost biomass residues and future energy crops.

Biomass co-firing is not a new technology; several utilities have taken advantage of opportunities and have or had been co-firing biomass for many years. One example is the AES (formerly NYSEG) Greenidge Station in Dresden, New York, that began co-firing tests in October 1994 [668]. The success of these tests led to the installation of a separate biomass preparation area in the fuel yard in early 1999, to separately feed biomass into the 108-MWe tangentially fired pulverized-coal unit. Another example is Steam Plant #2 at Tacoma Public Utilities' power plant in Tacoma, Washington. This plant was reconfigured in 1991 to co-fire coal, biomass, and refuse-derived fuels on a continuous basis in a 50-MWe atmospheric pressure fluidized-bed combustor. A third example is Northern States Power's King Station in Bayport, Minnesota. From 1987 through 1997, wood residues from the neighboring Anderson Window plant were continuously co-fired at a level of 5% (20 t/h) in a 550-MWe coal-fired cyclone boiler [668]. Co-firing was suspended at the King Station because Anderson Window found a higher value use for its wood residues, not for technical reasons.

Aside from these longer duration co-firing operations, several utilities have tested bio-mass/coal co-firing for short durations in utility boilers [669–673]. These various tests sought to systematically determine the impacts of biomass co-firing on such in-furnace parameters as boiler derating, boiler efficiency, emissions reductions, changes in fouling and slagging behavior, and corrosion. A number of demonstration projects are currently being funded to add to the knowledge base of utility-scale biomass/coal co-firing.

Several issues clearly remain regarding how blending biomass and coal will affect combustion performance, emissions, fouling and slagging propensities, corrosion, and ash salability. Nevertheless, several utilities have tested biomass/coal co-firing in utility boilers and the Electric Power Research Institute funded a study in 1997 to establish biomass co-firing guidelines [674].

Acknowledgment This work was supported by the U.S. Department of Energy Office of the Biomass Program.

References

1. Hall DO (1978) Solar energy use through biology—past, present and future. Solar Energy 22:307–328
2. USDOE international energy annual 2000, DOE/EIA-0219(2000). U.S. Department of Energy, Energy Information Administration, Washington, DC (2002)

3. Goldemberg J, Monaco L, Macedo I (1993) The Brazilian fuel-alcohol program. In: Johansson T, Kelly H, Reddy A, Williams R (eds) Renewable energy: sources for fuels and electricity. Island Press, Washington, DC
4. Rhodes A, Fletcher D (1966) Principles of industrial microbiology. Pergamon, New York
5. Grohmann K, Himmel M (1991) Chapter 1: enzymes for fuels and chemical feedstocks. In: Leatham GF, Himmel ME (eds) Enzymes in biomass conversion, vol 460. American Chemical Society, Washington, DC, pp 2–11
6. Sjostrom E (1993) Wood chemistry: fundamentals and applications. Academic, San Diego, pp 13–17
7. Fry SC, Miller JG (1989) Toward a working model of the growing plant cell wall. In: Lewis NG, Paice MG (eds) Plant cell wall polymers. American Chemical Society, Washington, DC, pp 33–46
8. Chum HL, Overend R (2003) Biomass and bioenergy in the United States. In: Goswami DY (ed) Advances in solar energy: an annual review, vol 15(3). American Solar Energy Society, Boulder, CO, pp 83–148
9. Haq Z (2002) Biomass for electricity. Energy Information Administration, U.S. Department of Energy, Washington, DC
10. Perry H (1974) The gasification of coal. Sci Am 230(3):19–25
11. McKendry P (2002) Energy production from biomass (part 3): gasification technologies. Bioresour Technol 83:55–63
12. Rapagna S, Jand N, Foscolo P (1998) Catalytic gasification of biomass to produce hydrogen rich gas. Int J Hydrogen Energy 23 (7):551–557
13. USDOE (2000) A biopower triumph—the gasification story, DOE/GO102000-1058. Office of Energy Efficiency and Renewable Energy, U.S. Department of Energy, Washington, DC
14. Sherrard EC, Gauger WH (1923) Effect of salts upon the acid hydrolysis of wood. Ind Eng Chem 15(1):63–64
15. Sherrard EC (1922) Ethyl alcohol from western larch—larix occidentalis, nuttal. Ind Eng Chem 14(10):948–949
16. LaForge FB, Hudson CS (1918) The preparation of several useful substances from corn cobs. J Ind Eng Chem 10(11):925–927
17. Braconnot H (1819) Verwandlungen des Holzstoffs mittelst Schwefelsaure in Gummi, Zucker und eine eigne Saure, und mittelst Kali in Ulmin. Ann Phys January:547–571
18. Dunning J, Lathrop E (1945) The saccharification of agricultural residues: a continuous process. Ind Eng Chem 37(1):24–29
19. Faith W (1945) Development of the Scholler process in the United States. Ind Eng Chem 37(1):9–11
20. Saeman J (1945) Kinetics of wood saccharification: hydrolysis of cellulose and decomposition of sugars in dilute acid at high temperature. Ind Eng Chem 37(1):43–51
21. Sherrard E, Kressman F (1945) Review of processess in the United States prior to World War II. Ind Eng Chem 37(1):5–8
22. Harris E, Berlinger E (1946) Madison wood-sugar process, R1617. U.S. Forest Product Laboratory, Madison, WI
23. Harris E (1949) Wood saccharification. In: Advances in carbohydrate chemistry. Academic, New York, vol 4, pp 163–188
24. Gilbert N, Hobbs I, Levine J (1952) Hydrolysis of wood using dilute sulfuric acid. Ind Eng Chem 44(7):1712–1720
25. Wenzl H (1970) Chapter IV: the acid hydrolysis of wood. In: The chemical technology of wood. Academic, New York, pp 157–252
26. Grethlein H (1978) Comparison of the economics of acid and enzymatic hydrolysis of newsprint. Biotechnol Bioeng 20:503–525
27. Wright J (1983) High-temperature acid hydrolysis of cellulose for alcohol fuel production, SERI/TR-231-1714. Solar Energy Research Institute, Golden, CO
28. Grohmann K, Himmel M, Rivard C, Tucker M, Baker J, Torget R, Graboski M (1984) Chemical–mechanical methods for the enhanced utilization of straw. In: Biotechnology and bioengineering symposium. Wiley, New York, pp 137–157
29. Church J, Woolbridge D (1981) Continuous high solids acid hydrolysis of biomass in a 1.5 in plug flow reactor. Ind Eng Chem Prod Res Devel 20:371–378
30. Wright J, d'Agnicourt C (1984) Evaluation of sulfuric acid hydrolysis for alcohol fuel production. In: Biotechnology and bioengineering symposium. Wiley, New York, pp 105–123
31. Harris J, Baker A, Conner A, Jefferies T, Minor J et al (1985) Two-stage dilute sulfuric acid hydrolysis of wood: an investigation of fundamentals. General technical report FPL-45. U.S. Forest Products Laboratory, Madison, WI
32. Wirght J, Power A, Bergeron P (1985) Evaluation of concentrated halogen acid hydrolysis processes for alcohol fuel production, SERI/TR-232-2386. Solar Energy Research Institute, Golden, CO
33. Farone WA, Cuzens JE (1996) Methods of producing sugars using strong acid hydrolysis of cellulosic and hemicellulosic materials. US Patent 5,562,777, 8 Oct 1996
34. Wright JD, D'Agincourt CG (1984) Evaluation of sulfuric acid hydrolysis processes for alcohol fuel production. Biotechnol Bioeng Symp 14:105–121
35. Grohmann K, Torget R, Himmel M (1985) Optimization of dilute acid pretreatment of biomass. Biotechnol Bioeng Symp 15:59–80
36. Kong F, Engler CR, Soltes EJ (1993) Effects of cell-wall acetate, xylan backbone, and lignin on enzymatic hydrolysis of aspen wood. Appl Biochem Biotechnol 34/35:23–35
37. Chang V, Holtzapple M (2000) Fundamental factors affecting biomass enzymatic reactivity. Appl Biochem Biotechnol 84–86:5–37
38. Vinzant TB, Ehrman CI, Himmel ME (1997) SSF of pretreated hardwoods: effect of native lignin content. Appl Biochem Biotechnol 62:97–102
39. Henrissat B (1994) Cellulases and their interaction with cellulose. Cellulose 1:169–196
40. Zhang M, Frandan M, Newman J, Finkelstein M, Picataggio S (1995) Promising ethanologens for xylose fermentation scientific note. Appl Biochem Biotechnol 51/52:527–536
41. Zhang M, Deanda K, Finkelstein M, Picataggion S (1995) Metabolic engineering of a pentose metabolism pathway in ethanologenic Zymomonas mobilis. Science 267:240–243
42. Zhang M, Chou Y, Piccatagio S, Finkelstein M (1996) Single Zymomonas mobilis strain for xylose and arabinose fermentation, NREL IR#95-51. National Renewable Energy Laboratory, Golden, CO
43. Ohta K, Beall D, Mejia J, Ingram L (1991) Genetic improvement of Escherichia for ethanol production: chromosol integration of Zymomonas mobilis genes encoding pyruvate decarboxylase and alcohol dehydrogenase II. Appl Environ Microbiol 57(4):893–900
44. Lawford H, Rousseau J (1991) Energy from biomass and waste. In: Klass D (ed) Xylose to ethanol: enhanced yield and productivity using genetically engineered Escherichia coli. Chicago, pp 583–623
45. Ingram L, Ohta K, Beall D (1991) Energy from biomass and wastes. In: Klass D (ed) Genetic modification of E. coli for ethanol production. Institute of Gas Technology, Chicago, pp 1105–1125
46. Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66(3):506–577
47. ISO (1997) ISO 14040: environmental management—life cycle assessment—principles and framework, ISO 14040: 1997(E). International Standardization Organization, Geneva, 15 June 1997
48. ISO (2000) ISO/TR 14049: environmental management—life cycle assessment—examples of application of ISO 14041 to goal and scope definition and inventory analysis, ISO/TR 14049. International Organization for Standardization, Geneva, 15 March 2000

49. ISO (1998) ISO 14041: environmental management—life cycle assessment—goal and scope definition and inventory analysis, ISO 14041:1998(E). International Organization for Standardization, Geneva, 1 Oct 1998
50. ISO (1996) ISO 14004: environmental management systems—general guidelines on principles, systems and supporting techniques, ISO 14004:1996(E). International Organization for Standardization, Geneva, 1 Sept 1996
51. ISO (2000) ISO 14043: environmental management systems—life cycle assessment—life cycle interpretation, ISO 14043:2000(E). International Organization for Standardization, Geneva, 1 March 2000
52. ISO (1996) ISO 14001: environmental management systems—specification with guidance for use, ISO 14001:1996(E). International Organization of Standardization, Geneva, 1 Sept 1996
53. Spath P, Mann M (2004) Biomass power and conventional fossil systems with and without CO_2 sequestration—comparing the energy balance, greenhouse gas emissions and economics, NREL/TP-510-32575. National Renewable Energy Laboratory, Golden, CO
54. Spath P, Dayton D (2003) Preliminary screening—technical and economic assessment of synthesis gas to fuels and chemicals with emphasis on the potential for biomass-derived syngas, NREL/TP-510-34929. National Renewable Energy Laboratory, Golden, CO
55. U.S. Department of Energy (1997) Energy information administration, annual energy review, vol DOE/EIA-0383. U.S. Department of Energy, Washington, DC
56. Lynd LR, Elander RT, Wyman CE (1996) Likely features and costs of mature biomass ethanol technology. Appl Biochem Biotechnol 57/58:741–761
57. Hoffert M, Caldeira K, Benford G, Crisell D, Green CJ et al (2002) Advanced technology paths to global climate stability: energy for a greenhouse plant. Science 298:981–987
58. NRC (2000) Bio-based industrial products: priorities for research and commercialization. Committee on Bio-Based Industrial Products, National Research Council, Washington, DC
59. Johansson T, Kelly H, Reddy A, Williams R (1996) Renewable fuels and electricity for a growing world economy. In: Johansson T, Kelly H, Reddy A, Williams R (eds) Renewable energy: sources for fuels and electricity. Island Press, Washington, DC
60. Lynd LR, Jin H, Michels JG, Wyman CE, Dale B (2003) Bioenergy: background, potential, and policy: a policy briefing prepared for the Center for Strategic and International Studies. Center for Strategic and International Studies, Washington, DC
61. Lovins AB, Datta EK, Bustnes O-E, Koomey JG, Glasgow N (2004) Winning the oil endgame: innovation for profits, jobs and security. Rocky Mountain Institute, Snowmass, CO
62. Greene N (2004) Growing energy: how biofuels Can help end America's oil dependence. Natural Resources Defense Council, New York
63. Helle S, Cameron D et al (2003) Effect of inhibitory compounds found in biomass hydrolysates on growth and xylose fermentation by a genetically engineered strain of *Saccharomyces cerevisiae*. Enzyme Microb Technol 33(6):786–792
64. Klinke HB, Thomsen AB et al (2004) Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. Appl Microbiol Biotechnol 66 (1):10–26
65. Dien BS, Cotta MA, Jeffries TW (2003) Bacteria engineered for fuel ethanol production: current status. Appl Microbiol Biotechnol 63(3):258–266
66. Dupreez JC, Bosch M, Prior BA (1986) The fermentation of hexose and pentose sugars by Candida-shehatae and Pichia-stipitis. Appl Microbiol Biotechnol 23(3–4):228–233
67. Ingram LO, Conway T, Clark DP, Sewell GW, Preston JF (1987) Genetic engineering of ethanol production in Escherichia coli. Appl Environ Microbiol 53(10):2420–2425
68. Ohta K, Beall DS, Mejia JP, Shanmugam KT, Ingram LO (1991) Metabolic engineering of Klebsiella oxytoca M5A1 for ethanol production from xylose and glucose. Appl Environ Microbiol 57 (10):2810–2815
69. Ohta K, Beall DS, Mejia JP, Shanmugam KT, Ingram LO (1991) Genetic improvement of Escherichia coli for ethanol production: chromosomal integration of Zymomonas mobilis genes encoding pyruvate decarboxylase and alcohol dehydrogenase II. Appl Environ Microbiol 57(4):893–900
70. Lawford HG, Rousseau JD (1992) Effect of acetic acid on xylose conversion to ethanol by genetically engineered E. coli. Appl Biochem Biotechnol 34–35:185–204
71. Hahn-Hagerdal B, Linden T, Senac T, Skoog K (1991) Ethanolic fermentation of pentoses in lignocellulose hydrolysates. Appl Biochem Biotechnol 28–29:131–144
72. Padukone N, Evans KW, McMillan JD, Wyman CE (1995) Characterization of recombinant E. coli ATCC 11303 (pLOI 297) in the conversion of cellulose and xylose to ethanol. Appl Microbiol Biotechnol 43(5):850–855
73. Asghari A, Bothast RJ, Doran JB, Ingram LO (1996) Ethanol production from hemicellulose hydrolysates of agricultural residues using genetically engineered Escherichia coli strain KO11. J Ind Microbiol 16(1):42–47
74. Zhang M, Eddy C, Deanda K, Franden MA, Finkelstein M, Picataggio S (1995) Metabolic engineering of Zymomonas-mobilis for ethanol-production from renewable feedstocks. Abstr Paper Am Chem Soc 209:115-BTEC
75. Zhang M, Franden MA, Newman M, McMillan J, Finkelstein M, Picataggio S (1995) Promising ethanologens for xylose fermentation—scientific note. Appl Biochem Biotechnol 51–2:527–536
76. Zhang M, Eddy C, Deanda K, Finkestein M, Picataggio S (1995) Metabolic engineering of a pentose metabolism pathway in ethanologenic Zymomonas-mobilis. Science 267(5195):240–243
77. Deanda K, Zhang M, Eddy C, Picataggio S (1996) Development of an arabinose-fermenting Zymomonas mobilis strain by metabolic pathway engineering. Appl Environ Microbiol 62 (12):4465–4470
78. Ho NWY, Chen ZD, Brainard AP (1998) Genetically engineered Saccharomyces yeast capable of effective cofermentation of glucose and xylose. Appl Environ Microbiol 64(5):1852–1859
79. Toon ST, Philippidis GP, Ho NWY, Chen ZD, Brainard A et al (1997) Enhanced cofermentation of glucose and xylose by recombinant Saccharomyces yeast strains in batch and continuous operating modes. Appl Biochem Biotechnol 63–5:243–255
80. Kuyper M, Harhangi HR, Stave AK, Winkler AA, Jetten MS et al (2003) High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by Saccharomyces cerevisiae? FEMS Yeast Res 4(1):69–78
81. Kuyper M, Hartog MM, Toirkens MJ, Almering MJ, Winkler AA, van Dijken JP, Pronk JT (2005) Metabolic engineering of a xylose-isomerase-expressing Saccharomyces cerevisiae strain for rapid anaerobic xylose fermentation. FEMS Yeast Res 5 (4–5):399–409
82. Harhangi HR, Akhmanova AS, Emmens R, van der Drift C, de Laat WT et al (2003) Xylose metabolism in the anaerobic fungus Piromyces sp. strain E2 follows the bacterial pathway. Arch Microbiol 180(2):134–141
83. Jeffries TW, Kurtzman CP (1994) Strain selection, taxonomy, and genetics of close-fermenting yeasts. Enzyme Microb Technol 16 (11):922–932
84. Jeffries TW, Dahn K, Kenealy WR, Pittman P, Sreenath HK, Davis BP (1994) Genetic-engineering of the xylose-fermenting

yeast Pichia-stipitis for improved ethanol-production. Abstr Paper Am Chem Soc 207:167-BTEC

85. Klapatch TR, Guerinot ML, Lynd LR (1996) Electrotransformation of Clostridium thermosaccharolyticum. J Ind Microbiol 16(6):342–347
86. Bothast RJ, Saha BC, Flosenzier AV, Ingram LO (1994) Fermentation of L-arabinose, D-xylose and D-glucose by ethanologenic recombinant *Klebsiella oxytoca* strain P2. Biotechnol Lett 16 (4):401–406
87. Sedlak M, Ho NWY (2001) Expression of *E. coli* araBAD operon encoding enzymes for metabolizing L-arabinose in *Saccharomyces cerevisiae*. Enzyme Microb Technol 28(1):16–24
88. Himmel ME, Adney WS, Baker JO, Elander R, McMillan JD et al (1997) Advanced bioethanol production technologies: a perspective. In: Saha BC, Woodward J (eds) Fuels and chemicals from biomass, vol 666. American Chemical Society, Washington, DC, pp 2–45
89. Toon ST, Philippidis GP, Ho NWY, Chen ZD, Brainard A et al (1997) Enhanced cofermentation of glucose and xylose by recombinant *Saccharomyces* yeast strains in batch and continuous operating modes. Appl Biochem Biotechnol 63(5):243–255
90. Zhang Y-HP, Lynd LR (2005) Cellulose utilization by Clostridium thermocellum: bioenergetics and hydrolysis product assimilation. Proc Natl Acad Sci U S A 102(20):7321–7325
91. VanRensburg P, Zyl WHV, Pretorius IS (1998) Engineering yeast for efficient cellulose degradation. Yeast 14(1):67–76
92. Godbole S, Decker SR, Nieves RA, Adney WS, Vinzant TB et al (1999) Cloning and expression of *Trichoderma reesei* CBH I in *Pichia pastoris*. Biotechnol Prog 15(3):828–833
93. VanRooyen R, Hahn-Hagerdal B, Grange DCL, VanZyl WH (2003) Comparative expression of novel beta-glucosidases in Saccharomyces cerevisiae. Yeast 20:S223
94. ASTM Methods available from American Society for Testing and Materials at http://www.ASTM.org; E1690-01 Determining ethanol extractives in biomass; E1721-01 Determining acid-insoluble residue in biomass; E1755-01 Determining ash in biomass; E1756-01 Determining total solids in biomass; E1757-0101 Preparing biomass for compositional analysis; E1758-0101 Determining carbohydrates in biomass by high performance liquid chromatography (HPLC); E1821-96 Determining carbohydrates in biomass by gas chromatograph (GC). http://www.ASTM.org. April
95. http://www.ott.doe.gov/biofuels/analytical_methods.html, n. R. e. l. p. a. Standard Laboratory Analytical Methods (LAPs) for biomass compositional analysis. http://www.ott.doe.gov/biofuels/analytical_methods.html. April 2005
96. Ehrman CI (1996) Methods for the chemical analysis of biomass process streams. In: Wyman CE (ed) Handbook on bioethanol: production and utilization. Taylor & Francis, Washington, DC
97. DiFoggio R (1995) Examination of some misconceptions about near-infrared analysis. Appl Spectrosc 49(1):67–75
98. Wisconsin Wisconsin Corn Hybrid Trials. http://corn.agronomy.wisc.edu/. Accessed 4 July 2005
99. Minnesota Minnesota Agricultural Experiment Station. http://www.maes.umn.edu/maespubs/vartrial/cropages/cornpage.html. Accessed 4 July 2005
100. Shull GH (1909) A pure-line method in corn breeding. Am Breed Assoc Rep 5:51–59
101. East EM (1936) Heterosis. Genetics 21:375–397
102. United States Department of Agriculture National Agricultural Statistics Service. http://www.nass.usda.gov/. Accessed 4 July 2005
103. Hallauer AR, Russell WA, Lamkey KR (1988) Corn breeding. In: Sprague GF, Dudley JW (eds) Corn and corn improvement, 3rd edn. Crop Science Society of America, Madison, WI, pp 463–564
104. Peterson PA (1993) Transposable elements in maize: their role in creating plant genetic variability. Adv Agron 51:79–124
105. Alexander DE (1988) Breeding special nutritional and industrial types. In: Sprague GF, Dudley JW (eds) Corn and corn improvement, 3rd edn. American Society of Agronomy, Madison, WI, pp 869–880
106. Dudley JW (1977) Seventy-six generations of selection for oil and protein in maize. In: Pollak E et al (eds) Proceedings of the international conference on quantitative genetics. Iowa State University Press, Ames, pp 459–473
107. Dudley JW, Lambert RJ (2004) 100 generations of selection for oil and protein content in corn. Plant Breed Rev 24(1):79–110
108. Miller RL, Dudley JW, Alexander DE (1981) High intensity selection for percent oil in corn. Crop Sci 21:455–457
109. Dudley JW, Dijkhuizen A, Paul C, Coates ST, Rocheford TR (2004) Effects of random mating on marker-QTL associations in the cross of the Illinois high protein x Illinois low protein maize strains. Crop Sci 44:1419–1428
110. Seki MMN, Kamiya A, Ishida J, Satou M, Sakurai T, Nakajima M et al (2002) Functional annotation of a full-length *Arabidopsis* cDNA collection. Science 296:141–145
111. Yu JSH, Wang J, Wong GKS, Li S, Liu B et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). Science 296:79–92
112. Goff SA, Ricke D, Lan TH, Presting G, Wang R et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). Science 296:92–100
113. JGI, J. G. I. Populus trichocarpa v.1.0. http://genome.jgi-psf.org/Poptrl/Poptrl.info.html. Accessed 4 June 2005
114. TIGR TIGR Maize Gene Index. http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species=maize. Accessed 4 May 2005
115. Carpita NC, Tierney M, Campbell M (2001) Molecular biology of the plant cell wall: searching for the genes that define structure, architecture and dynamics. Plant Mol Biol 47:1–5
116. Carpita cell wall genomics. http://cellwall.genomics.purdue.edu/. Accessed 4 June 2005
117. Keegstra WallBioNet. http://xyloglucan.prl.msu.edu/. Accessed 4 June 2005
118. Somerville Chris Somerville. http://www-ciwdpb.stanford.edu/research/research_csomerville.php. Accessed 4 June 2005
119. Pear JR, Kawagoe Y, Schreckengost WE, Delmer DP, Stalker DM (1996) Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase. Proc Natl Acad Sci U S A 93(22):12637–12642
120. Appenzeller L, Doblin M, Barreiro R, Wang H, Niu X et al (2004) Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (CesA) gene family. Cellulose 11:287–299
121. Liepman AH, Wilkerson CG, Keegstra K (2005) Expression of cellulose synthase-like (Csl) genes in insect cells reveals that CslA family members encode mannan synthases. Proc Natl Acad Sci U S A 102(6):2221–2226
122. Reiter WD, Vanzin GF (2001) Molecular genetics of nucleotide sugar interconversion pathways in plants. Plant Mol Biol 47 (1–2):95–113
123. Gillmor CS, Lukowitz W, Brininstool G, Sedbrook JC, Hamann T et al (2005) Glycosylphosphatidylinositolanchored proteins are required for cell wall synthesis and morphogenesis in Arabidopsis. Plant Cell 17(4):1128–1140
124. Milne TA, Chum HL, Agblevor F, Johnson DK (1992) Standardized analytical methods. Biomass Bioenergy 2 (1–6):341–366
125. Milne TA, Brennan AH, Glenn BH (1990) Sourcebook of methods of analysis for biomass and biomass conversion processes. Elsevier, New York, p 341
126. Tkachuk R (1977) Calculation of the nitrogen-to-protein conversion factor. In: Hulse JH, Rachie KO, Billingsley LW (eds)

Nutritional standards and methods of evaluation for food legume breeders. Bernan Associates, Lanham, pp 78–82
127. Mossé J (1990) Nitrogen to protein conversion factor for 10 cereals and 6 legumes or oilseeds—a reappraisal of its definition and determination—variation according to species and to seed protein content. J Agric Food Chem 38(1):18–24
128. Hames BR, Thomas SR, Sluiter AD, Roth CJ, Templeton DW (2003) Rapid biomass analysis: new tools for the compositional analysis of corn stover feedstocks and process intermediates from ethanol production. Appl Biochem Biotechnol 105–108, Proceedings of the 24th international symposium for the biotechnology of fuels and chemicals, pp 1–12
129. Hames BR, Thomas SR, Sluiter AD, Roth CJ, Templeton DW (2003) Rapid biomass analysis. New tools for compositional analysis of corn stover feedstocks and process intermediates from ethanol production. Appl Biochem Biotechnol 105–108:5–16
130. Aden A, Ruth M, Ibsen K, Jechura J, Neeves K et al (2002) Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis for corn stover. NREL report no. TP-510-32438 available online at URL http://www.nrel.gov/publications/. National Renewable Energy Laboratory, Golden, CO, June 2002, p 154
131. NREL. A calculator for theoretical ethanol yields form biomass. http://www.ott.doe.gov/biofuels/ethanol_calculator.html. Accessed April 2005
132. Satoh H, Nishi A, Fujita N, Kubo A, Nakamura Y, Kawasaki T, Okita TW (2003) Isolation and characterization of starch mutants in rice. J Appl Glycosci 50(2):225–230
133. Nelson OE (1966) Mutant genes that change the composition of maize endosperm proteins. Fed Proc 25(6):1676–1678
134. Ohlrogge JB, Browse J, Somerville CR (1991) The genetics of plant lipids. Biochim Biophys Acta 1082(1):1–26
135. Reiter W-D, Chappie C, Somerville CR (1997) Mutants of Arabidopsis thaliana with altered cell wall polysaccharide composition. Plant J 12(2):335–345
136. Mouille G, Robin S, Lecomte M, Pagant S, Hofte H (2003) Classification and identification of Arabidopsis cell wall mutants using Fourier-transform infrared (FT-IR) microspectroscopy. Plant J 35(3):393–404
137. Yong WD, Link B, O'Malley R, Tewari J, Hunter CT III et al (2005) Genomics of plant cell wall biogenesis. Planta 221:747–751
138. Baker JO, Thomas SR, Adney WS, Nieves RA, Himmel ME (1994) The cellulase synergistic effect—binary and ternary-systems. Abstr Paper Am Chem Soc 207:50-AGFD
139. Nidetzky B, Steiner W, Hayn M, Claeyssens M (1994) Cellulose hydrolysis by the cellulases from Trichoderma reesei: a new model for synergistic interaction. Biochem J 298(pt 3):705–710
140. Gatenholm P, Tenkanen M (2004) Hemicelluloses: science and technology, vol 864. American Chemical Society, Washington, DC, p 388
141. Galante Y, Formantici C (2003) Enzyme applications in detergency and in manufacturing industries. Curr Org Chem 7 (13):1399–1422
142. Hoondal G, Tiwari R, Tewari R, Dahiya N, Beg Q (2002) Microbial alkaline pectinases and their industrial applications: a review. Appl Microbiol Biotechnol 59(4–5):409–418
143. Kashyap DR, Vohra PK, Chopra S, Tewari R (2001) Applications of pectinases in the commercial sector: a review. Bioresour Technol 77(3):215–227
144. Lebeda A, Luhova L, Sedlarova M, Jancova D (2001) The role of enzymes in plant-fungal pathogens interactions—review. Z Pflanzenk Pflanzen 108(1):89–111
145. Naidu G, Panda T (1998) Production of pectolytic enzymes—a review. Bioprocess Eng 19(5):355–361
146. Burke R, Cairney J (2002) Laccases and other polyphenol oxidases in ecto- and ericoid mycorrhizal fungi. Mycorrhiza 12 (3):105–116
147. Shah V, Nerud F (2002) Lignin degrading system of white-rot fungi and its exploitation for dye decolorization. Canad J Microbiol 48(10):857–870
148. Tuomela M, Vikman M, Hatakka A, Itavaara M (2000) Biodegradation of lignin in a compost environment: a review. Bioresour Technol 72(2):169–183
149. Garg S, Modi D (1999) Decolorization of pulp-paper mill effluents by white-rot fungi. Crit Rev Biotechnol 19(2):85–112
150. Cairney J, Burke R (1998) Extracellular enzyme activities of the ericoid mycorrhizal endophyte Hymenoscyphus ericae (Read) Korf & Kernan: their likely roles in decomposition of dead plant tissue in soil. Plant Soil 205(2):181–192
151. Cullen D (1997) Recent advances on the molecular genetics of ligninolytic fungi. J Biotechnol 53(2–3):273–289
152. Zadrazil F, Kamra D, Isikhuemhen O, Schuchardt F, Flachowsky G (1996) Bioconversion of lignocellulose into ruminant feed with white rot fungi—review of work done at the FAL, Braunschweig. J Appl Animal Res 10(2):105–124
153. Moss GP (1992) Recommendations of the nomenclature committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzyme-catalysed reactions. URL:http://www.chem.qmul.ac.uk/iubmb/enzyme/
154. Eriksson K-EL, Cavaco-Paulo A (1998) Enzyme applications in fiber processing. In: ACS symposium series, vol 687. American Chemical Society, Washington, DC
155. Saddler JN, Penner MH (1994) Enzymatic degradation of insoluble carbohydrates. In: ACS symposium series, vol 618. American Chemical Society, Washington, DC
156. Mansfield SD, Saddler JN (2003) Applications of enzymes to lignocellulosics. In: ACS symposium series, vol 855. American Chemical Society, Washington, DC
157. Leatham GF, Himmel ME (1991) Enzymes in biomass conversion. In: ACS symposium series, vol 460. American Chemical Society, Washington, DC
158. Jeffries TW, Viikari L (1996) Enzymes for pulp and paper processing. In: ACS symposium series, vol 655. American Chemical Society, Washington, DC
159. Himmel ME, Baker JO, Saddler JN (2001) Glycosyl hydrolases for biomass conversion. In: ACS symposium series, vol 769. American Chemical Society, Washington, DC
160. Himmel ME, Baker JO, Overend RP (1994) Enzymatic conversion of biomass for fuels production. In: ACS symposium series, vol 566. American Chemical Society, Washington, DC
161. Schomburg I, Chang A, Ebeling C, Gremse M, Heldt C et al (2004) BRENDA, the enzyme database: updates and major new developments. Nucleic Acids Res 32(Database issue):D431–D433
162. Boeckmann B, Bairoch A, Apweiler R, Blatter M-C, Estreicher A et al (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res 31(1):365–370
163. Coutinho PM, Henrissat B (1999) CAZy—carbohydrate active enzymes. URL:http://afmb.cnrs-mrs.fr/CAZY/
164. Bairoch A (2000) The ENZYME database in 2000. Nucleic Acids Res 28:304–305
165. Mandels M, Miller GL, Slater RW (1961) Separation of fungal carbohydrases by starch block zone electrophoresis. Arch Biochem Biophys 93:115–121
166. Pere J, Puolakka A, Nousiainen P, Buchert J (2001) Action of purified Trichoderma reesei cellulases on cotton fibers and yarn. J Biotechnol 89(2–3):247–255
167. Samejima M, Sugiyama J, Igarashi K, Eriksson K-EL (1997) Enzymatic hydrolysis of bacterial cellulose. Carbohydr Res 305 (2):281–288

168. Helbert W, Sugiyama J, Ishihara M, Yamanaka S (1997) Characterization of native crystalline cellulose in the cell walls of Oomycota. J Biotechnol 57(1–3):29–37
169. Zauscher S, Klingenberg DJ (2000) Normal forces between cellulose surfaces measured with colloidal probe microscopy. J Colloid Interface Sci 229(2):497–510
170. Lee I, Evans BR, Woodward J (2000) The mechanism of cellulase action on cotton fibers: evidence from atomic force microscopy. Ultramicroscopy 82(1–4):213–221
171. Rutland MW, Carambassis A, Willing GA, Neuman RD (1997) Surface force measurements between cellulose surfaces using scanning probe microscopy. Colloids Surf A Physicochem Eng Asp 123–124:369–374
172. Gustafsson J, Ciovica L, Peltonen J (2003) The ultrastructure of spruce kraft pulps studied by atomic force microscopy (AFM) and X-ray photoelectron spectroscopy (XPS). Polymer 44(3):661–670
173. Li H, Rief M, Oesterhelt R, Gaub HE, Zhang X, Shen J (1999) Single-molecule force spectroscopy on polysaccharides by AFM—nanomechanical fingerprint of α-(1,4)-linked polysaccharides. Chem Phys Lett 305(3–4):197–201
174. Ramos LP, Zandon A, Filho A, Deschamps FC, Saddler JN (1999) The effect of *Trichoderma* cellulases on the fine structure of a bleached softwood kraft pulp. Enzyme Microb Technol 24 (7):371–380
175. Decker SR, Adney WS, Jennings E, Vinzant TB, Himmel ME (2003) An automated filter paper assay for determination of cellulase activity. Appl Biochem Biotechnol 105:689–703
176. Kong F, Engler CR, Soltes EJ (1992) Effects of cell-wall acetate, xylan backbone, and lignin on enzymatic hydrolysis of aspen wood. Appl Biochem Biotechnol 34/35:23–35
177. Tenkanen M, Poutanen K (1992) Significance of esterases in the degradation of xylans. In: Visser J, Beldman G, Kusters-van Someren MA, Voragen AGJ (eds) Xylans and xylanases. Elsevier, New York, pp 203–212
178. Poutanen K, Sundberg M (1988) An acetyl esterase of Trichoderma reesei and its role in the hydrolysis of acetyl xylans. Appl Microbiol Biotechnol 28:419–424
179. Poutanen K, Sundberg M, Korte H, Puls J (1990) Deacetylation of xylans by acetyl esterases of Trichoderma reesei. Appl Microbiol Biotechnol 33:506–510
180. Biely P, Vrsanska M, Kremnicky L, Alfoldi J, Tenkanen M et al (1994) Family F and G endo-beta-1,4-xylanases—differences in their performance on glucuronoxylan and rhodymenan. Abstr Paper Am Chem Soc 207:148-BTEC
181. Coughlan MP, Hazlewood GP (1993) Hemicellulose and hemicellulases. Portland, London, UK
182. Reese ET, Siu RGH, Levinson HS (1950) The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. J Bacteriol 59:485–497
183. Ryu D, Mandels M (1980) Cellulases: biosynthesis and applications. In: Reviews, pp 1–12
184. Himmel ME, Decker SR, Adney WS, Baker JO (2000) Cellulase animation. Copyright DOE/MRIPAU2-568-354, p 11
185. Miller GL (1959) Dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem 31:426–428
186. Ghose TK (1987) Measurement of cellulase activities. Pure Appl Chem 59(2):257–268
187. Philippidis GP (1994) Cellulase production technology—evaluation of current status. Enzym Convers Biomass Fuel Prod 566:188–217
188. Esterbauer H, Steiner W, Kreiner W, Sattler W, Hayn M (1992) Comparison of enzymatic hydrolysis in a worldwide round robin assay. Bioresour Technol 39:117–123
189. Adney WS, Mohagheghi A, Thomas SR, Himmel ME (1995) Comparison of protein contents of cellulase preparations in a worldwide round-robin assay. In: Saddler JN, Penner MH (eds) Enzymatic degradation of insoluble carbohydrates. American Chemical Society, Washington, DC, pp 256–271
190. Esterbauer H, Steiner LI, Hermann A, Hayn M (1991) Production of Trichoderma cellulase in laboratory and pilot scale. Bioresour Technol 36:51–65
191. Irwin DC, Spezio M, Walker LP, Wilson DB (1993) Activity studies of 8 purified cellulases—specificity, synergism, and binding domain effects. Biotechnol Bioeng 42(8):1002–1013
192. Ghose TK, Bisaria VS (1987) Measurement of hemicellulase activities. 1. Xylanases. Pure Appl Chem 59(12):1739–1751
193. Mandels M, Andreotti R, Roche C (1976) Measurement of saccharifying cellulase. Biotechnol Bioeng Symp 6:21–33
194. Wood TM, McCrae SI (1977) Cellulase from Fusarium solani purification and properties of the C-1 component. Carbohydr Res 57:117–133
195. Ghose TK, Pathak AN, Bisaria VS (1975) In: Bailey M, Enari T-M, Linko M (eds) Proceedings of the symposium on enzymatic hydrolysis of cellulose. VTT, Aulanko, Finland, p 111
196. Shoemaker SP, Brown RD Jr (1978) Characterization of endo-1,4-b-D-glucanases purified from Trichoderma viride. Biochim Biophys Acta 523:147–161
197. Sheir-Neiss G, Montenecourt BS (1984) Characterization of the secreted cellulases of Trichoderma reesei wild type and mutants during controlled fermentations. Appl Microbiol Biotechnol 20:46–53
198. Johnson EA, Sakajoh M, Halliwell G, Madia A, Demain AL (1982) Saccharification of complex cellulosic substrates by the cellulase system from Clostridium thermocellum. Appl Environ Microbiol 43:1125–1132
199. Somogyi M (1952) Notes on sugar determination. J Biol Chem 195:19–23
200. Nelson N (1944) A photometric adaptation of the Somogyi method for the determination of glucose. J Biol Chem 153:375–380
201. Johnston DB, Shoemaker SP, Smith GM, Whitaker JR (1998) Kinetic measurements of cellulase activity on insoluble substrates using disodium 2,2′ bicinchoninate. J Food Biochem 22:301–319
202. Vlasenko EY, Ryan AI, Shoemaker CF, Shoemaker SP (1998) The use of capillary viscometry, reducing end-group analysis, and size exclusion chromatography combined with multi-angle laser light scattering to characterize endo-1,4-beta-D-glucanases on carboxymethylcellulose: a comparative evaluation of the three methods. Enzyme Microb Technol 23:350–359
203. Baker JO, Vinzant TB, Ehrman CI, Adney WS, Himmel ME (1997) Use of a new membrane-reactor saccharification assay to evaluate the performance of cellulases under simulated SSF conditions—effect on enzyme quality of growing Trichoderma reesei in the presence of targeted lignocellulosic substrate. Appl Biochem Biotechnol 63–5:585–595
204. Rescigno A, Rinaldi AC, Curreli N, Olianas A, Sanjust E (1994) A dyed substrate for the assay of endo-1,4-beta-glucanases. J Biochem Biophys Methods 28(2):123–129
205. Sharrock KR (1988) Cellulase assay methods: a review. J Biochem Biophys Methods 17:81–106
206. Bartley TD, Murphy-Holland K, Eveleigh DE (1984) A method for the detection and differentiation of cellulase components in polyacrylamide gels. Anal Biochem 140(1):157–161
207. Carder JH (1986) Detection and quantitation of cellulase by Congo red staining of substrates in a cup-plate diffusion assay. Anal Biochem 153(1):75–79
208. Sattler W, Esterbauer H, Glatter O, Steiner W (1989) The effect of enzyme concentration on the rate of the hydrolysis of cellulose. Biotechnol Bioeng Symp 33:1221–1234
209. Adney WS, Ehrman CI, Baker JO, Thomas SR, Himmel ME (1994) Cellulase assays—methods from empirical mathematical-

models. In: Himmel ME, Baker JO, Overend RP (eds) Enzymatic conversion of biomass for fuels production, vol 566. ACS, Washington, DC, pp 218–235

210. Sieben A (1975) Cellulase and other hydrolytic enzyme assays using an oscillating tube viscometer. Anal Biochem 63(1):214–219
211. Kim CH (1995) Characterization and substrate specificity of an endo-beta-1,4-D-glucanase I (Avicelase I) from an extracellular multienzyme complex of Bacillus circulans. Appl Environ Microbiol 61(3):959–965
212. Chee KK (1990) Kinetic study of random chain scission by viscometry. J Appl Polymer Sci 41:985–994
213. Demeester J, Bracke M, Lauwers A (1979) Absolute viscometric method for the determination of endocellulase (Cx) activities based upon light-scattering interpretations of gel chromatographic fractionation data. Adv Chem Sen 181:91–125
214. Manning K (1981) Improved viscometric assay for cellulase methods. J Biochem Biotechnol 5:189–202
215. Gilkes NR, Kwan E, Kilburn DG, Miller RC, Warren RAJ (1997) Attack of carboxymethylcellulose at opposite ends by two cellobiohydrolases from Cellulomonas fimi. J Biotechnol 57 (1–3):83–90
216. Limam F, Chaabouni SE, Ghrir R, Marzouki N (1995) Two cellobiohydrolases of Penicillium occitanis mutant Pol 6: purification and properties. Enzyme Microb Technol 17(4):340–346
217. Hoshino E, Shiroishi M, Amano Y, Nomura M, Kanda T (1997) Synergistic actions of exo-type cellulases in the hydrolysis of cellulose with different crystallinities. J Ferment Bioeng 84 (4):300–306
218. Sadana JC, Patil RV (1985) Synergism between enzymes of *Sclerotium rolfsii* involved in the solubilization of crystalline cellulose. Carbohydr Res 140(1):111–120
219. van Tilbeurgh H, Claeyssens M, DeBruyne CK (1982) The use of 4-methylumbelliferyl and other chromophoric glycosides in the study of cellulolytic enzymes. FEBS Lett 149:152–156
220. Vrsanska M, Biely P (1992) The cellobiohydrolase I from Trichoderma reesei QM 9414: action on cello-oligosaccharides. Carbohydr Res 227:19–27
221. Konstantinidis AK, Marsden I, Sinnott ML (1993) Hydrolyses of alpha- and beta-cellobiosyl fluorides by cellobiohydrolases of Trichoderma reesei. Biochem J 291(pt 3):883–888
222. Claeyssens M, van Tilbeurgh H, Tomme P, Wood TM, McCrae SI (1989) Fungal cellulase systems: comparison of the specificities of the cellobiohydrolases isolated from Penicillium pinophilum and Trichoderma reesei. Biochem J 261:819–826
223. Beldman G, Searle-Van Leeuwen M, Rombouts F, Voragen F (1985) The cellulase of Trichoderma viride; purification, characterization and comparison of all detectable endoglucanases, exoglucanases and B-glucosidases. Eur J Biochem 146:301–308
224. Karlsson J, Siika-aho M, Tenkanen M, Tjerneld F (2002) Enzymatic properties of the low molecular mass endoglucanases Cell2A (EG III) and Cel45A (EG V) of *Trichoderma reesei*. J Biotechnol 99(1):63–78
225. Tuohy MG, Walsh DJ, Murray PG, Claeyssens M, Cuffe MM et al (2002) Kinetic parameters and mode of action of the cellobiohydrolases produced by *Talaromyces emersonii*. Biochim Biophys Acta 1596(2):366–380
226. Lemos MA, Teixeira JA, Domingues MRM, Mota M, Gama FM (2003) The enhancement of the cellulolytic activity of cellobiohydrolase I and endoglucanase by the addition of cellulose binding domains derived from *Trichoderma reesei*. Enzyme Microb Technol 32(1):35–40
227. Levy I, Shoseyov O (2002) Cellulose-binding domains: biotechnological applications. Biotechnol Adv 20(3–4):191–213
228. Linder M, Teeri TT (1997) The roles and function of cellulose-binding domains. J Biotechnol 57(1–3):15–28
229. Wood TM (1971) The cellulase of Fusarium solani: purification and specificity of the b-(164)-glucanase and the *β*-D-glucosidase components. Biochem J 121:353–362
230. Deshpande MV, Eriksson K-E, Pettersson LG (1984) *β*-glucanases in a mixture of cellulolytic enzymes. Anal Biochem 138:481–487
231. Adney WS, Baker JO, Vinzant TB, Thomas SR, Himmel ME (1995) Kinetic comparison of beta-D-glucosidases of industrial importance. Abstr Paper Am Chem Soc 209:119-BTEC
232. Biely P, Mislovicova D, Toman R (1985) Soluble chromogenic substrates for the assay of endo-1,4-beta-xylanases and endo-1,4-beta-glucanases. Anal Biochem 144(1):142–146
233. Biely P (1988) Detection and differentiation of cellulases and xylanases. Abstr Paper Am Chem Soc 195:202-CELL
234. Prade RA (1996) Xylanases: from biology to biotechnology. Biotechnol Genet Eng Rev 13:101–131
235. Gielkens MM, Visser J, de Graaff LH (1997) Arabinoxylan degradation by fungi: characterization of the arabinoxylan-arabinofuranohydrolase encoding genes from Aspergillus niger and Aspergillus tubingensis. Curr Genet 31(1):22–29
236. Tenkanen M (1998) Action of Trichoderma reesei and Aspergillus oryzae esterases in the deacetylation of hemicelluloses. Biotechnol Appl Biochem 27:19–24
237. Fillingham IJ, Kroon PA, Williamson G, Gilbert HJ, Hazlewood GP (1999) A modular cinnamoyl ester hydrolase from the anaerobic fungus Piromyces equi acts synergistically with xylanase and is part of a multi-protein cellulose-binding cellulase-hemicellulase complex. Biochem J 343(pt 1):215–224
238. Borneman WS, Ljungdahl LG, Hartley RD, Akin DE (1991) Isolation and characterization of p-coumaroyl esterase from the anaerobic fungus Neocallimastix strain MC-2. Appl Environ Microbiol 57(8):2337–2344
239. Cybinski DH, Layton I, Lowry JB, Dalrymple BP (1999) An acetylxylan esterase and a xylanase expressed from genes cloned from the ruminal fungus Neocallimastix patriciarum act synergistically to degrade acetylated xylans. Appl Microbiol Biotechnol 52(2):221–225
240. de Vries RP, Kester HC, Poulsen CH, Benen JA, Visser J (2000) Synergy between enzymes from Aspergillus involved in the degradation of plant cell wall polysaccharides. Carbohydr Res 327 (4):401–410
241. Saha BC (2000) α-1-Arabinofuranosidases: biochemistry, molecular biology and application in biotechnology. Biotechnol Adv 18 (5):403–423
242. Christov LP, Prior BA (1993) Esterases of xylan-degrading microorganisms: production, properties, and significance. Enzyme Microb Technol 15(6):460–475
243. Wood TM, McCrae SI (1996) Arabinoxylan-degrading enzyme system of the fungus Aspergillus awamori: purification and properties of an alpha-L-arabinofuranosidase. Appl Microbiol Biotechnol 45(4):538–545
244. Sakamoto T, Sakai T (1995) Analysis of structure of sugar-beet pectin by enzymatic methods. Phytochemistry 39(4):821–823
245. McKie VA, Black GW, Millward-Sadler SJ, Hazlewood GP, Laurie JI, Gilbert HJ (1997) Arabinanase A from Pseudomonas fluorescens subsp. cellulosa exhibits both an endo- and an exo-mode of action. Biochem J 323(pt 2):547–555
246. Matsuo N, Kaneko S, Kuno A, Kobayashi H, Kusakabe I (2000) Purification, characterization and gene cloning of two alpha-L-arabinofuranosidases from Streptomyces chartreusis GS901. Biochem J 346(pt 1):9–15
247. Yanai T, Sato M (2000) Purification and characterization of a novel alpha-L-arabinofuranosidase from Pichia capsulata X91. Biosci Biotechnol Biochem 64(6):1181–1188

248. Degrassi G, Vindigni A, Venturi V (2003) A thermostable α-arabinofuranosidase from xylanolytic *Bacillus pumilus*: purification and characterisation. J Biotechnol 101(1):69–79
249. Gomes J, Gomes I, Terler K, Gubala N, Ditzelm AG, Steiner W (2000) Optimisation of culture medium and conditions for α-1-arabinofuranosidase production by the extreme thermophilic eubacterium *Rhodothermus marinus*. Enzyme Microb Technol 27(6):414–422
250. Lee RC, Burton RA, Hrmova M, Fincher GB (2001) Barley arabinoxylan arabinofuranohydrolases: purification, characterization and determination of primary structures from cDNA clones. Biochem J 356(pt 1):181–189
251. Blum DL, Li XL, Chen H, Ljungdahl LG (1999) Characterization of an acetyl xylan esterase from the anaerobic fungus Orpinomyces sp. strain PC-2. Appl Environ Microbiol 65 (9):3990–3995
252. Castanares A, Wood TM (1992) Purification and characterization of a feruloyl/p-coumaroyl esterase from solid-state cultures of the aerobic fungus Penicillium pinophilum. Biochem Soc Trans 20 (3):275S
253. McDermid KP, Forsberg CW, MacKenzie CR (1990) Purification and properties of an acetylxylan esterase from Fibrobacter succinogenes S85. Appl Environ Microbiol 56(12):3805–3810
254. Khandke KM, Vithayathil PJ, Murthy SK (1989) Purification and characterization of an alpha-D-glucuronidase from a thermophilic fungus, Thermoascus aurantiacus. Arch Biochem Biophys 274 (2):511–517
255. Tenkanen M, Siika-aho M (2000) An α-glucuronidase of *Schizophyllum commune* acting on polymeric xylan. J Biotechnol 78(2):149–161
256. Nagy T, Nurizzo D, Davies GJ, Biely P, Lakey JH, Bolam DN, Gilbert HJ (2003) The α-glucuronidase, GlcA67A, of Cellvibrio japonicus utilises the carboxylate and methyl groups of aldobiouronic acid as important substrate recognition determinants. J Biol Chem 278(22):286–292
257. Jeffries TW (1996) Biochemistry and genetics of microbial xylanases. Curr Opin Biotechnol 7(3):337–342
258. Choi ID, Kim HY, Choi YJ (2000) Gene cloning and characterization of alpha-glucuronidase of Bacillus stearothermophilus no. 236. Biosci Biotechnol Biochem 64(12):2530–2537
259. Saraswat V, Bisaria VS (1997) Biosynthesis of xylanolytic and xylan-debranching enzymes in Melanocarpus albomyces IIS 68. J Ferment Bioeng 83(4):352–357
260. Nagy T, Emami K, Fontes CM, Ferreira LM, Humphry DR, Gilbert HJ (2002) The membrane-bound alpha-glucuronidase from Pseudomonas cellulosa hydrolyzes 4-O-methyl-D-glucuronoxylooligosaccharides but not 4-O-methyl-D-glucuronoxylan. J Bacteriol 184(17):4925–4929
261. de Vries RP, Poulsen CH, Madrid S, Visser J (1998) aguA, the gene encoding an extracellular alpha-glucuronidase from Aspergillus tubingensis, is specifically induced on xylose and not on glucuronic acid. J Bacteriol 180(2):243–249
262. Castanares A, Hay AJ, Gordon AH, McCrae SI, Wood TM (1995) D-xylan-degrading enzyme system from the fungus Phanerochaete chrysosporium: isolation and partial characterisation of an alpha-(4-O-methyl)-D-glucuronidase. J Biotechnol 43(3):183–194
263. Nishitani K, Nevins DJ (1991) Glucuronoxylan xylanohydrolase. A unique xylanase with the requirement for appendant glucuronosyl units. J Biol Chem 266(10):6539–6543
264. Christov L, Biely P, Kalogeris E, Christakopoulos P, Prior BA, Bhat MK (2000) Effects of purified endo-*β*-1,4-xylanases of family 10 and 11 and acetyl xylan esterases on eucalypt sulfite dissolving pulp. J Biotechnol 83(3):231–244
265. Biely P, Vrsanska M, Tenkanen M, Kluepfel D (1997) Endo-beta-1,4-xylanase families: differences in catalytic properties. J Biotechnol 57(1–3):151–166
266. Schwarz WH, Bronnenmeier K, Krause B, Lottspeich F, Staudenbauer WL (1995) Debranching of arabinoxylan: properties of the thermoactive recombinant alpha-L-arabinofuranosidase from Clostridium stercorarium (ArfB). Appl Microbiol Biotechnol 43(5):856–860
267. Bailey MJ, Poutanen K (1989) Production of xylanolytic enzymes by strains of Aspergillus. Appl Microbiol Biotechnol 30(1):5–10
268. Bronnenmeier K, Kern A, Liebl W, Staudenbauer WL (1995) Purification of Thermotoga maritima enzymes for the degradation of cellulosic materials. Appl Environ Microbiol 61(4):1399–1407
269. Dupont C, Daigneault N, Shareck F, Morosoli R, Kluepfel D (1996) Purification and characterization of an acetyl xylan esterase produced by Streptomyces lividans. Biochem J 319(pt 3):881–886
270. Green F III, Clausen CA, Highley TL (1989) Adaptation of the Nelson-Somogyi reducing-sugar assay to a microassay using microtiter plates. Anal Biochem 182(2):197–199
271. Lin J, Ndlovu LM, Singh S, Pillay B (1999) Purification and biochemical characteristics of beta-D-xylanase from a thermophilic fungus, Thermomyces lanuginosus-SSBP. Biotechnol Appl Biochem 30(pt 1):73–79
272. Ruiz-Arribas A, Fernandez-Abalos JM, Sanchez P, Garda AL, Santamaria RI (1995) Overproduction, purification, and biochemical characterization of a xylanase (Xysl) from Streptomyces halstedii JM8. Appl Environ Microbiol 61(6):2414–2419
273. Taguchi H, Hamasaki T, Akamatsu T, Okada H (1996) A simple assay for xylanase using o-nitrophenyl-beta-D-xylobioside. Biosci Biotechnol Biochem 60(6):983–985
274. Milagres AMF, Sales RM (2001) Evaluating the basidiomycetes *Poria medula-panis* and *Wolfiporia cocos* for xylanase production. Enzyme Microb Technol 28(6):522–526
275. Lin LL, Thomson JA (1991) An analysis of the extracellular xylanases and cellulases of Butyrivibrio fibrisolvens H17c. FEMS Microbiol Lett 68(2):197–203
276. Bray MR, Clarke AJ (1992) Action pattern of xylo-oligosaccharide hydrolysis by Schizophyllum commune xylanase A. Eur J Biochem 204(1):191–196
277. Christakopoulos P, Nerinckx W, Kekos D, Macris B, Claeyssens M (1997) The alkaline xylanase III from Fusarium oxysporum F3 belongs to family F/10. Carbohydr Res 302(3–4):191–195
278. Debeire P, Priem B, Strecker G, Vignon M (1990) Purification and properties of an endo-1,4-xylanase excreted by a hydrolytic thermophilic anaerobe, Clostridium thermolacticum. A proposal for its action mechanism on larchwood 4-O-methylglucuronoxylan. Eur J Biochem 187(3):573–580
279. Suzuki T, Kitagawa E, Sakakibara F, Ibata K, Usui K, Kawai K (2001) Cloning, expression, and characterization of a family 52 beta-xylosidase gene (xysB) of a multiple-xylanase-producing bacterium, Aeromonas caviae ME-1. Biosci Biotechnol Biochem 65(3):487–494
280. Saluzzi L, Flint HJ, Stewart CS (2001) Adaptation of Ruminococcus flavefaciens resulting in increased degradation of ryegrass cell walls. FEMS Microbiol Ecol 36(2–3):131–137
281. La Grange DC, Pretorius IS, Claeyssens M, van Zyl WH (2001) Degradation of xylan to D-xylose by recombinant Saccharomyces cerevisiae coexpressing the Aspergillus niger beta-xylosidase (xlnD) and the Trichoderma reesei xylanase II (xyn2) genes. Appl Environ Microbiol 67(12):5512–5519
282. Ratanakhanokchai K, Kyu KL, Tanticharoen M (1999) Purification and properties of a xylan-binding endoxylanase from alkaliphilic bacillus sp. Strain K-1. Appl Environ Microbiol 65 (2):694–697

283. York WS, Kumar Kolli VS, Orlando R, Albersheim P, Darvill AG (1996) The structures of arabinoxyloglucans produced by solanaceous plants. Carbohydr Res 285:99–128
284. Vanzin GF, Madson M, Carpita NC, Raikhel NV, Keegstra K, Reiter WD (2002) The mur2 mutant of Arabidopsis thaliana lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. Proc Natl Acad Sci U S A 99(5):3340–3345
285. Dunand C, Gautier C, Chambat G, Lienart Y (2000) Characterization of the binding of alpha-L-Fuc (1–>2)-beta-D-Gal (1–>), a xyloglucan signal, in blackberry protoplasts. Plant Sci 151 (2):183–192
286. Watt DK, Brasch DJ, Larsen DS, Melton LD, Simpson J (1996) Oligosaccharides related to xyloglucan: synthesis and X-ray crystal structure of methyl 2-0-(alpha-L-fucopyranosyl)-beta-D-galactopyranoside. Carbohydr Res 285:1–15
287. Watt DK, Brasch DJ, Larsen DS, Melton LD, Simpson J (2000) Oligosaccharides related to xyloglucan: synthesis and X-ray crystal structure of methyl alpha-L-fucopyranosyl-(1–>2)-beta-D-galactopyranosyl-(1–>2)-alpha-D-xylopyranoside and the synthesis of methyl alpha-L-fucopyranosyl-(1–>2)-beta-D-galactopyranosyl-(1–>2)-beta-D-xylopyranoside. Carbohydr Res 325(4):300–312
288. Busato AP, Vargas-Rechia CG, Reicher F (2001) Xyloglucan from the leaves of Hymenaea courbaril. Phytochemistry 58 (3):525–531
289. Vierhuis E, York WS, Kolli VS, Vincken J, Schols HA, Van Alebeek GW, Voragen AG (2001) Structural analyses of two arabinose containing oligosaccharides derived from olive fruit xyloglucan: XXSG and XLSG. Carbohydr Res 332(3):285–297
290. Maruyama K, Goto C, Numata M, Suzuki T, Nakagawa Y et al (1996) O-acetylated xyloglucan in extracellular polysaccharides from cell-suspension cultures of Mentha. Phytochemistry 41 (5):1309–1314
291. Wu CT, Leubner-Metzger G, Meins F Jr, Bradford KJ (2001) Class I beta-1,3-glucanase and chitinase are expressed in the micropylar endosperm of tomato seeds prior to radicle emergence. Plant Physiol 126(3):1299–1313
292. Hrmova M, Fincher GB (2001) Structure-function relationships of beta-D-glucan endo- and exohydrolases from higher plants. Plant Mol Biol 47(1–2):73–91
293. Hu G, Rijkenberg FH (1998) Subcellular localization of beta-1,3-glucanase in Puccinia recondita f.sp. tritici-infected wheat leaves. Planta 204(3):324–334
294. Kotake T, Nakagawa N, Takeda K, Sakurai N (1997) Purification and characterization of wall-bound exo-1,3-beta-D-glucanase from barley (Hordeum vulgare L.) seedlings. Plant Cell Physiol 38(2):194–200
295. Brummell DA, Catala C, Lashbrook CC, Bennett AB (1997) A membrane-anchored E-type endo-1,4-beta-glucanase is localized on Golgi and plasma membranes of higher plants. Proc Natl Acad Sci U S A 94(9):4794–4799
296. Keitel T, Thomsen KK, Heinemann U (1993) Crystallization of barley (1-3,1-4)-beta-glucanase, isoenzyme II. J Mol Biol 232 (3):1003–1004
297. Wang G, Marquardt RR, Xiao H, Zhang Z (1999) Development of a 96-well enzyme-linked solid-phase assay for beta- glucanase and xylanase. J Agric Food Chem 47(3):1262–1267
298. Mestechkina NM, Anulov OV, Smirnova NI, Shcherbukhin VD (2000) Composition and structure of a galactomannan macromolecule from seeds of Astragalus lehmannianus Bunge. Appl Biochem Microbiol 36(5):502–506
299. Chaubey M, Kapoor VP (2001) Structure of a galactomannan from the seeds of Cassia angustifolia Vahl. Carbohydr Res 332 (4):439–444
300. Teramoto A, Fuchigami M (2000) Changes in temperature, texture, and structure of konnyaku (konjac glucomannan gel) during high-pressure-freezing. J Food Sci 65(3):491–497
301. Zhang H, Yoshimura M, Nishinari K, Williams MAK, Foster TJ, Norton IT (2001) Gelation behaviour of konjac glucomannan with different molecular weights. Biopolymers 59(1):38–50
302. Zhang HS, Yang JG, Zhao Y (2002) The glucomannan from ramie. Carbohydr Polym 47(1):83–86
303. Smirnova NI, Mestechkina NM, Shcherbukhin VD (2001) The structure and characteristics of glucomannans from Eremurus iae and E-zangezuricus: assignment of acetyl group localization in macromolecules. Appl Biochem Microbiol 37(3):287–291
304. Petkowicz CLD, Reicher F, Chanzy H, Taravel FR, Vuong R (2001) Linear mannan in the endosperm of Schizolobium amazonicum. Carbohydr Polym 44(2):107–112
305. Parker KN, Chhabra SR, Lam D, Callen W, Duffaud GD et al (2001) Galactomannanases Man2 and Man5 from Thermotoga species: growth physiology on galactomannans, gene sequence analysis, and biochemical properties of recombinant enzymes. Biotechnol Bioeng 75(3):322–333
306. Parker KN, Chhabra S, Lam D, Snead MA, Mathur EJ, Kelly RM (2001) Beta-mannosidase from Thermotoga species. Methods Enzymol 330:238–246
307. Chhabra S, Parker KN, Lam D, Callen W, Snead MA, Mathur EJ et al (2001) Beta-mannanases from Thermotoga species. Methods Enzymol 330:224–238
308. Ganter JL, Sabbi JC, Reed WF (2001) Real-time monitoring of enzymatic hydrolysis of galactomannans. Biopolymers 59 (4):226–242
309. Ademark P, de Vries RP, Hagglund P, Stalbrand H, Visser J (2001) Cloning and characterization of Aspergillus niger genes encoding an alpha-galactosidase and a beta-mannosidase involved in galactomannan degradation. Eur J Biochem 268 (10):2982–2990
310. Tenkanen M, Puls J, Ratto M, Viikari L (1993) Enzymatic deacetylation of galactoglucomannans. Appl Microbiol Biotechnol 39(2):159–165
311. Tenkanen M, Thornton J, Viikari L (1995) An acetylglucomannan esterase of Aspergillus-oryzae-purification, characterization and role in the hydrolysis of O-acetyl-galactoglucomannan. J Biotechnol 42(3):197–206
312. Berens S, Kaspari H, Klemme JH (1996) Purification and characterization of two different xylanases from the thermophilic actinomycete Microtetraspora flexuosa SIIX. Antonie Van Leeuwenhoek 69(3):235–241
313. Bolam DN, Hughes N, Virden R, Lakey JH, Hazlewood GP et al (1996) Mannanase A from Pseudomonas fluorescens ssp. cellulosa is a retaining glycosyl hydrolase in which E212 and E320 are the putative catalytic residues. Biochemistry 35(50):16195–16204
314. Ghangas GS, Hu YJ, Wilson DB (1989) Cloning of a Thermomonospora fusca xylanase gene and its expression in Escherichia coli and Streptomyces lividans. J Bacteriol 171 (6):2963–2969
315. McKie VA, Vincken JP, Voragen AG, van den Broek LA, Stimson E, Gilbert HJ (2001) A new family of rhamnogalacturonan lyases contains an enzyme that binds to cellulose. Biochem J 355 (pt 1):167–177
316. Braithwaite KL, Black GW, Hazlewood GP, Ali BR, Gilbert HJ (1995) A non-modular endo-beta-1,4-mannanase from Pseudomonas fluorescens subspecies cellulosa. Biochem J 305(pt 3):1005–1010
317. Sipat A, Taylor KA, Lo RY, Forsberg CW, Krell PJ (1987) Molecular cloning of a xylanase gene from Bacteroides succinogenes and its expression in Escherichia coli. Appl Environ Microbiol 53(3):477–481

318. Takahashi R, Mizumoto K, Tajika K, Takano R (1992) Production of oligosaccharides from hemicellulose of woody biomass by enzymatic-hydrolysis.1. A simple method for isolating beta-D-mannanase-producing microorganisms. Mokuzai Gakkaishi 38 (12):1126–1135
319. Whitehead TR, Hespell RB (1989) Cloning and expression in Escherichia coli of a xylanase gene from Bacteroides ruminicola 23. Appl Environ Microbiol 55(4):893–896
320. Markovic O, Mislovicov AD, Biely P, Heinrichov K (1992) Chromogenic substrate for endo-polygalacturonase detection in gels. J Chromatogr A 603(1–2):243–246
321. Zantinge JL, Huang HC, Cheng KJ (2002) Microplate diffusion assay for screening of beta-glucanase-producing microorganisms. Biotechniques 33(4):798–ff
322. Skjot M, Kauppinen S, Kofod LV, Fuglsang C, Pauly M et al (2001) Functional cloning of an endo-arabinanase from Aspergillus aculeatus and its heterologous expression in A-oryzae and tobacco. Mol Genet Genomics 265(5):913–921
323. Zantinge JL, Huang HC, Cheng KJ (2002) Microplate diffusion assay for screening of beta-glucanase-producing microorganisms. Biotechniques 33(4):798, 800, 802 passim
324. Ridley BL, O'Neill MA, Mohnen D (2001) Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. Phytochemistry 57(6):929–967
325. de Vries RP, Visser J (2001) Aspergillus enzymes involved in degradation of plant cell wall polysaccharides. Microbiol Mol Biol Rev 65(4):497–522, table of contents
326. Cardoso SM, Silva AM, Coimbra MA (2002) Structural characterisation of the olive pomace pectic polysaccharide arabinan side chains. Carbohydr Res 337(10):917–924
327. McCartney L, Ormerod AP, Gidley MJ, Knox JP (2000) Temporal and spatial regulation of pectic (1–>4)-beta-D-galactan in cell walls of developing pea cotyledons: implications for mechanical properties. Plant J 22(2):105–113
328. Willats WG, Orfila C, Limberg G, Buchholt HC, van Alebeek GJ et al (2001) Modulation of the degree and pattern of methyl-esterification of pectic homogalacturonan in plant cell walls. Implications for pectin methyl esterase action, matrix properties, and cell adhesion. J Biol Chem 276(22):19404–19413
329. Fransen CT, Haseley SR, Huisman MM, Schols HA, Voragen AG et al (2000) Studies on the structure of a lithium-treated soybean pectin: characteristics of the fragments and determination of the carbohydrate substituents of galacturonic acid. Carbohydr Res 328 (4):539–547
330. Mazeau K, Perez S (1998) The preferred conformations of the four oligomeric fragments of Rhamnogalacturonan II. Carbohydr Res 311(4):203–217
331. Gainvors A, Nedjaoum N, Gognies S, Muzart M, Nedjma M, Belarbi A (2000) Purification and characterization of acidic endo-polygalacturonase encoded by the PGLI-1 gene from Saccharomyces cerevisiae. FEMS Microbiol Lett 183(1):131–135
332. Kennedy JF, Methacanon P (1999) Enzymes for carbohydrate engineering. In: Park K-H, Robyt JF, Choi Y-D (eds) Carbohydrate polymers, vol 39(3), p 292. Elsevier, Amsterdam, The Netherlands, 1996, pp Vii + 215
333. Mutter M, Renard CM, Beldman G, Schols HA, Voragen AG (1998) Mode of action of RG-hydrolase and RG-lyase toward rhamnogalacturonan oligomers. Characterization of degradation products using RG-rhamnohydrolase and RG-galacturonohydrolase. Carbohydr Res 311(3):155–164
334. Antov MG, Pericin DM, Dimic GR (2001) Cultivation of Polyporus squamosus for pectinase production in aqueous two-phase system containing sugar beet extraction waste. J Biotechnol 91(1):83–87
335. Bhattacharya S, Rastogi NK (1998) Rheological properties of enzyme-treated mango pulp. J Food Eng 36(3):249–262
336. Castilho LR, Alves TLM, Medronho RA (1999) Recovery of pectolytic enzymes produced by solid state culture of Aspergillus niger. Process Biochem 34(2):181–186
337. Fanta N, Quaas A, Zulueta P, Rez LM (1992) Release of reducing sugars from Citrus seedlings, leaves and fruits. Effect of treatment with pectinase and cellulase from Alternaria and Trichoderma. Phytochemistry 31(10):3359–3364
338. Kapoor M, Khalil Beg Q, Bhushan B et al (2000) Production and partial purification and characterization of a thermo-alkali stable polygalacturonase from Bacillus sp. MG-cp-2. Process Biochem 36(5):467–473
339. Hadj-Taieb N, Ayadi M, Trigui S, Bouabdallah F, Gargouri A (2002) Hyperproduction of pectinase activities by a fully constitutive mutant (CT1) of Penicillium occitanis. Enzyme Microb Technol 30(5):662–666
340. Weissermel K, Arpe H-J (1997) Industrial organic chemistry, 3rd edn. VCH, New York
341. Murphy DJ (1994) Designer oil crops: breeding, processing and biotechnology. VCH, Weinheim
342. Robbelen G, Downey RK, Ashri A (1989) Oil crops of the world: their breeding and utilization. McGraw-Hill, New York
343. Scott DS, Piskorz J, Radlein D (1994) Production of levoglucosan as an industrial chemical. In: Witczak ZJ (ed) Levoglucosenone and levoglucosans. Chemistry and applications. ATL Press Science Publishers, Mt. Prospect, IL, pp 179–188
344. Eggersdorfer M, Meyer J, Eckes P (1992) Use of renewable resources for nonfood materials. FEMS Microbiol Rev 103 (2–4):355–364
345. Donaldson TL, Culberson OL (1984) An industry model of commodity chemicals from renewable resources. Energy 9 (8):693–707
346. Hanselmann KW (1982) Lignochemicals. Experientia 38 (2):176–188
347. Lipinsky ES (1981) Chemicals from biomass—petrochemical substitution options. Science 212(4502):1465–1471
348. Indergaard M, Johansson A, Crawford B (1989) Biomass technologies. Chimia 43(7–8):230–232
349. Kirk–Othmer encyclopedia of chemical technology, vol 12, 4th edn. Wiley, New York, 1994
350. Goldstein IS (1981) Chemicals from biomass—present status. Forest Prod 31(10):63–68
351. Goldstein IS (1981) Organic chemicals from biomass. CRC Press, Boca Raton, FL
352. Kuhad RC, Singh A (1993) Lignocellulose biotechnology—current and future-prospects. Crit Rev Biotechnol 13(2):151–172
353. Kerr RA (1998) The next oil crisis looms large—and perhaps close. Science 281(5380):1128–1131
354. Campbell CJ, Laherrere JH (1998) Preventing the next oil crunch—the end of cheap oil. Sci Am 278(3):77–83
355. U. S. Department of Energy (1997) Energy information administration, annual energy review, vol DOE/EIA-0384(97). U. S. Department of Energy, Washington, DC
356. Holderich WF, Roseler J, Heitmann G, Liebens AT (1997) The use of zeolites in the synthesis of fine and intermediate chemicals. Catal Today 37(4):353–366
357. Donnini GP, Blain TJ, Holton HH, Kutney GW (1983) 300 Alkaline pulping additives—structure-activity-relationships. Pulp Pap Canada 84(11):R134–R140
358. McMillan JD (1994) Pretreatment of lignocellulosic biomass. In: Himmel ME, Baker JO, Overend RP (eds) Enzymatic conversion of biomass for fuels production, vol 566. ACS, Washington, DC, pp 292–324
359. Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, Ladisch M (2005) Features of promising technologies for pretreatment of lignocellulosic biomass. Bioresour Technol 96 (6):673–686

360. Torget R, Walter P, Himmel M, Grohmann K (1991) Dilute-acid pretreatment of corn residues and short-rotation woody crops. Appl Biochem Biotechnol 28–9:75–86
361. Torget R, Werdene P, Himmel M, Grohmann K (1990) Dilute acid pretreatment of short rotation woody and herbaceous crops. Appl Biochem Biotechnol 24–5:115–126
362. Singh A, Das K, Sharma DK (1984) Production of xylose, furfural, fermentable sugars and ethanol from agricultural residues. J Chem Technol Biotechnol A Chem Technol 34(2):51–61
363. Singh A, Das K, Sharma DK (1984) Integrated process for production of xylose, furfural, and glucose from bagasse by 2-step acid-hydrolysis. Ind Eng Chem Prod Res Dev 23(2):257–262
364. Vila C, Santos V, Parajo JC (2003) Recovery of lignin and furfural from acetic acid-water-HCl pulping liquors. Bioresour Technol 90 (3):339–344
365. Lehnen R, Saake B, Nimz HH (2001) Furfural and hydroxymethyl furfural as by-products of FORMACELL pulping. Holzforschung 55(2):199–204
366. Lignosulfonates. In: Chemical economics handbook. SRI Consulting, Menlo Park, CA
367. Ward OP, Singh A (2002) Bioethanol technology: developments and perspectives. In: Laskin AI, Bennett JW, Gadd GM (eds) Advances in applied microbiology, vol 51. Academic, Amsterdam, pp 53–80
368. Sun Y, Cheng JY (2002) Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresour Technol 83(1):1–11
369. Galbe M, Zacchi G (2002) A review of the production of ethanol from softwood. Appl Microbiol Biotechnol 59(6):618–628
370. Nguyen QA, Tucker MP, Keller FA, Eddy FP (2000) Two-stage dilute-acid pretreatment of softwoods. Appl Biochem Biotechnol 84–6:561–576
371. Nguyen QA, Tucker MP, Keller FA, Beaty DA, Connors KM, Eddy FP (1999) Dilute acid hydrolysis of softwoods. Appl Biochem Biotechnol 77–9:133–142
372. Garrote G, Dominguez H, Parajo JC (1999) Hydrothermal processing of lignocellulosic materials. Holz Roh Werkst 57 (3):191–202
373. Schell DJ, Torget R, Power A, Walter PJ, Grohmann K, Hinman ND (1991) A technical and economic-analysis of acid-catalyzed steam explosion and dilute sulfuric-acid pretreatments using wheat straw or aspen wood chips. Appl Biochem Biotechnol 28–9:87–97
374. Aziz S, Sarkanen K (1989) Organosolv pulping—a review. Tappi J 72(3):169–175
375. Johansson A, Aaltonen O, Ylinen P (1987) Organosolv pulping—methods and pulp properties. Biomass 13(1):45–65
376. Kleinert TN (1974) Organosolvent pulping with aqueous alcohol. Tappi 57(8):99–102
377. Lora JH, Aziz S (1985) Organosolv pulping—a versatile approach to wood refining. Tappi J 68(8):94–97
378. McDonough TJ (1993) The chemistry of organosolv delignification. Tappi J 76(8):186–193
379. Sarkanen KV (1990) Chemistry of solvent pulping. Tappi J 73 (10):215–219
380. Stockburger P (1993) An overview of near-commercial and commercial solvent-based pulping processes. Tappi J 76(6):71–74
381. Kishimoto T, Sano Y (2001) Delignification mechanism during high-boiling solvent pulping Part 1. Reaction of guaiacylglycerol-beta-guaiacyl ether. Holzforschung 55(6):611–616
382. Thring RW, Chornet E, Overend RP (1993) Thermolysis of glycol lignin in the presence of tetralin. Canad J Chem Eng 71(1):107–115
383. Demirbas A (1998) Aqueous glycerol delignification of wood chips and ground wood. Bioresour Technol 63(2):179–185
384. Jimenez L, de la Torre MJ, Maestre F, Ferrer JL, Perez I (1997) Organosolv pulping of wheat straw by use of phenol. Bioresour Technol 60(3):199–205
385. Nimz HH, Berg A, Granzow C, Casten R, Muladi S (1989) Pulping and bleaching by the acetosolv process. Papier 43(10A): V102–V108
386. Nimz HH, Granzow C, Berg A (1986) Acetosolv pulping. Holz Roh Werkst 44(9):362
387. Parajo JC, Alonso JL, Vazquez D, Santos V (1993) Optimization of catalyzed acetosolv fractionation of pine wood. Holzforschung 47(3):188–196
388. Saake B, Lummitsch S, Mormanee R, Lehnen R, Nimz HH (1995) production of pulps using the Formacell process. Papier 49(10A): V1–V7
389. Sundquist J (1996) Chemical pulping based on formic acid—summary of Milox research. Pap Puu Paper Tim 78(3):92–95
390. Poppius-Levlin K, Mustonen R, Huovila T, Sundquist J (1991) Milox pulping with acetic-acid peroxyacetic acid. Pap Puu Paper Tim 73(2):154–158
391. Dahlmann G, Schroeter MC (1990) The organocell process—pulping with the environment in mind. Tappi J 73(4):237–240
392. Pye EK, Lora JH (1991) The Alcell process—a proven alternative to kraft pulping. Tappi J 74(3):113–118
393. Schroeter MC (1991) Possible lignin reactions in the organocell pulping process. Tappi J 74(10):197–200
394. Black NP (1991) ASAM alkaline sulfite pulping process shows potential for large-scale application. Tappi J 74(4):87–93
395. Kirci H, Bostanci S, Yalinkilic MK (1994) A new modified pulping process alternative to sulfate method alkali-sulfite-antraquinone-ethanol (ASAE). Wood Sci Technol 28(2):89–99
396. Patt R, Knoblauch J, Faix O, Kordsachia O, Puls J (1991) Lignin and carbohydrate reactions in alkaline sulfite, anthraquinone, methanol (ASAM) pulping. Papier 45(7):389–396
397. Schubert HL, Fuchs K, Patt R, Kordsachia O, Bobik M (1993) The ASAM-process—a pulp technology ready for industry—experiences gained through 3 years operation of the pilot-plant. Papier 47(10A):V6–V15
398. Yawalata D, Paszner L (2004) Cationic effect in high concentration alcohol organosolv pulping: the next generation biorefinery. Holzforschung 58(1):7–13
399. Oliet M, Garcia J, Rodriguez F, Gilarrranz MA (2002) Solvent effects in autocatalyzed alcohol-water pulping comparative study between ethanol and methanol as delignifying agents. Chem Eng J 87(2):157–162
400. Goyal GC, Lora JH, Pye EK (1992) Autocatalyzed organosolv pulping of hardwoods—effect of pulping conditions on pulp properties and characteristics of soluble and residual lignin. Tappi J 75(2):110–116
401. Oliet M, Rodriguez F, Garcia J, Gilarranz MA (2001) The effect of autocatalyzed ethanol pulping on lignin characteristics. J Wood Chem Technol 21(1):81–95
402. Yawalata D, Paszner L (2004) Anionic effect in high concentration alcohol organosolv pulping. Holzforschung 58(1):1–6
403. Paszner L, Cho HJ (1989) Organosolv pulping—acidic catalysis options and their effect on fiber quality and delignification. Tappi J 72(2):135–142
404. Paszner L, Behera NC (1989) Topochemistry of softwood delignification by alkali earth metal salt catalyzed organosolv pulping. Holzforschung 43(3):159–168
405. Black SK, Hames BR, Myers MD (1998) US 5,730,837
406. Ibrahim M, Glasser WG (1999) Steam-assisted biomass fractionation. Part III: a quantitative evaluation of the “clean fractionation” concept. Bioresour Technol 70(2):181–192
407. Avellar BK, Glasser WG (1998) Steam-assisted biomass fractionation. I. Process considerations and economic evaluation. Biomass Bioenergy 14(3):205–218
408. Heitz M, Capekmenard E, Koeberle PG, Gagne J, Chornet E et al (1991) Fractionation of populus-tremuloides at the pilot-plant

scale—optimization of steam pretreatment conditions using the stake-Ii technology. Bioresour Technol 35(1):23–32

409. Schultz TP, Blermann CJ, McGinnis GD (1983) Steam explosion of mixed hardwood chips as a biomass pretreatment. Ind Eng Chem Product Res Dev 22(2):344–348
410. Sun XF, Xu F, Sun RC, Wang YX, Fowler P, Baird MS (2004) Characteristics of degraded lignins obtained from steam exploded wheat straw. Polym Degrad Stab 86(2):245–256
411. Shevchenko SM, Chang K, Dick DG, Gregg DJ, Saddler JN (2001) Structure and properties of lignin in softwoods after SO2-catalyzed steam explosion and enzymatic hydrolysis. Cell Chem Technol 35(5–6):487–502
412. Shevchenko SM, Beatson RP, Saddler JN (1999) The nature of lignin from steam explosion enzymatic hydrolysis of softwood—structural features and possible uses. Appl Biochem Biotechnol 77–9:867–876
413. Ramos LP, Mathias AL, Silva FT, Cotrim AR, Ferraz AL, Chen CL (1999) Characterization of residual lignin after SO2-catalyzed steam explosion and enzymatic hydrolysis of Eucalyptus viminalis wood chips. J Agricult Food Chern 47(6):2295–2302
414. Fernandez-Bolanos J, Felizon B, Heredia A, Guillen R, Jimenez A (1999) Characterization of the lignin obtained by alkaline delignification and of the cellulose residue from steam-exploded olive stones. Bioresour Technol 68(2):121–132
415. Emmel A, Mathias AL, Wypych F, Ramos LP (2003) Fractionation of Eucalyptus grandis chips by dilute acid-catalysed steam explosion. Bioresour Technol 86(2):105–115
416. Schell D, Nguyen Q, Tucker M, Boynton B (1998) Pretreatment of softwood by acid-catalyzed steam explosion followed by alkali extraction. Appl Biochem Biotechnol 70–2:17–24
417. Bura R, Bothast RJ, Mansfield SD, Saddler JN (2003) Optimization of SO2-catalyzed steam pretreatment of corn fiber for ethanol production. Appl Biochem Biotechnol 105:319–335
418. Bura R, Mansfield SD, Saddler JN, Bothast RJ (2002) SO2-catalyzed steam explosion of com fiber for ethanol production. Appl Biochem Biotechnol 98:59–72
419. McDonald AG, Clark TA (1992) Characterization of oligosaccharides released by steam explosion of sulfurdioxide impregnated pinus-radiata. J Wood Chem Technol 12(1):53–78
420. Gellerstedt G, Lindfors EL (1984) Structural-changes in lignin during kraft pulping. Holzforschung 38(3):151–158
421. Gierer J (1980) Chemical aspects of kraft pulping. Wood Sci Technol 14(4):241–266
422. Kleppe PJ (1970) Kraft pulping. Tappi 53(1):35ff
423. Holton H (1977) Soda additive softwood pulping—major new process. Pulp Pap Canada 78(10):19–24
424. Holton HH, Chapman FL (1977) Kraft pulping with anthraquinone—laboratory and full-scale mill trials. Tappi 60(11):121–125
425. Dimmel DR (1985) Electron-transfer reactions in pulping systems. 1. Theory and applicability to anthraquinone pulping. J Wood Chem Technol 5(1):1–14
426. Dimmel DR, Perry LF, Palasz PD, Chum HL (1985) Electron-transfer reactions in pulping systems. 2. Electrochemistry of anthraquinone lignin model quinone methides. J Wood Chem Technol 5(1):15–36
427. Saeman JF (1945) Ind Eng Chem 37:43
428. Mosier NS, Sarikaya A, Ladisch CM, Ladisch MR (2001) Characterization of dicarboxylic acids for cellulose hydrolysis. Biotechnol Prog 17(3):474–480
429. Torget RW, Kim JS, Lee YY (2000) Fundamental aspects of dilute acid hydrolysis/fractionation kinetics of hardwood carbohydrates. 1. Cellulose hydrolysis. Ind Eng Chem Res 39(8):2817–2825
430. Witczak ZJ (ed) (1994) Levoglucosenone and levoglucosans. Chemistry and applications. ATL, Mt. Prospect, IL
431. Trahanovsky WS, Ochaoda JM, Wang C, Revell KD, Arvidson KB et al (2003) A convenient procedure for the preparation of levoglucosenone and its conversion to novel chiral derivatives. In: Carbohydrate synthons in natural products chemistry: synthesis, functionalization and applications, vol 841, pp 21–31
432. Shibagaki M, Takahashi K, Kuno H, Honda I, Matsushita H (1990) Synthesis of levoglucosenone. Chem Lett 2:307–310
433. Fung DPC (1976) Further investigation on effect of H3p04 on pyrolysis of cellulose. Wood Sci 9(1):55–57
434. Halpern Y, Riffer R, Broido A (1973) Levoglucosenone (1,6-anhydro-3,4-dideoxy-delta3-beta-Dpyranosen-2-one)—major product ofacid-catalyzed pyrolysis of cellulose and related carbohydrates. J Org Chem 38(2):204–209
435. Shafizadeh F, Furneaux RH, Cochran TG, Scholl JP, Sakai Y (1979) Production of levoglucosan and glucose from pyrolysis of cellulosic materials. J Appl Polym Sci 23(12):3525–3539
436. Dobele G, Dizhbite T, Rossinskaja G, Telysheva G, Mier D, Radtke S, Faix O (2003) Pre-treatment of biomass with phosphoric acid prior to fast pyrolysis—a promising method for obtaining 1,6-anhydrosaccharides in high yields. Anal Appl Pyrol 68–9:197–211
437. Dobele G, Rossinskaja G, Telysheva G, Meier D, Faix O (1999) Cellulose dehydration and depolymerization reactions during pyrolysis in the presence of phosphoric acid. J Anal Appl Pyrol 49(1–2):307–317
438. Edye LA, Richards GN, Zheng G (1993) Transition-metals as catalysts for pyrolysis and gasification of biomass. ACS Symp Ser 515:90–101
439. Kleen M, Gellerstedt G (1995) Influence of inorganic species on the formation of polysaccharide and lignin degradation products in the analytical pyrolysis of pulps. J Anal Appl Pyrol 35(1):15–41
440. Richards GN, Zheng GC (1991) Influence of metal-ions and of salts on products from pyrolysis of wood applications to thermochemical processing of newsprint and biomass. J Anal Appl Pyrol 21(1–2):133–146
441. Denooy AEJ, Besemer AC, Vanbekkum H (1995) Selective oxidation of primary alcohols mediated by nitroxyl radical in aqueous-solution—kinetics and mechanism. Tetrahedron 51 (29):8023–8032
442. Denooy AEJ, Besemer AC, Vanbekkum H (1995) Highly selective nitroxyl radical-mediated oxidation of primary alcohol groups in water-soluble glucans. Carbohydr Res 269(1):89–98
443. Bragd PL, van Bekkum H, Besemer AC (2004) TEMPO-mediated oxidation of polysaccharides: survey of methods and applications. Top Catal 27(1–4):49–66
444. Kato Y, Matsuo R, Isogai A (2003) Oxidation process of water-soluble starch in TEMPO-mediated system. Carbohydr Polym 51 (1):69–75
445. Saito T, Isogai A (2004) TEMPO-mediated oxidation of native cellulose. The effect of oxidation conditions on chemical and crystal structures of the water-insoluble fractions. Biomacromolecules 5(5):1983–1989
446. Perez DD, Montanari S, Vignon MR (2003) TEMPO-mediated oxidation of cellulose III. Biomacromolecules 4(5):1417–1425
447. Tahiri C, Vignon MR (2000) TEMPO-oxidation of cellulose: synthesis and characterisation of polyglucuronans. Cellulose 7 (2):177–188
448. Gomez-Bujedo S, Fleury E, Vignon MR (2004) Preparation of cellouronic acids and partially acetylated cellouronic acids by TEMPO/NaCIO oxidation of water-soluble cellulose acetate. Biomacromolecules 5(2):565–571
449. Montanari S, Rountani M, Heux L, Vignon MR (2005) Topochemistry of carboxylated cellulose nanocrystals resulting from TEMPO-mediated oxidation. Macromolecules 38 (5):1665–1671
450. Kato Y, Kaminaga J, Matsuo R, Isogai A (2004) TEMPO-mediated oxidation of chitin, regenerated chitin and N-acetylated chitosan. Carbohydr Polym 58(4):421–426

451. Koga T, Taniguchi I (2004) Electrochemical oxidation of glucose to glucarate using TEMPO as a mediator in an alkaline solution. Electrochemistry 72(12):858–860
452. Schamann M, Schafer HJ (2003) TEMPO-mediated anodic oxidation of methyl glycosides and l-methyl and l-azido disaccharides. Eur Org Chem 2:351–358
453. Thaburet JF, Merbouh N, Ibert M, Marsais F, Queguiner G (2001) TEMPO-mediated oxidation of maltodextrins and D-glucose: effect of pH on the selectivity and sequestering ability of the resulting polycarboxylates. Carbohydr Res 330(1):21–29
454. Merbouh N, Thaburet JF, Ibert M, Marsais F, Bobbitt JM (2001) Facile nitroxide-mediated oxidations of D-glucose to D-glucaric acid. Carbohydr Res 336(1):75–78
455. Bozell JJ (ed) (2001) Chemicals and materials from renewable resources, vol 784. American Chemical Society, Washington, DC, p 64
456. Trombotto S, Violet-Courtens E, Cottier L, Queneau Y (2004) Oxidation of two major disaccharides: sucrose and isomaltulose. Top Catal 27(1–4):31–37
457. Antal MJ, Leesomboon T, Mok WS, Richards GN (1991) Kinetic-studies of the reactions of ketoses and aldoses in water at high-temperature. 3. Mechanism of formation of 2-furaldehyde from D-xylose. Carbohydr Res 217:71–85
458. Ahmad T, Kenne L, Olsson K, Theander O (1995) The formation of 2-furaldehyde and formic-acid from pentoses in slightly acidic deuterium-oxide studied by H-I-Nmr spectroscopy. Carbohydr Res 276(2):309–320
459. Oefner PJ, Lanziner AH, Bonn G, Bobleter O (1992) Quantitative studies on furfural and organic-acid formation during hydrothermal, acidic and alkaline-degradation of deuterium-xylose. Monatsh Chem 123(6–7):547–556
460. Montane D, Salvado J, Torras C, Farriol X (2002) High-temperature dilute-acid hydrolysis of olive stones for furfural production. Biomass Bioenerg 22(4):295–304
461. Mansilla HD, Baeza J, Urzua S, Maturana G, Villasenor J, Duran N (1998) Acid-catalysed hydrolysis of rice hull: evaluation of furfural production. Bioresour Technol 66(3):189–193
462. Carrasco E (1993) Production of furfural by dilute-acid hydrolysis of wood—methods for calculating furfural yield. Wood Fiber Sci 25(1):91–102
463. Riera EA, Alvarez R, Coca J (1991) Production of furfural by acid-hydrolysis of corncobs. J Chem Technol Biotechnol 50 (2):149–155
464. Zeitsch KJ (2000) Furfural production needs chemical innovation. Chem Innov 30(4):29–32
465. Basta AH, EI-Saied H (2003) Furfural production and kinetics of pentosans hydrolysis in corn cobs. Cell Chem Technol 37 (1–2):79–94
466. Zeitsch KJ (2001) Gaseous acid catalysis: an intriguing new process. Chem Innov 31(1):41–44
467. Dias AS, Pillinger M, Valente AA (2005) Dehydration of xylose into furfural over micro-mesoporous sulfonic acid catalysts. J Catal 229(2):414–423
468. Gamse T, Marr R, Froschl E, Siebenhofer M (1997) Extraction of furfural with carbon dioxide. Separation Sci Technol 32 (1–4):355–371
469. Sako T, Sugeta T, Nakazawa N, Okubo T, Sato M et al (1992) Kinetic study of furfural formation accompanying supercritical carbon-dioxide extraction. J Chem Eng Jpn 25(4):372–377
470. Kim YC, Lee HS (2001) Selective synthesis of furfural from xylose with supercritical carbon dioxide and solid acid catalyst. J Ind Eng Chem 7(6):424–429
471. Moreau C, Durand R, Peyron D, Duhamet J, Rivalier P (1998) Selective preparation of furfural from xylose over microporous solid acid catalysts. Ind Crops Prod 7(2–3):95–99
472. Schraufnagel RA, Rase HF (1975) Levulinic acid from sucrose using acidic ion-exchange resins. Ind Eng Chem Prod Res Dev 14 (1):40–44
473. Bozell JJ, Moens L, Elliott DC, Wang Y, Neuenscwander GG et al (2000) Production of levulinic acid and use as a platform chemical for derived products. Resour Conserv Recycling 28(3–4):227–239
474. Horvat J, Klaic B, Metelko B, Sunjic V (1985) Mechanism of levulinic acid formation. Tetrahedron Lett 26(17):2111–2114
475. Lourvanij K, Rorrer GL (1994) Dehydration of glucose to organic-acids in microporous pillared clay catalysts. Appl Catal A Gen 109 (1):147–165
476. Dahlmann J (1968) Hydrolytic method for the production of levulinic acid and its derivatives from biomass and sugars. Chem Ber 101:4251–4253
477. Farone WA, Cuzens JE (1999) Method for the production of levulinic acid. US Patent 5,892,107, 6 April 1999
478. Jow J, Rorrer GL, Hawley MC, Lamport DTA (1987) Dehydration of D-fructose to levulinic acid over Lzy zeolite catalyst. Biomass 14(3):185–194
479. Bozell JJ, Moens L, Elliott DC, Wang Y, Neuenscwander GG, Fitzpatrick SW, Bilski RJ, Jamefeld JL (2000) Production of levulinic acid and use as a platform chemical for derived products. Resour Conserv Recycling 28:227–239
480. Fitzpatrick SW (2006) The biofine technology A “bio-refinery” concept based on thermochemical conversion of cellulosic biomass. Feedstocks for the future: renewables for the production of chemicals and materials. ACS Symp Ser 921:271–287
481. Kuster BFM (1990) 5-Hydroxymethylfurfural (HMF)—a review focusing on its manufacture. Starch-Starke 42(8):314–321
482. Moreau C, Belgacem MN, Gandini A (2004) Recent catalytic advances in the chemistry of substituted furans from carbohydrates and in the ensuing polymers. Top Catal 27 (1–4):11–30
483. Kroger M, Prusse U, Vorlop KD (2000) A new approach for the production of 2,5-furandicarboxylic acid by in situ oxidation of 5-hydroxymethylfurfural starting from fructose. Top Catal 13 (3):237–242
484. Seri K, Sakaki T, Shibata M, Inoue Y, Ishida H (2002) Lanthanum (III)-catalyzed degradation of cellulose at 250 degrees C. Bioresour Technol 81(3):257–260
485. Seri K, Inoue Y, Ishida H (2001) Catalytic activity of lanthanide (III) ions for the dehydration of hexose to 5hydroxymethyl-2-furaldehyde in water. Bull Chem Soc Jpn 74(6):1145–1150
486. Ishida H, Seri K (1996) Catalytic activity of lanthanoide(III) ions for dehydration of D-glucose to 5-(hydroxymethyl)furfural. J Mol Catal A Chem 112(2):L163–L165
487. Seri K, Inoue Y, Ishida H (2000) Highly efficient catalytic activity of lanthanide(III) ions for conversion of saccharides to 5-hydroxymethyl-2-furfural in organic solvents. Chem Lett 1:22–23
488. Vinke P, Vanbekkum H (1992) The dehydration of fructose towards 5-hydroxymethylfurfural using activated carbon as adsorbent. Starch-Starke 44(3):90–96
489. Vandam HE, Kieboom APG, Vanbekkum H (1986) The conversion of fructose and glucose in acidic media formation of hydroxymethylfurfural. Starch-Starke 38(3):95–101
490. Kuster BFM, Vanderbaan HS (1977) Dehydration of D-fructose (formation of 5-hydroxymethyl-2-furaldehyde and levulinic acid). 2. Influence of initial and catalyst concentrations on dehydration of D-fructose. Carbohydr Res 54(2):165–176
491. Kuster BFM, Temmink HMG (1977) Dehydration Of D-fructose (formation of 5-hydroxymethyl-2-furaldehyde and levulinic acid) 4 influence of PH and weak-acid anions on dehydration of D-fructose. Carbohydr Res 54(2):185–191
492. Rigal L, Gaset A, Gorrichon JP (1981) Selective conversion of D-fructose to 5-hydroxymethyl-2-furancarboxaldehyde using a

water-solvent-ion-exchange resin triphasic system. Ind Eng Chem Prod Res Dev 20(4):719–721

493. Elhajj T, Masroua A, Martin JC, Descotes G (1987) Synthesis of 5-hydroxymethylfuran-2-carboxaldehyde and its derivatives by acidic treatment of sugars on ion-exchange resins. Bull Soc Chim Fr 5:855–860

494. Mercadier D, Rigal L, Gaset A, Gorrichon JP (1981) Synthesis of 5-hydroxymethyl-2-furancarboxaldehyde catalyzed by cationic exchange resins. 1. Choice of the catalyst and the characteristics of the reaction medium. Chem Technol Biotechnol 31(8):489–496

495. Bicker M, Hirth J, Vogel H (2003) Dehydration of fructose to 5-hydroxymethylfurfural in sub-and supercritical acetone. Green Chem 5(2):280–284

496. Lansalot-Matras C, Moreau C (2003) Dehydration of fructose into 5-hydroxymethylfurfural in the presence of ionic liquids. Catal Commun 4(10):517–520

497. Carlini C, Patrono P, Galletti AMR, Sbrana G (2004) Heterogeneous catalysts based on vanadyl phosphate for fructose dehydration to 5-hydroxymethyl-2-furaldehyde. Appl Catal A Gen 275 (1–2):111–118

498. Armaroli T, Busca G, Carlini C, Giuttari M, Galletti AMR, Sbrana G (2000) Acid sites characterization of niobium phosphate catalysts and their activity in fructose dehydration to 5-hydroxymethyl-2-furaldehyde. J Mol Catal A Chem 151 (1–2):233–243

499. Carlini C, Giuttari M, Galletti AMR, Sbrana G, Armaroli T, Busca G (1999) Selective saccharides dehydration to5-hydroxymethyl-2-furaldehyde by heterogeneous niobium catalysts. Appl Catal A Gen 183(2):295–302

500. Moreau C, Durand R, Razigade S, Duhamet J, Faugeras P et al (1996) Dehydration of fructose to 5-hydroxymethylfurfural over H-mordenites. Appl Catal A Gen 145(1–2):211–224

501. Lourvanij K, Rorrer GL (1993) Reactions of aqueous glucose solutions over solid-acid Y-zeolite catalyst at 110–160 degrees-C. Ind Eng Chem Res 32(1):11–19

502. Koda K, Goto H, Shintani H, Matsumoto Y, Meshitsuka G (2001) Oxidative cleavage of lignin aromatics during chlorine bleaching of kraft pulp. J Wood Sci 47(5):362–367

503. Gaspar AR, Evtuguin DV, Neto CP (2004) Polyoxometalate-catalyzed oxygen delignification of kraft pulp: a pilot-plant experience. Ind Eng Chem Res 43(24):7754–7761

504. Gaspar A, Evtuguin DV, Neto CP (2004) Lignin reactions in oxygen delignification catalysed by Mn(II)-substituted molybdovanadophosphate polyanion. Holzforschung 58 (6):640–649

505. Crestini C, Pastorini A, Tagliatesta P (2004) Metalloporphyrins immobilized on motmorillonite as biomimetic catalysts in the oxidation of lignin model compounds. Mol Catal A Chem 208 (1–2):195–202

506. Gaspar A, Evtuguin DV, Neto CP (2003) Oxygen bleaching of kraft pulp catalysed by Mn(III)-substituted polyoxometalates. Appl Catal A Gen 239(1–2):157–168

507. Chen CL, Capanema EA, Gracz HS (2003) Reaction mechanisms in delignification of pine Kraft-AQ pulp with hydrogen peroxide using Mn(IV)-Me4DTNE as catalyst. J Agric Food Chem 51 (7):1932–1941

508. Chen CL, Capanema EA, Gracz HS (2003) Comparative studies on the delignification of pine kraft anthraquinone pulp with hydrogen peroxide by binucleus Mn(IV) complex catalysis. J Agric Food Chem 51(21):6223–6232

509. Alves V, Capanema E, Chen CL, Gratzl J (2003) Comparative studies on oxidation of lignin model compounds with hydrogen peroxide using Mn(IV)-Me(3)TACN and Mn(IV)-Me4DTNE as catalyst. J Mol Catal A Chem 206(1–2):37–51

510. Balakshin MY, Evtuguin DV, Neto CP, Cavaco-Paulo A (2001) Polyoxometalates as mediators in the laccase catalyzed delignification. J Mol Catal B Enzym 16(3–4):131–140

511. Evtuguin DV, Neto CP, Rocha J (2000) Lignin degradation in oxygen delignification catalysed by [PM07V5040](8-) polyanion. Part I. Study on wood lignin. Holzforschung 54(4):381–389

512. Cui Y, Puthson P, Chen CL, Gratzl JS, Kirkman AG (2000) Kinetic study on delignification of kraft-AQ pine pulp with hydrogen peroxide catalyzed by Mn(IV)-Me4DTNE. Holzforschung 54 (4):413–419

513. Crestini C, Saladino R, Tagliatesta P, Boschi T (1999) Biomimetic degradation of lignin and lignin model compounds by synthetic anionic and cationic water soluble manganese and iron porphyrins. Bioorgan Med Chem 7(9):1897–1905

514. Glasser WG, Northey RA, Schultz TP (1999) Lignin: historical, biological and material perspectives, vol 740. American Chemical Society, Washington, DC

515. Sippola V, Krause O, Vuorinen T (2004) Oxidation of lignin model compounds with cobalt-sulphosalen catalyst in the presence and absence of carbohydrate model compound. J Wood Chem Technol 24(4):323–340

516. Sippola VO, Krause AOI (2003) Oxidation activity and stability of homogeneous cobalt-sulphosalen catalyst Studies with a phenolic and a non-phenolic lignin model compound in aqueous alkaline medium. J Mol Catal A Chem 194(1–2):89–97

517. Xiang Q, Lee YY (2001) Production of oxychemicals from precipitated hardwood lignin. Appl Biochem Biotechnol 91–3:71–80

518. Villar JC, Caperos A, Garcia-Ochoa F (2001) Oxidation of hardwood kraft-lignin to phenolic derivatives with oxygen as oxidant. Wood Sci Technol 35(3):245–255

519. Embree HD, Chen TH, Payne GF (2001) Oxygenated aromatic compounds from renewable resources: motivation, opportunities, and adsorptive separations. Chem Eng J 84(2):133–147

520. Koehler JA, Brune BJ, Chen TH, Glemza AJ, Vishwanath P et al (2000) Potential approach for fractionating oxygenated aromatic compounds from renewable resources. Ind Eng Chem Res 39 (9):3347–3355

521. Bozell JJ, Hames BR, Dimmel DR (1995) Cobalt-Schiff base complex-catalyzed oxidation of parasubstituted phenolics—preparation of benzoquinones. J Org Chem 60(8):2398–2404

522. Bozell JJ, Hoberg JO, Dimmel DR (2000) Heteropolyacid catalyzed oxidation of lignin and lignin models to benzoquinones. J Wood Chem Technol 20(1):19–41

523. Dimmel DR, Althen E, Savidakis M, Courchene C, Bozell JJ (1999) New quinone-based pulping catalysts. Tappi J 82 (12):83–89

524. Bozell JJ, Hoberg JO, Dimmel DR (1998) Catalytic oxidation of para-substituted phenols with nitrogen dioxide and oxygen. Tetrahedron Lett 39(16):2261–2264

525. Shore SG, Ding E, Park C, Keane MA (2004) The application of {(DMF)(I 0)Yb-2[TM(CN)(4)](3) }(infinity) (TM = Ni, Pd) supported on silica to promote gas phase phenol hydrogenation. J Mol Catal A Chem 212(1–2):291–300

526. Shore SG, Ding ER, Park C, Keane MA (2002) Vapor phase hydrogenation of phenol over silica supported Pd and Pd-Yb catalysts. Catal Commun 3(2):77–84

527. Scire S, Minico S, Crisafulli C (2002) Selective hydrogenation of phenol to cyclohexanone over supported Pd and Pd-Ca catalysts: an investigation on the influence of different supports and Pd precursors. Appl Catal A Gen 235(1–2):21–31

528. Mahata N, Vishwanathan V (2000) Influence of palladium precursors on structural properties and phenol hydrogenation characteristics of supported palladium catalysts. J Catal 196 (2):262–270

529. Claus P, Berndt H, Mohr C, Radnik J, Shin EJ, Keane MA (2000) Pd/MgO: catalyst characterization and phenol hydrogenation activity. J Catal 192(1):88–97
530. Mahata N, Raghavan KV, Vishwanathan V (1999) Influence of alkali promotion on phenol hydrogenation activity of palladium alumina catalysts. Appl Catal A Gen 182(1):183–187
531. Mathias AL, Rodrigues AE (1995) Production of vanillin by oxidation of pine kraft lignins with oxygen. Holzforschung 49 (3):273–278
532. Wu GX, Heitz M, Chornet E (1994) Improved alkaline oxidation process for the production of aldehydes (vanillin and syringaldehyde) from steam-explosion Hardwood lignin. Ind Eng Chem Res 33(3):718–723
533. Tarabanko VE, Petukhov DV, Selyutin GE (2004) New mechanism for the catalytic oxidation of lignin to vanillin. Kinet Catal 45 (4):569–577
534. Fargues C, Mathias A, Silva J, Rodrigues A (1996) Kinetics of vanillin oxidation. Chem Eng Technol 19(2):127–136
535. Fargues C, Mathias A, Rodrigues A (1996) Kinetics of vanillin production from om Kraft lignin oxidation. Ind Eng Chem Res 35 (1):28–36
536. Wu GX, Heitz M (1995) Catalytic mechanism of Cu2+ and Fe3+ in alkaline 0–2 oxidation of lignin. J Wood Chem Technol 15 (2):189–202
537. Tarabanko VE, Fornova NA, Kuznetsov BN, Ivanchenko NM, Kudryashev AV (1995) On the mechanism of vanillin formation in the catalytic-oxidation of lignin with oxygen. React Kinet Catal Lett 55(1):161–170
538. McKendry P (2002) Energy production from biomass (part 1): overview of biomass. Bioresour Technol 83(1):37–46
539. McKendry P (2002) Energy production from biomass (part 2): conversion technologies. Bioresour Technol 83(1):47–54
540. Bridgwater AV (1999) Principles and practice of biomass fast pyrolysis processes for liquids. J Anal Appl Pyrol 51(1–2):3–22
541. Bridgwater AV, Peacocke GVC (2000) Fast pyrolysis processes for biomass. Renew Sustain Energ Rev 4(1):1–73
542. Czernik S, Bridgwater AV (2004) Overview of applications of biomass fast pyrolysis oil. Energy Fuel 18(2):590–598
543. Yaman S (2004) Pyrolysis of biomass to produce fuels and chemical feedstocks. Energy Convers Manage 45(5):651–671
544. Antal MJ, Gronli M (2003) The art, science, and technology of charcoal production. Ind Eng Chem Res 42(8):1619–1640
545. Balci S, Dogu T, Yucel H (1993) Pyrolysis kinetics of lignocellulosic materials. Ind Eng Chem Res 32(11):2573–2579
546. Branca C, Di Blasi C (2003) Kinetics of the isothermal degradation of wood in the temperature range 528–708 K. J Anal Appl Pyrol 67(2):207–219
547. Brown AL, Dayton DC, Daily JW (2001) A study of cellulose pyrolysis chemistry and global kinetics at high heating rates. Energy Fuel 15(5):1286–1294
548. Di Blasi C, Branca C (1999) Global degradation kinetics of wood and agricultural residues in air. Can J Chem Eng 77(3):555–561
549. Fisher T, Hajaligol M, Waymack B, Kellogg D (2002) Pyrolysis behavior and kinetics of biomass derived materials. J Anal Appl Pyrolysis 62(2):331–349
550. Koufopanos CA, Papayannakos N, Maschio G, Lucchesi A (1991) Modeling of the pyrolysis of biomass particles—studies on kinetics, thermal and heat-transfer effects. Canad J Chem Eng 69 (4):907–915
551. Manya JJ, Velo E, Puigjaner L (2003) Kinetics of biomass pyrolysis: a reformulated three-parallel-reactions model. Ind Eng Chem Res 42(3):434–441
552. Miller RS, Bellan J (1997) A generalized biomass pyrolysis model based on superimposed cellulose, hemicellulose and lignin kinetics. Combust Sci Technol 126(1–6):97–137
553. Diebold JP (1994) A unified, global-model for the pyrolysis of cellulose. Biomass Bioenergy 7(1–6):75–85
554. Bradbury AGW, Sakai Y, Shafizadeh F (1979) Kinetic-model for pyrolysis of cellulose. J Appl Polymer Sci 23(11):3271–3280
555. Antal MJ, Mok WSL, Varhegyi G, Szekely T (1990) Review of methods for improving the yield of charcoal from biomass. Energy Fuel 4(3):221–225
556. Wenzl HJA. Chemical technology of wood. Academic, Saint Louis, MO
557. Bridgwater AV (2003) Renewable fuels and chemicals by thermal processing of biomass. Chem Eng J 91(2–3):87–102
558. Bridgwater AV, Cottam ML (1992) Opportunities for biomass pyrolysis liquids production and upgrading. Energy Fuel 6 (2):113–120
559. Elliott DC, Beckman D, Bridgwater AV, Diebold JP, Gevert SB, Solantausta Y (1991) Developments in direct thermochemical liquefaction of biomass—1983–1990. Energy Fuel 5(3):399–410
560. Demirbas A, Arin G (2002) An overview of biomass pyrolysis. Energ Source 24(5):471–482
561. Shafizadeh F (1984) The chemistry of pyrolysis and combustion. Adv Chem Ser 207:491–529
562. Shafizadeh F (1982) Introduction to pyrolysis of biomass. J Anal Appl Pyrolysis 3(4):283–305
563. Bridgwater AE (2002) Fast pyrolysis of biomass: vol. 2: a handbook. CPL Scientific, Newbury, Berks
564. Bridgwater AA, Czernik SA, Diebold JA, Meier DA, Oasmaa AA, Peacocke CA, Piskorz JA (1999) Fast pyrolysis of biomass: a handbook. CPL Scientific Newbury Berks, Newbury, Berks
565. Oasmaa A, Czernik S (1999) Fuel oil quality of biomass pyrolysis oils—state of the art for the end user. Energy Fuel 13(4):914–921
566. Boucher ME, Chaala A, Roy C (2000) Bio-oils obtained by vacuum pyrolysis of softwood bark as a liquid fuel for gas turbines. Part I: properties of bio-oil and its blends with methanol and a pyrolytic aqueous phase. Biomass Bioenergy 19(5):337–350
567. Chiaramonti D, Bonini A, Fratini E, Tondi G, Gartner K, Bridgwater AV et al (2003) Development of emulsions from biomass pyrolysis liquid and diesel and their use in engines—part 2: tests in diesel engines. Biomass Bioenergy 25(1):101–111
568. Czernik S, Scahill J, Diebold J (1995) The production of liquid fuel by fast pyrolysis of biomass. J Sol Energ Eng Trans ASME 117(1):2–6
569. Ganesh A, Banerjee R (2001) Biomass pyrolysis for power generation—a potential technology. Renew Energy 22(1–3):9–14
570. Gayubo AG, Aguayo AT, Atutxa A, Aguado R, Bilbao J (2004) Transformation of oxygenate components of biomass pyrolysis oil on a HZSM-5 zeolite. I. Alcohols and phenols. Ind Eng Chem Res 43(11):2610–2618
571. Gayubo AG, Aguayo AT, Atutxa A, Aguado R, Olazar M, Bilbao J (2004) Transformation of oxygenate components of biomass pyrolysis oil on a HZSM-5 zeolite. H. Aldehydes, ketones, and acids. Ind Eng Chem Res 43(11):2619–2626
572. Grassi G, Bridgwater AV (1993) The opportunities for electricity production from biomass by advanced thermal-conversion technologies. Biomass Bioenergy 4(5):339–345
573. Peacocke GVC, Bridgwater AV (1994) Ablative plate pyrolysis of biomass for liquids. Biomass Bioenergy 7(1–6):147–154
574. Shihadeh A, Hochgreb S (2000) Diesel engine combustion of biomass pyrolysis oils. Energy Fuel 14(2):260–274
575. Solantausta Y, Beckman D, Bridgwater AV, Diebold JP, Elliott DC (1992) Assessment of liquefaction and pyrolysis systems. Biomass Bioenergy 2(1–6):279–297
576. Solantausta Y, Bridgwater AT, Beckman D (1995) Feasibility of power production with pyrolysis and gasification systems. Biomass Bioenergy 9(1–5):257–269
577. Vitolo S, Bresci B, Seggiani M, Gallo MG (2001) Catalytic upgrading of pyrolytic oils over HZSM-5 zeolite: behaviour of

the catalyst when used in repeated upgrading-regenerating cycles. Fuel 80(1):17–26

578. Wornat MJ, Porter BG, Yang NYC (1994) Single droplet combustion of biomass pyrolysis oils. Energy Fuel 8(5):1131–1142
579. Cottam ML, Bridgwater AV (1994) Technoeconomic modeling of biomass flash pyrolysis and upgrading systems. Biomass Bioenergy 7(1–6):267–273
580. Horne PA, Williams PT (1996) Upgrading of biomass-derived pyrolytic vapours over zeolite ZSM-5 catalyst: effect of catalyst dilution on product yields. Fuel 75(9):1043–1050
581. Williams PT, Horne PA (1995) The influence of catalyst regeneration on the composition of zeolite-upgraded biomass pyrolysis oils. Fuel 74(12):1839–1851
582. Williams PT, Horne PA (1995) The influence of catalyst type on the composition of upgraded biomass pyrolysis oils. J Anal Appl Pyrol 31:39–61
583. Radlein D (1999) The production of chemicals from fast pyrolysis bio-oils. In: Bridgwater A (ed) Fast pyrolysis of biomass: a handbook. CPL Press, Newbury, UK, pp 164–188
584. Tabatabaieraissi A, Trezek GJ (1987) Parameters governing biomass gasification. Ind Eng Chem Res 26(2):221–228
585. Beenackers AACM, Van Swaaij WPM (1984) Gasification of biomass, a state of the art review. In: Bridgwater AV (ed) Thermochemical processing of biomass. Butterworths, London, UK, pp 91–136
586. Hos JJ, Groeneveld MJ (1987) Biomass gasification. In: Hall DO, Overend RP (eds) Biomass. Wiley, Chichester, UK, pp 237–255
587. Beenackers A (1999) Biomass gasification in moving beds, a review of European technologies. Renew Energy 16 (1–4):1180–1186
588. Li XT, Grace JR, Lim CJ, Watkinson AP, Chen HP, Kim JR (2004) Biomass gasification in a circulating fluidized bed. Biomass Bioenergy 26(2):171–193
589. Scala F, Chirone R (2004) Fluidized bed combustion of alternative solid fuels. Exp Therm Fluid Sci 28(7):691–699
590. Scala F, Salatino P (2002) Modelling fluidized bed combustion of high-volatile solid fuels. Chem Eng Sci 57(7):1175–1196
591. Lanauze RD (1987) A review of the fluidized-bed combustion of biomass. J Inst Energy 60(443):66–76
592. Asif M, Ibrahim AA (2002) Minimum fluidization velocity and defluidization behavior of binary-solid liquid-fluidized beds. Powder Technol 126(3):241–254
593. Stultz SC, Kitto JB, Rahn CH (1992) Chapter 16: atmospheric pressure fluidized-bed boilers. In: Stulz S, Kitto J (eds) Steam: its generation and use, vol 40. Babcock Wilcox Co, Barberton, OH, p 1064
594. Maschio G, Lucchesi A, Stoppato G (1994) Production of syngas from biomass. Bioresour Technol 48(2):119–126
595. Littlewood K (1977) Gasification—theory and application. Prog Energy Combust Sci 3(1):35–71
596. De Bari I, Barisano D, Cardinale M, Matera D, Nanna F, Viggiano D (2000) Air gasification of biomass in a downdraft fixed bed: a comparative study of the inorganic and organic products distribution. Energy Fuel 14(4):889–898
597. Wyman CE, Hinman ND, Bain RL, Stevens DJ (1993) Ethanol and methanol from cellulosic materials, Chapter 21. In: Johansson TB, Kelly H, Reddy AKN, Williams RH (eds) Renewable energy, sources for fuels and electricity. Island Press, Washington, DC, pp 865–923. Acc No. 12161
598. Milne TA, Abatzoglou N, Evans RJ (1998) Biomass gasifier "tars": their nature, formation, and conversion. NREL report no. TP-570-25357. National Renewable Energy Laboratory, Golden, CO, Nov 1998, p 202
599. Simell P, Stahlberg P, Kurkela E, Albrecht J, Deutsch S, Sjostrom K (2000) Provisional protocol for the sampling and analysis of tar and particulates in the gas from large-scale biomass gasifiers. Version 1998. Biomass Bioenergy 18(1):19–38
600. Knoef HAM, Koele HJ (2000) Survey of tar measurement protocols. Biomass Bioenergy 18(1):55–59
601. Maniatis K, Beenackers A (2000) Tar protocols. IEA bioenergy gasification task. Biomass Bioenergy 18(1):1–4
602. Caballero MA, Corella J, Aznar MP, Gil J (2000) Biomass gasification with air in fluidized bed. Hot gas cleanup with selected commercial and full-size nickel-based catalysts. Ind Eng Chem Res 39(5):1143–1154
603. Cummer KR, Brown RC (2002) Ancillary equipment for biomass gasification. Biomass Bioenergy 23(2):113–128
604. Devi L, Ptasinski KJ, Janssen F (2003) A review of the primary measures for tar elimination in biomass gasification processes. Biomass Bioenergy 24(2):125–140
605. Simell PA, Hepola JO, Krause AOI (1997) Effects of gasification gas components on tar and ammonia decomposition over hot gas cleanup catalysts. Fuel 76(12):1117–1127
606. Consonni S, Larson ED (1996) Biomass-gasifier/aeroderivative gas turbine combined cycles. A. Technologies and performance modeling. J Eng Gas Turb Power Trans ASME 118(3):507–515
607. Consonni S, Larson ED (1996) Biomass-gasifier/aeroderivative gas turbine combined cycles. B. Performance calculations and economic assessment. J Eng Gas Turb Power Trans ASME 118 (3):516–525
608. Jurado F, Ortega M, Cano A, Carpio J (2001) Biomass gasification, gas turbine, and diesel engine. Energ Source 23 (10):897–905
609. Kinoshita CM, Turn SQ, Overend RP, Bain RL (1997) Power generation potential of biomass gasification systems. J Energ Eng ASCE 123(3):88–99
610. Abu El-Rub Z, Bramer EA, Brem G (2004) Review of catalysts for tar elimination in biomass gasification processes. Ind Eng Chem Res 43(22):6911–6919
611. Gerhard SC, Wang DN, Overend RP, Paisley MA (1994) Catalytic conditioning of synthesis gas produced by biomass gasification. Biomass Bioenergy 7(1–6):307–313
612. Kinoshita CM, Wang Y, Zhou J (1994) Tar formation under different biomass gasification conditions. J Anal Appl Pyrol 29 (2):169–181
613. Marsak J, Skoblja S (2002) Role of catalysts in tar removal from biomass gasification. Chem Listy 96(10):813–820
614. Simell P, Kurkela E, Stahlberg P, Hepola J (1996) Catalytic hot gas cleaning of gasification gas. Catal Today 27(1–2):55–62
615. Sutton D, Kelleher B, Ross JRH (2001) Review of literature on catalysts for biomass gasification. Fuel Process Technol 73 (3):155–173
616. Baker EG, Mudge LK, Brown MD (1987) Steam gasification of biomass with nickel secondary catalysts. Ind Eng Chem Res 26 (7):1335–1339
617. Brown MD, Mudge LK, Baker EG (1984) Catalysts for gasification of biomass. Biotechnol Bioeng 14:125–136
618. Mudge LK, Sealock LJ, Weber SL (1979) Catalyzed steam gasification of biomass. Anal Appl Pyrol 1(2):165–175
619. Stevens DJ (2001) Hot gas conditioning: recent progress with larger-scale biomass gasification systems. NREL/SR-510-29952. Department of Energy, National Renewable Energy Laboratory, Golden, CO, Aug 2001, p 88
620. Corella J, Aznar MP, Gil J, Caballero MA (1999) Biomass gasification in fluidized bed: where to locate the dolomite to improve gasification? Energy Fuel 13(6):1122–1127
621. Corella J, Toledo JM, Padilla R (2004) Olivine or dolomite as in-bed additive in biomass gasification with air in a fluidized bed: which is better? Energy Fuel 18(3):713–720

622. Delgado J, Aznar MP, Corella J (1997) Biomass gasification with steam in fluidized bed: effectiveness of CaO, MgO, and CaO-MgO for hot raw gas cleaning. Ind Eng Chem Res 36(5):1535–1543
623. Devi L, Ptasinski KJ, Janssen F, van Paasen SVB, Bergman PCA, Kiel JHA (2005) Catalytic decomposition of biomass tars: use of dolomite and untreated olivine. Renew Energy 30(4):565–587
624. Orio A, Corella J, Narvaez I (1997) Performance of different dolomites on hot raw gas cleaning from biomass gasification with air. Ind Eng Chem Res 36(9):3800–3808
625. Simell PA, Leppalahti JK, Kurkela EA (1995) Tar-decomposing activity of carbonate rocks under high CO_2 partial-pressure. Fuel 74(6):938–945
626. Aznar MP, Caballero MA, Gil J, Martin JA, Corella J (1998) Commercial steam reforming catalysts to improve biomass gasification with steam-oxygen mixtures. 2. Catalytic tar removal. Ind Eng Chem Res 37(7):2668–2680
627. Caballero MA, Aznar MP, Gil J, Martin JA, Frances E, Corella J (1997) Commercial steam reforming catalysts to improve biomass gasification with steam-oxygen mixtures. 1. Hot gas upgrading by the catalytic reactor. Ind Eng Chem Res 36(12):5227–5239
628. Courson C, Makaga E, Petit C, Kiennemann A (2000) Development of Ni catalysts for gas production from biomass gasification. Reactivity in steam- and dry-reforming. Catal Today 63 (2–4):427–437
629. Bridgwater AV (1995) The technical and economic-feasibility of biomass gasification for power-generation. Fuel 74(5):631–653
630. Stahl K, Neergaard M (1998) IGCC power plant for biomass utilisation, Varnamo, Sweden. Biomass Bioenergy 15(3):205–211
631. Lundqvist RG (1993) The IGCC demonstration plant at Varnamo. Bioresour Technol 46(1–2):49–53
632. Dasappa S, Paul PJ, Mukunda HS, Rajan NKS, Sridhar G, Sridhar HV (2004) Biomass gasification technology—a route to meet energy needs. Curr Sci 87(7):908–916
633. Dasappa S, Sridhar HV, Sridhar G, Paul PJ, Mukunda HS (2003) Biomass gasification—a substitute to fossil fuel for heat application. Biomass Bioenergy 25(6):637–649
634. Fuel Cell Handbook Corp Author(s), EG and G Services Staff, Author, Parsons, Inc. Staff, Author, SA/C Staff, Author, 5th edn. Business/Technology (B/T), Orinda, 2000, p 292
635. Appleby AJ (1996) Fuel cell technology: status and future prospects. Energy 21(7–8):521–653
636. Cappadonia M, Stimming U, Kordesch KV, de Oliveira JCT (2005) Fuel cells. In: Ullmann's encyclopedia of industrial chemistry, 7th edn. Wiley, Hoboken, p 400
637. Kordesch KVA, Simader GA (2000) Fuel cells and their applications. Wiley, Hoboken, NJ, p 375
638. Larminie J (2002) Fuel cells. In: Kirk-Othmer encyclopedia of chemical technology. vol 12, 5th edn. Wiley, Hoboken, NJ, p 850
639. Eichenberger PH (1998) The 2 MW Santa Clara project. J Power Sources 71(1–2):95–99
640. Figueroa RA, Otahal J (1998) Utility experience with a 250-kW molten carbonate fuel cell cogeneration power plant at NAS Miramar, San Diego. J Power Sources 71(1–2):100–104
641. Huijsmans JPP, Kraaij GJ, Makkus RC, Rietveld G, Sitters EF, Reijers HTJ (2000) An analysis of endurance issues for MCFC. J Power Sources 86(1–2):117–121
642. Huijsmans JPP, van Berkel FPF, Christie GM (1998) Intermediate temperature SOFC—a promise for the 21st century. J Power Sources 71(1–2):107–110
643. Lobachyov KV, Richter HJ (1998) An advanced integrated biomass gasification and molten fuel cell power system. Energ Convers Manage 39(16–18):1931–1943
644. Gesser HD, Hunter NR (1998) A review of C-1 conversion chemistry. Catal Today 42(3):183–189
645. Green AES (1991) Overview of fuel conversion. FACT (American Society of Mechanical Engineers), 12 (Solid Fuel Conversion for the Transportation Sector), 3–15
646. Keim, W., C1 chemistry: present status and aspects for the future. In: Chemistry for the future, Proceedings of the IUPAC congress, 29, 1984, pp 53–62
647. Rostrup-Nielsen JR (2002) Syngas in perspective. Catal Today 71 (3–4):243–247
648. Wender I (1996) Reactions of synthesis gas. Fuel Process Technol 48(3):189–297
649. Spath PL, Dayton DC (2003) Preliminary screening—technical and economic assessment of synthesis gas to fuels and chemicals with emphasis on the potential for biomass-derived syngas, TP-510-34929. National Renewable Energy Laboratory, Golden, CO, p 160
650. Werther J, Saenger M, Hartge EU, Ogada T, Siagi Z (2000) Combustion of agricultural residues. Prog Energy Combust Sci 26(1):1–27
651. van den Broek R, Faaij A, van Wijk A (1996) Biomass combustion for power generation. Biomass Bioenergy 11(4):271–281
652. Prasad SB (1995) Biomass-fired steam power cogeneration system—a theoretical-study. Energ Convers Manage 36(1):65–77
653. Baxter LL, Miles TR, Miles TR, Jenkins BM, Milne T et al (1998) The behavior of inorganic material in biomass-fired power boilers: field and laboratory experiences. Fuel Process Technol 54 (1–3):47–78
654. Nussbaumer T (2003) Combustion and co-combustion of biomass: fundamentals, technologies, and primary measures for emission reduction. Energy Fuel 17(6):1510–1521
655. Demirbas A (2003) Toxic air emissions from biomass combustion. Energ Source 25(5):419–427
656. Hubbard AJ (1995) Hazardous air emissions potential from a wood-fired furnace. Combust Sci Technol 108(4–6):297–309
657. Baxter LL (1993) Ash deposition during biomass and coal combustion—a mechanistic approach. Biomass Bioenergy 4 (2):85–102
658. Blander M, Milne TA, Dayton DC, Backman R, Blake D et al (2001) Equilibrium chemistry of biomass combustion: a round-robin set of calculations using available computer programs and databases. Energy Fuel 15(2):344–349
659. Dayton DC, Frederick WJ (1996) Direct observation of alkali vapor release during biomass combustion and gasification. 2. Black liquor combustion at 1100 degrees C. Energy Fuel 10 (2):284–292
660. Dayton DC, French RJ, Milne TA (1995) Direct observation of alkali vapor release during biomass combustion and gasification. 1. Application of molecular-beam mass-spectrometry to switchgrass combustion. Energy Fuel 9(5):855–865
661. Glarborg P, Jensen AD, Johnsson JE (2003) Fuel nitrogen conversion in solid fuel fired systems. Prog Energy Combust Sci 29 (2):89–113
662. Bakker RR, Jenkins BM, Williams RB (2002) Fluidized bed combustion of leached rice straw. Energy Fuel 16(2):356–365
663. Nordin A (1994) Chemical elemental characteristics of biomass fuels. Biomass Bioenergy 6(5):339–347
664. Brouwer J, Owens WD, Harding NS, Heap MP, Pershing DW (1995) Cofiring waste biofuels and coal for emissions reduction. Abstr Paper Am Chem Soc 209:32-FUEL
665. Demirbas A (2003) Sustainable cofiring of biomass with coal. Energ Convers Manage 44(9):1465–1479
666. Harding NS, Adams BR (2000) Biomass as a reburning fuel: a specialized cofiring application. Biomass Bioenergy 19 (6):429–445
667. Niksa S, Liu GS, Felix L, Bushy PV, Boylan DM (2003) Predicting NOX emissions from biomass cofiring. Abstr Paper Am Chem Soc 226:U540

668. Hughes EE, Tillman DA (1998) Biomass cofiring: status and prospects 1996. Fuel Process Technol 54(1–3):127–142
669. Boylan D, Bush V, Bransby DI (2000) Switchgrass cofiring: pilot scale and field evaluation. Biomass Bioenergy 19(6):411–417
670. Boylan DM (1996) Southern company tests of wood/coal cofiring in pulverized coal units. Biomass Bioenergy 10(2–3):139–147
671. Gold BA, Tillman DA (1996) Wood cofiring evaluation at TVA power plants. Biomass Bioenergy 10(2–3):71–78
672. Tillman DA (2000) Biomass cofiring: the technology, the experience, the combustion consequences. Biomass Bioenergy 19 (6):365–384
673. Tillman DA, Stahl RW (1995) Wood cofiring experience in cyclone boilers. Abstr Paper Am Chem Soc 209:31-FUEL
674. Tillman DA (1997) Biomass cofiring guidelines, EPRITR-108952. Foster Wheeler Environmental Corporation for the Electric Power Research Institute, Sacramento, CA, p 107
675. Nieves RA, Ehrman CI, Adney WS, Elander RT, Himmel ME (1998) Survey and analysis of commercial cellulase preparations suitable for biomass conversion to ethanol. World J Microbiol Biotechnol 14(2):301–304
676. Shoemaker S, Raymond J, Bruner R (1981) Cellulases: diversity amongst improved Trichoderma strains. In: Hollander A (ed) Trends in the biology of fermentations for fuels and chemicals. Plenum Press, New York, pp 89–129
677. Enari TM (1983) Microbial cellulases, 1st edn. Applied Science, London, pp 183–223
678. Mandels M, Reese ET (1964) Fungal cellulases and the microbial decomposition of cellulosic fabric. Dev Ind Microbiol 5:5–20
679. Selig, M. J., W. S. Adney, M. E. Himmel and S. R. Decker. (2009). “The Impact of Cell Wall Acetylation on Corn Stover Hydrolysis by Cellulolytic and Xylanolytic Enzymes”. Cellulose. 16:711–722
680. Selig, M. J., T. B. Vinzant, M. E. Himmel, and S. R. Decker. (2009). “The Effect of Lignin Removal by Alkaline Peroxide Pretreatment on the Susceptibility of Corn Stover to Purified Cellulolytic and Xylanolytic Enzymes”. Appl. Biochem. Biotechnol. 155:397–406.

Animal and Vegetable Fats, Oils, and Waxes

34

Edmund W. Lusas, Mohammad S. Alam,
Richard C. Clough, and Mian N. Riaz

Biotechnology in Agriculture and Processing

Human Survival Is Biotechnology

Biotechnology has been defined by various groups and broadly includes technologies that utilize living organisms or parts of biological systems. The nurture of man and animals, and provision of replenishable industrial materials, typically includes: (1) growing selected species or their genetic modifications; (2) harvest, preprocess storage, conversion into useful products, and protection until use; and (3) utilization or disposal of byproducts and wastes in the most beneficial or least-cost manner. Specific actions may be taken to suppress residual enzymes and contaminating microorganisms that could degrade product value. Also, *remediation* (restoration) of air and water used in processing to near-pristine condition often is mandated today.

The first *transgenic* (across genera) oilseed crops were planted in the United States and Canada in 1996 and were followed by rapid acreage expansion. Transgenic oilseeds generally do not require special processing unless they contain higher melting oils. In this chapter, the reader is first introduced to modern biology principles, and industry terms are presented throughout.

Human understanding of life processes, and competition between genera, continually broadens. Many biological reactions are catabolic and split large compounds into smaller units to obtain energy (carried by ATP, adenosine triphosphate) for heat, work, or reinvestment in syntheses and chemical structures for making required compounds. Plants are the original source of nutrients for man and animals. Microscopic plants, plankton, and larger members synthesize proteins, carbohydrates, lipids, and other compounds from inorganic elements and water, using ATP obtained from the sun by photosynthesis. In the process, carbon dioxide also is reduced to oxygen. Some products are used to build the plant's structures and maintain its functions, and others are stored in the seed as enzymes, proteins, and energy reserves (starches and lipids) for reproduction as new plants, or as sugars or starches in roots of perennials for later growth cycles. Humans rely on plants and animals, which consume plants (herbivores), or carnivores which eat other animals, for their food. Animals vary in their abilities to synthesize intermediate chemical structures needed for their development and life; generally, those lower in the food chain are more self-sufficient. As examples, amino acids are the building blocks of proteins and 11 are *dietary essential* (must be obtained from food) by humans; approximately 12–14 are dietary essential for economic animals (grown for meat or other food products), the exact needs varying with species, age, and physiological state. Carnivorous fish obtain long-chain highly unsaturated fatty acids in the wild by feeding on plankton eaters and often lose the ability to synthesize them directly. When grown in captivity, as in salmon farming, oils of plankton-eating fish are added to their feed. Diunsaturated linoleic acid and triunsaturated linolenic acid are considered dietary essential for humans. Also, longer-chain higher-unsaturated fish oils sometimes are prescribed for people with lipid metabolism deficiencies.

All organically made compounds, fossilized or not, were once synthesized by specific enzymes and pathways. They must remain degradable by enzymes of the same organism to perpetuate its life and by enzymes of other species for their nutrition and for biodegradation and carbon recycling. Life processes come to a halt if the laws of biochemistry and physical chemistry aren't satisfied. This chapter is limited to

---

This chapter was originally conceived and written by Dr. Lusas and appeared in the immediately preceding three editions.

E.W. Lusas (✉)
Ed Lusas, Consultant, 3604 Old Oaks Drive, Bryan, TX 77802, USA

M.S. Alam • R.C. Clough • M.N. Riaz
Food Protein Research and Development Center, Texas A&M University, 2476 TAMU, College Station, TX 77843-2476, USA
e-mail: mnriaz@tamu.edu

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_34,

lipids, but business and policy decisions about food, feed, and replenishable industrial raw materials also require holistic information about protein and carbohydrate technologies, about needs and preferences of biological users, and about human enterprises throughout the world.

Biotechnology has been implemented in two phases. Macro-biotechnology examples are commonplace and include agriculture, baking, and various fermentations. Implementation began about 10,000 years ago, even before written history, when humans initiated symbiotic relations with selected wild animals and plants to ensure a more reliable food supply than possible by hunting/fishing and gathering. Animals were refashioned through selection, protection, crossbreeding, and increase, to meet man's food, transportation, security, and compatibility needs. Higher-yielding seeds, which also were easier to dehull and process into food, were replanted for succeeding harvests. In time, the face of the earth's habitable lands was changed. Except in remote areas or designated refuges, humans: (1) eliminated wild animals dangerous to them and their protected crops and animals; (2) cleared lands and jungles for firewood, building materials, and growing additional crops, as the population increased; but (3) sometimes also destroyed arable land by over-grazing and erosion of uncovered top soil.

Biotechnology currently emphasizes microscale aspects that were not visible or envisioned until about 150 years ago, including: (1) microbes; (2) stereochemistry of reactions, molecular biology including transcription of genetic codes and translation for protein production, and cell multiplication; and (3) enzymes as catalysts. Humans came to only partially understand the nature of systems (animals, plants, insects, and soil) they handled in agriculture. Modern biotechnology must address needs of these systems, as well as those of people. Animals and plants must be healthy and adequately nourished, the pests (weeds and insects) controllable, the soil kept productive, and all must function in acceptable harmony if the quality of human life is to be sustained or improved as the population increases.

Histories of biotechnology progress in agriculture, genetics, medicine, pharmacy, industrial fermentations, food processing, and nutrition are reported in many books and on the Internet. Archeologists tell us the Babylonians brewed beer in 6000 BC, Egyptians baked yeast-leavened bread in 4000 BC, and other early cultures made wine and vinegar by fermentation and preserved cheeses by salt and acid. Sophisticated herbal medicines were developed. But, without scientific knowledge of why certain practices were successful, early progress through trial and error was earned slowly. Some of the less-mentioned milestones relevant to food production and utilization are described next.

Development of the microscope by Anton van Leeuwenhoek (1673) and others enabled man to see a new world of "microcreatures," including bacteria, yeasts, molds, blood cells, and spermatozoa, and microstructures such as muscle fibers and plant and seed tissues. This added credibility to later claims by Louis Pasteur and Robert Koch that microorganisms can spoil foods and cause disease at a time when many influential learned men still clung to theories of spontaneous generation of life.

Louis Pasteur (1863) invented a process to pasteurize wine by heating to inactivate microbes and prevent its turning sour from fermentation to acid. Nicolas Appert had already won a 12,000-franc prize from Napoleon I in 1809 by inventing a process for "canning" foods (actually packing precooked foods in glass bottles, sealing, and heating in boiling water). But, spoilage occurred frequently because few understood the roles of acidity or osmotic pressure in restricting microbial growth, problems of recontamination, or heating inactivation requirements for different bacteria, especially spore formers. The roles of bacteria and yeast in fermentations had been accepted when Eduard Buchner (1897) demonstrated that filtered extracts also could catalyze fermentation in the absence of intact yeast cells [1]. This led to modern enzymology where more efficient enzymes are extracted and purified from a variety of nontraditional sources, and more recently from "engineered" microorganisms that carry gene sequences from relatively unrelated species.

Genetic Messages

Gregor Mendel, an Austrian botanist and monk working with sweet pea plants, proposed that "factors" (units of information responsible for observable traits) are passed from one generation to the next (1866). But, his work essentially went unnoticed until rediscovered by several scientists in the early 1900s. Walther Flemming reported the discovery of chromosomes and mitosis in 1882. William Sutton (1902) announced that chromosomes are paired and may be the carriers of heredity. He named Mendel's factors "genes" and suggested they occur on chromosomes.

After observing meiosis, he developed the Chromosomal Theory of Heredity that gametes (egg and sperm cells) are haploid and carry only one of each pair of chromosomes from each parent in forming the new being (a diploid zygote) which then holds on to life as it best can. Details of genetics continued to develop. Alfred Hershey and Marsha Chase showed that deoxyribonucleic acid (DNA) is the genetic code carrier rather than proteins as was thought by some (1952). James Watson and Francis Crick proposed the double-stranded, complementary, antiparallel structure model for DNA (1953) [1].

Development of recombinant DNA techniques to produce engineered, genetically engineered (GE), or genetically modified organism (GMO) products is among the boldest achievements of cellular biologists. Paul Berg (1972) first demonstrated use of a selected enzyme to cut DNA strands into sections and reattach them to produce recombinant DNA molecules. In 1973, Stanley Cohen, Annie Chang, and Herbert Boyer announced splicing a viral DNA section with a bacterial DNA section to form a recombinant DNA molecule, which then was spliced into the DNA of a bacteria to produce the first *transgenic* (across genera) recombinant DNA organism. Herbert Boyer and Robert Swanson founded Gene tech, Incorporated, a biotechnology company dedicated to developing and marketing products based on recombinant DNA technology in 1976. It was the first company to receive US Food and Drug Administration (FDA) approval for marketing a genetically engineered drug (a form of human insulin produced by bacteria) in 1982, and since has developed many other products [1].

Research in genetic and protein structures created tremendous amounts of data. Fortunately, means to handle it also were being developed on another front. Norbert Wiener, a mathematician at the Massachusetts Institute of Technology and a founder of the science of cybernetics, published his book *Cybernetics; or, Control and Communication in the Animal and the Machine* in 1948 [2]. In this volume and its 1961 edition [3], Wiener, a specialist in communications and later in artificial intelligence [4], knew that messages degrade and can be fouled in retransmission and suggested that aging and diseases are analogous degradations of the genetic code. Realizing that a revolution in information handling was occurring, Wiener became concerned that cybernetics and automation not degrade the quality of life and published *Human Use of Human Beings: Cybernetics and Society* in 1954 [5]. It also was reissued posthumously in 1967 [6]. Cybernetics, and especially digital computing, have made the massive record keeping, data summarization, inventory handling, sequencing, scheduling, and control systems of today's business, government, and science possible. Complete amino acid sequences of proteins have been mapped, as well as genomes of microorganisms, plants, animals, and man. As medical research workers become interested in specific genetic traits or problems, or desire to investigate modified-protein drugs, they will more readily know where to locate the corresponding genes and structures.

Cell culture techniques, developed during research leading to transgenic DNA organisms, greatly assisted in the development of many nontransgenic *cultivars* (uniform subvarieties of a recognized genus and species) by selective cross-breeding of closely related germplasms.

Fatty acids in plants and animals typically are elongated during synthesis to chain lengths of 18 carbons by successive additions of two carbon units. The resulting product, stearic acid, is fully saturated, but has a melting point too high for most physiological reactions. Nature remedies this situation by also providing *desaturase* systems that enable organisms to introduce 1, 2, or 3 unsaturated bonds in the 18-carbon chain, thus lowering melting point of fatty acids and triglycerides. Melting points differ between fat deposit locations in animals, change seasonally in animal bodies, and vary with seed maturation temperatures of the same cultivar.

Stability of fatty acids against oxidation decreases as unsaturation increases, a prime example being tri-unsaturated 18-carbon linolenic acid in soybean oil. Biotechnologists developing soybean with greater oil stability have sought to do so by seeking germplasms with reduced desaturation for selective crossing, and by altering or suppressing expression of desaturation [7], to produce oils with reduced linolenic acid and increased oleic or stearic acid contents.

Water Activity

W.J. Scott, an Australian microbiologist, published a classic food science chapter "Water Relations of Food Spoilage Microorganisms" in 1957 [8]. He recognized that microorganisms compete with other solutes in mixed systems for water required for their metabolism and summarized ranges within which most bacteria, yeasts, and molds grow. However, water was not expressed as percent of system mass, but rather as *water activity* ($a_W = p/p_o$), where p is the vapor pressure of the solution (or product), and p_o is the vapor pressure of the pure solvent (water). Some prefer to use the term relative vapor pressure (RVP), which can be determined by measuring percent relative humidity in the headspace above a product that has equilibrated within its container and the relationship RVP = (percent ERH)/100. Water activity, osmotic pressure, lowering of freezing point, and rising of boiling point are colligative properties of solutions that follow Raoult's Law.

The guides indicated that relatively few bacteria grow at a_W below 0.91, few yeasts below 0.87, and few molds below 0.75–0.65 a_W and helped explain why many foods do not spoil even though they are not sterile. Microbial growth ability can be further reduced by increasing osmotic pressure by addition of soluble salts and sugars, lowering pH, use of microbial inhibitors, and reducing product temperatures. Exceptions to the guides exist, for example, halophilic bacteria tolerate high salt concentrations and osmophilic yeasts can grow on high sugar content products.

Food-induced diseases are mainly of two types: *food poisoning,* the effect of toxins produced during growth of a microorganism in the food before ingestion, and *infection,* resulting from food-borne microorganisms taking up residence in the body before doing their damage. Low a_W has been reported to inhibit production of toxins by various food-poisoning bacteria, even though the live organisms are present. However, some low a_W processed meats have been implicated as carriers of viable food infection bacteria, which apparently resume growth when a_W and temperature increase. Water activity can be controlled by: product formulation; drying fruits, vegetables, and meat; and by storing products like meats and fish in salt brines and fruit preserves in sugar syrups.

Over time, diagrams were developed relating water activity with: enzyme activity, dormancy of stored seed, loss of dry product crispness by moisture absorption, pigments and vitamins degradation, nonenzymatic browning, and fat oxidation. Response curves generally are not linear, and readers working with food or feed formulations are referred to the technical literature about their products.

Herbicide- and Insect-Resistant Oilseed Crops

Crop producers and processors were provided with powerful tools during the last four decades of biotechnology evolution. Seed producers can develop new cultivars by crossing two selected parents, or by *hybridization* (crossing two genetically diverse genotypes to obtain additional vigor). Generally, transgenic crops also are hybrids. The world's current major biotech oilseed crops are herbicide-tolerant (HT) soybean, corn, cotton, and canola (a type of rapeseed), and insect-resistant corn (maize) and cotton.

United States and European pharmaceutical and chemical companies have taken the lead in biotechnology, frequently by purchasing plant breeders and growers of planting seeds. Various producers offer transgenic crop seeds resistant to several herbicides. The Monsanto Company (St. Louis, MO) has often been among the first to get its products to the marketplace and is used as an example here. Roundup Ready® Soybean (planting seed) was introduced in the United States, and Roundup Ready Canola in Canada, in 1996. These transgenic crops are resistant to Roundup,® Monsanto's brand of glyphosate, a *nonselective* (broad-spectrum) herbicide for killing *weeds* (unwanted plants). A grower of a Roundup Ready crop needs to make fewer spray passes across the field, often with one herbicide, to control a season's weeds. Roundup Ready Cotton was commercialized in 1997 and Roundup Ready Corn in 1998. Monsanto introduced Bollgard® Cotton in 1996, and YieldGard® Corn in 1998. The Bollgard and YieldGard seeds included a gene sequence from the soil bacterium *Bacillus thurigiensis* var. kurstaki ("Btk") which enables the plant to make *systemic pesticides* toxic to larvae of specific insects. These compounds are known as plant-incorporated protectants (PIP) and formerly were called "plant pesticides"; the crops categorically are called "B.t."

By 2005, the Monsanto Company marketed: Roundup Ready Soybeans, Roundup Ready Canola, Roundup Ready Cotton, Roundup Ready Corn, and Roundup Ready Corn 2. Additional gene sequences from "Btk," which produce other PIP specific for other insect larvae, had been inserted into seeds, and Monsanto also offered: Bollgard Cotton for controlling tobacco budworm and pink bollworm, and high suppression of cotton bollworm; and Bollgard II® for control of tobacco budworm, pink bollworm, cotton bollworm, fall armyworm, cabbage and soybean loopers, and other secondary leaf- or fruit-feeding caterpillar pests in cotton. (It is ironic that none of the commercial B.t. cotton seeds control the cotton boll weevil, a long-time nemesis, but concurrent boll weevil eradication programs have made good progress toward this objective.) Monsanto's B.t. Corn line included: YieldGard Corn Borer®, effective against European corn borer, southwest corn borer, sugarcane borer, southern cornstalk borer, corn earworm, fall armyworm, and stalk borer; and YieldGard Rootworm®, effective against western, northern, and Mexican corn rootworm. In addition, YieldGard Plus® offered the protective properties of both YieldGard Corn Borer and YieldGard Rootworm®. The inclusion of two or more transgenic properties in a seed is known as "stacking." Other seeds offered include: YieldGard Corn Borer with Roundup Ready Corn; YieldGard Corn Borer with Roundup Ready Corn 2; YieldGard Rootworm with Roundup Ready Corn 2; YieldGard Plus® with Roundup Ready Corn 2; and Bollgard II with Roundup Ready Cotton. The Monsanto Company also licenses the use of Roundup Ready traits to other seed producers who offer their Roundup Ready lines in the marketplace. The growing of transgenic crops is approved variety-by-variety and state-by-state. The Monsanto Company sees its relation with growers as licensor-licensees of its technologies, rather than as traditional seller-buyers, and establishes licensor-monitoring rights before selling seed.

The *2005 Technology Use Guide*, downloaded from the Monsanto Company Web site, states that growers are required to sign a Monsanto Technology Stewardship Agreement to: (1) comply with all EPA (US Environmental Protection Agency) mandated Insect Resistance Management (IRM) requirements; (2) use all purchased seed with biotech traits for planting a single crop; and (3) sell harvested corn with biotech traits not yet approved by the European Union only to grain handlers who confirm their acceptance, or use that grain as on-farm feed. Finding an initial crop buyer is the grower's responsibility, and prior discussions with grain elevators and Monsanto seed dealers are encouraged.

Refuges (20% minimum to 50% of the total crop acreage) planted to the same crop (corn or cotton) with non-B.t. seed are required by the EPA as part of the IRM effort to slow development of B.t.-resistant insects. The concept is to provide an area for presumably B.t.-resistant insects escaping the B.t. crop area to mate with non-B.t.-resistant insects in an area where other insect control practices are used, and thus slow development of B.t.-resistant insects that survive to the following year. Options for refuge patterns and location are described, and growers must agree that Monsanto personnel or its agents will have access to monitor the transgenic crop and refuge areas.

Growers may use glyphosate herbicides, other than those designated by Monsanto, but only if they have been approved for use over Roundup Ready crops, and have been labeled for this use by all required governmental agencies.

If Monsanto suspects a grower may have planted saved seed containing a Monsanto genetic trait, it may request invoices or otherwise confirm that newly purchased seed has been planted. If information is not provided within 30 days, Monsanto may inspect and test all the grower's fields to determine if saved seed has been planted. Inspections are to be scheduled in advance at a reasonable time so the grower can be present if desired.

The US Department of Agriculture (USDA) has estimated that herbicide-tolerant (HT) soybeans reached 85% of total US soybean acreage and HT cotton reached 60% of total in 2004, whereas the HT share of corn reached 18% [9]. The ISAAA (International Service for the Acquisition of Agri-Biotech Applications) has reported that world cultivation of biotech crops increased 47-fold, from 1.7 million hectares (4.2 million acres) to 81.0 million hectares (1 ha = 2.47 acres) in the 9-year period 1996–2004, and was up by 20% from 2003. The number of countries growing 50,000 ha (123,500 acres) or more annually increased from the original United States and Canada to 14 in 2004, with the US still planting the most land (59% of global GMO plantings), followed by Argentina (20%), Canada (6%), Brazil (6%), China (5%), Paraguay (2%), India (1%), South Africa (1%), and Uruguay, Australia, Romania, Mexico, Spain, and The Philippines each planting less than 1%. Approximately 34% of the global biotech crop area in 2004 was in developing countries, with use in Southern Hemisphere developing countries increasing about three times more rapidly than northern industrial countries [10]. Growth in the five principal developing countries (China, India, Argentina, Brazil, and South Africa) portends significant changes in world food/feed trade patterns, which already are unfolding.

Concerns About Transgenic Crops

Benefits of biotech crops have not always materialized as expected. Introductions of many glyphosate-resistant transgenic crops have been met with complaints of "yield drag" in the initial years, with as much as 6–10% yield reduction for soybean compared to nontransgenic hybrids in test plots reported. Damage to root inoculants or nitrogen-fixing nodules by glyphosate and incomplete cleaning of mixed herbicide residues from weed sprayer tanks have been proposed as causes. However, problems seem to lessen in succeeding years, with no yield drag observed in glyphosate-resistant corn, and ~3% yield drag reported in more recent soybean tests. Yield is determined by genes other than those for glyphosate resistance and by adaptability of cultivars to specific soil conditions. It takes longer to optimize crops as the number of desired traits increases, and improved nontransgenic hybrids may be brought to market in shorter time. Soils differ in fertility, drainage, compacting, and temperature, and some growers chose several seed varieties for different soil conditions on the same farm. Reduced costs of tillage, including opportunities for drilling seed into no-tilled fields (nonplowed stubble of previous crops) and fewer trips across fields for herbicide application, are offset by higher seed costs in calculating overall costs of producing a crop. Economical justification of herbicide-tolerant crops seems confirmed by rapid increases in acres planted in the United States in recent years. Insect larvae carry fungi that produce mycotoxins in corn and cottonseed that can cause illness and death in economic animals and cancer in humans. Elimination of such problems also is welcomed, in addition to increased yields of B.t. crops protected against insect invasion.

Experience showed that some early visions of one-pass herbicide treatment of HT crops per season were overly optimistic. Various practices for controlling weeds are needed across the country, including an herbicide "burn down" of sprouted weeds before crop drilling, dealing with weeds that emerge at different times during the year and well into the growth of cotton, and control of perennial weeds and those resistant to glyphosate. Weed control practices must be tailored to local conditions, even though fewer applications may be required when planting glyphosate-resistant crops.

Many early skeptics predicted development of weed resistance to glyphosate and other herbicides used with HT crops, and insect resistance to PIP insecticides. But, resistance occurs as a natural adaptation for survival and has continuously led to the need for more powerful pest control methods in agriculture. Very strong weed resistance was experienced in growing HT canola in parts of Canada, also suggesting triple stacking of genetic resistance from exposures to previous sequences of herbicides. Canola itself, volunteering from previous plantings, could be a resistant super-weed if a locality switched to growing other crops. Various antibiotech concerns, and activist groups and their causes, can be found on the Internet. Glyphosate has been registered as an herbicide for nearly 30 years and is claimed to have attracted less resistance than other herbicides.

In early days of transgenic crop research, various scientific groups passed resolutions that extreme care should be taken to avoid biotech materials escaping into the environment. But, handling transgenic seeds in open fields by uninstructed personnel is far different than aseptic microbe transfers between test tubes in laboratories by trained technicians. Broad contamination of seed supplies and germplasm resources by transgenic sequences has become major concerns. Although unplanted borders around experimental and seed increase plots are recommended to reduce cross-contamination by wind-blown pollen, occasional winds may be stronger and pollinating insects may fly farther than expected. Floods can carry seed to other fields. Bulk seed can spill from trucks, sprout along roadways and fencerows, and the resulting plants pollinate related plants and later plantings. Inadequate cleaning of seed handling, storage, and transporting equipment can contaminate other crops (even at low levels of <0.5%) and spread transgenic seeds that eventually germinate and pollinate their own kind or wild relatives.

Protecting gene pools at *centers of diversity* (locales where a wild plant was first domesticated) from contamination by transgenic sequences has become a major concern among plant scientists. Farmers growing soybean for the organic foods market have difficulty finding planting seed uncontaminated by transgenic sequences and keeping the crop genetically uncontaminated during growth [11].

Concerns exist about potential gene flow problems arising from producing nonfood products such as industrial chemicals and pharmaceuticals in transgenic plants grown in open fields. The inability of current practices to keep such materials out of the food supply was demonstrated by the StarLink® corn incident. The EPA approved sale of StarLink, a B.t.-type corn developed by Aventis CropScience for animal consumption and industrial production of ethanol, in 1998. However, it was not approved for human consumption because of potential allergenic effects. In 2000, a coalition of environmentalists sent a collection of corn food products to a private laboratory and paid for its testing. In September, newspapers reported that StarLink corn was detected in taco shells sold in grocery stores, and Aventis suspended sales soon after. Although only about 0.4% of the United States corn crop had been planted to StarLink in the most popular of its 2-year cultivation, and much of the crop remained in feed channels, it had broadly contaminated the United States food corn supply and some export lots. Many corn-based foods were recalled because of concerns about contamination, and considerable unrest occurred throughout the food processing and exporting industries. Costs to Aventis for numerous analytical bills experienced by processors and traders holding corn in storage, and buying back the remaining StarLink corn and recalled products, reportedly, were approximately $100 million. No evidence was found that StarLink had produced an allergy reaction in any person. The EPA stopped granting split registrations for genetically engineered crops [12]. The concern is that far greater damages might occur if a more noxious chemical entered the food or feed supply through misplaced shipments of a common-looking transgenic crop.

The feasibility of producing plant made pharmaceuticals (PMPs) has been demonstrated with common crops such as corn in a practice sometimes called *biopharming* or simply *pharming*. Production of vaccines, pharmaceuticals, or their precursors in plants, rather than by animal cell culture or transgenic animals, has the advantage of avoiding potential transfer of bacterial or viral diseases. The need for greatly increased oversight in this emerging industry was demonstrated by the Prodigene incident in 2001. The company planted corn, genetically modified to produce pharmaceutical components, at various field sites. The following year, conventional soybean was grown at one of the sites. Seed from the previous year's experimental crop germinated as volunteers, and the corn plants were harvested with the soybean. Pieces of the transgenic corn plants were found with the soybean in elevators, and USDA inspectors quarantined 500,000 bushels of soybean [12]. The pharmaceuticals were in the corn plant trash and not in the soybean seed. It is difficult to predict what may have happened if commercial practices had run their course. If the soybean was extracted for oil, it may have been cleaned and the trash (including corn stalk pieces and any corn seed) sold for feed use. Depending on digestibility, the active compound may have entered the milk or meat supply. A public health disaster did not occur, but some feel the potential for misdirection of intended drugs has been demonstrated.

Over 85% of domestic foods are estimated to contain biotech crops. Whether, or how, to label GMO-containing foods has been an issue since their initial introduction. The EPA, which rules on safety and use of herbicides and pesticides, and the FDA, have adopted the *substantial equivalence* principle when granting approvals and look for closeness of similarities of genetically modified crops or foods with existing products based on chemical analyses. Acute toxicity and suspected allergenicity tests are run in initial screenings. European and some US scientists prefer multigeneration animal testing, reminiscent of earlier food additives approval practices, before release.

Attempts to tag EPA/FDA-registered GMO foods by product labeling have consistently been denied by the government and courts, apparently on the principle they cannot be identified as different from traditional crops because they already have been declared equivalent. Attempts to limit growing of GMO crops in specific areas also have failed, with state programs overruled as interfering with interstate commerce. The alternative of labeling nontransgenic products as “non-GMO” also has been denied on the principle they cannot be implied to be better.

Additionally, formulated non-GMO foods would have to carry the burden of ensuring that all ingredients were non-GMO, including cheeses, ice cream, and other products, which may contain dried nonfat milk solids from cows treated with recombinant bovine somatotropin (rBST). It would be extremely hard to track such products from originating farms to the point of formulation. Currently in the United States, organic foods, produced under certification programs, appear the most reliable non-GMO option (with non-GMO labeling not allowed). GMO products have been declared kosher for Jews and halal for Muslims, provided they do not contain genes from non-kosher or non-halal sources (e.g., pigs).

In Europe and other countries, contents of GMO ingredients are limited, or labeling may be required under the principle that the “right to know” what is in food is a citizen’s right. Domestic consumer activist groups initially strongly supported GMO labeling. The issue has become less vocal with the passing of years without recognized major health problems, but continues to be challenged as in current marketing of “natural hormone-free milk” and “organic” milk.

Still, diseases do not follow governmental decrees. The United States does not have mechanisms in place for specifically following food ingredients or additives approved by the FDA after they enter the marketplace as occurs with drugs. If chronic negative effects from long-term consumption of transgenic crops occurred across the broad population, they might not be attributed to the cause for years because hardly anyone is watching. The National Research Council has recognized needs for in-market follow-up of certain types of new foods [13] and has suggested legislation.

Drivers and Tools for the Future

The basic driver for the biotechnology evolution is rapid growth of the population and its needs. World population doubled (from 3 to 6 billion persons) between 1960 and 1999. In theory, this meant learning to grow as much food in 39 years as was learned during the first 10,000 years of agriculture. In practice, much more was accomplished, including increased life expectancy and improved health and living standards. A slowdown in population growth is expected. Currently, few forecasts venture beyond a world population of nine billion in 2043, still a sizeable increase. Significant quantities of easily accessed fossil energy (petroleum, gas, and coal) and minerals already have been exhausted. With high population densities, solutions to disposal of wastes can no longer be dilution in air, water, or landfills. But, waste disposal and air and water reclamation, by known means, are energy-intensive and place additional loads on resources. There seems to be little choice but to more completely harness recurring energy sources such as sunshine, wind, falling water, and possibly ocean tides and to produce more recyclable food, textiles, building materials, coatings, plastics, and moldable materials through applied biotechnology. Whereas in the past, chemical processing options often were chosen because they are faster, enzymatic routes are more likely in the future because they require less energy. But we must become better applied biophysical chemists to implement such changes.

Our tools include the already discussed abilities to tailor transgenic plants, animals, microorganisms, enzymes, and pharmaceuticals to our needs. We live at a time when communications and transportation are developing rapidly, and the most promising insurance for peace appears to be an international trade so interdependent that rational nations would hesitate before upsetting the balance and linkages by wars. Confidence in the rapidly globalized trade requires reliable trading standards, uniform analytical methods and quality control practices such as ISO 9001–2000 14001–2004, and appropriate enforcement by exporting nations.

The digital computer age has brought us sophisticated analytical instruments and computation abilities to delve even deeper into basic sciences. It also has enabled close monitoring and feedback control of processes, even in remote inhospitable atmospheres, to ensure that operations, and materials and products storage, are continuing as intended.

Government policies do not always turn out as expected. It long was common practice in some countries to direct uneducated workers to production agriculture and to hand operations in food processing. But modern crop and animal production facilities are capital-intensive and require skilled operators and knowledgeable supervisors. The same is true in crop, animal, and seafood processing, where machines and sorters never tire and prove more effective than hand laborers. Where sanitation is critical, quality usually improves as fewer hands touch the product. Whereas some governments in developing nations have sought to create manual jobs in processing foods for local consumption, traders, available alternate suppliers, and eventually consumers will decide what will be carried to the international marketplace.

Many industries have adopted integrated planning. It is a management by objective (MBO)-type technique, often without a name. Integrated pest management (IPM) has been one of the most successful examples. The term was introduced in 1967 by R.F. Smith and R. van den Bosch and formalized by the US National Academy of Sciences in1969 [14]. The initial objective was to find ways to reduce amounts of insecticides used on cotton. As it developed, implementation included: (1) an exhaustive review of the factors leading to application of insecticides, and opportunities for using other options; (2) removal of refuges such as crop stubbles, which protect insects during winter, and host plants in fencerows; (3) consideration of life cycles of beneficial insects, which prey on the species to be controlled by the insecticide; and (4) realization that major crop damage occurs only after several generations of insect build-up, and relatively little is accomplished by spraying early, except killing beneficial insects that would help in control. The solution is closely monitoring numbers of undesirable insects on the crop and delaying spraying until costs are warranted by benefits. Significant reductions occurred in amounts and costs of insecticides used in applying these simple concepts. The principle has caught on quickly. Today, IPM systems are also used to control insects in schools, libraries, museums of natural history, and operations. But the main principles, (1) review the holistic system and interactions between alternative options and (2) continue to monitor the problem closely, but take remedial action only when benefits warrant the costs, are applicable to other management situations. Such approaches have led to energy savings, reduction of waste disposal costs, improvement of oil yield and stability, and increased profits in the edible fats and oils industry.

Biotechnology Practices in Soybean Production and Processing

A brief summary of biotechnology practices in modern production of soybean and oil follows. Farmers select planting seeds, which have been produced by breeding, hybridization, or transgenic means already described. Most varieties grown are "daylight determinant," and flower and produce seed only when day lengths shorten to their optimum photoperiod. The United States has been divided into ten parallel regions for soybean, each about 100–150 miles wide in latitude. Local dealers stock seed of the maturity group appropriate for their region [15]. The soil is prepared by plowing-harrowing, or disk harrowing, for planting. In "no-till" planting, the seed is drilled into the stubble of the previous year's crop, but an herbicide "burn-down" is applied first if weeds are obvious before planting. Farmers typically inoculate the seed with commercial rhizobia inoculum (*Rhizobium* and *Bradyrhizobium* genera) if soybean has not been grown recently or to refresh inoculation. These bacteria establish a symbiotic relationship and create root nodules on the legume which convert atmospheric nitrogen to ammonia usable by the plant. Efforts to induce nitrogen-fixing nodulation in the grasses (specifically corn) have been unsuccessful, but inoculation has been extended to other legume crops.

Farmers also may deposit fertilizer near the seed when planting, or drill it in the previous fall. Depths of topsoil and moisture levels vary in fields in slightly rolling country. The effects often can be seen by aerial and infrared photography. Programs for mapping fertility levels within fields by global positioning systems (GPS), and adjusting rates of fertilizer application as planting–fertilizing equipment moves through the field, have become available. GPS is also used for variable-rate application of other chemicals [16] and for autoguidance of tractors across fields with four-inch precision in strip tilling and planting of close rows [17].

Plant nutrition, seasonal temperatures, and moisture availability affect final composition of soybean and yield/acre. Plants typically create the protein systems in their seed first and add energy reserves (oil in the case of soybean) later before maturing and dying. If the plant is lightly frosted before maturing of the oilseed, it may reinitiate growth and deposit additional chlorophyll in the seed. This compound is oil-soluble and a very strong light-catalyzed prooxidant, which reduces shelf life of bottled oil. Additional efforts must be made to remove chlorophyll by absorption with bleaching clays and silica gels during oil refining. During the summer, the farmer may decide to speculate on the soybean futures market and commit some of the expected crop at a guaranteed price to ensure recovery of at least critical expenses.

Harvest (combining) of soybean continues in some localities even after light snowfall, although quality (as monitored by free fatty acids increases) and market value decrease. Price is also discounted for water-mottled soybean. Whenever soybean is harvested, drying to the 0.65–0.75 a_W range puts it into a dormant state and maximizes the remaining storage life. Drying can occur at the farm, at elevators, or at a processor's holding facilities. Seeds continue to respire in the dormant state and must be aerated, the amount dependent on storage temperature.

By law, organic foods should not be made from transgenic seed. Organic growers cannot use chemical fertilizers, herbicides, or insecticides, but "organic pesticides" are available. Standards may exist about acceptable previous crops on the same soil. Precautions must be taken to ensure the crop is not contaminated with transgenic seed and is identity-preserved (IP) during trading and shipping. The Agricultural Marketing Service of USDA supervises organic foods as a marketing alternative and regulates the National Organic Program.

Although soybean is the world's major oil currently, it is the secondary product after feed meals and food proteins in percent of weight yield and value per bushel processed. Oil extraction processes should not degrade the more valuable protein fractions, which usually is not a problem with soybean processed to make feed protein meals. Enzymatic degradations during extraction and processing of oils can include: (1) a variety of lipases, which can cleave free fatty acids from triglycerides and reduce yield of saleable neutral oil; (2) phospholipases, which can render the phosphatides (lecithins) water-insoluble, difficult to remove when refining, and shorten oil frying life; and (3) two lipoxygenases which can cause off-flavors in refined oils. Maintaining the seed at low water activity, equivalent to about 10–11% during processing, is helpful in slowing enzyme activity. The ideal approach would be to inactivate the enzymes very early in soybean processing, and effectiveness of such processing already has been demonstrated in small extraction plants.

Enzymes have been used to assist oil extraction and in *degumming* (phosphatides removal), splitting fatty acids from trigly-cerides, *interesterification* (rearranging fatty acids on triglyceride molecules), and preparation of specialty oils. These processes are described later in this chapter.

Introduction to Lipids

Fats and oils predominantly *are triesters* (triglycerides, triacylglycerols, TAG) of glycerol and aliphatic fatty acids containing up to 22 carbon atoms. *Waxes are esters of long-chain fatty acids,* usually containing 24–28 carbon atoms, with long-chain primary alcohols (16–36 carbon atoms) or with alcohols of the steroid group [18].

Fats and oils are members of a broader group of chemical substances called *lipids*, which has been classified by the National Research Council into: (1) nonpolar lipids, including esters of fatty acids (triacylglycerols and cholesteryl esters) that are virtually insoluble in water, but soluble in most organic solvents and enter metabolic pathways only after hydrolysis; and (2) polar or amphipathic lipids, including fatty acids, cholesterol, sphingolipids, and glycerolphospholipids (mainly lecithins). The term phospholipids (phosphatides) includes lecithins and sphingomyelins [19]. Other minor natural compounds, also extracted by low-polarity organic solvents, include fat-soluble vitamins, colors, and flavors.

Fats and oils have major roles in human nutrition. They are concentrated dietary sources of energy, providing approximately 9 kcal/g when metabolized compared with 4 kcal/g for carbohydrates and proteins, and account for about 36% of domestic caloric intake per capita [19]. Dietary lipids also can provide essential molecular structures that are synthesized by the body into compounds required for selective functioning of cell membranes and regulation of life processes.

Fats and oils modify product texture in preparation of foods, serve as heat transfer media in food frying, carry flavors, colors, and oil-soluble vitamins, improve mouthfeel, provide a sensation of product richness, and induce satiety. They are used as energy sources in feeds for domesticated animals, and as components of many industrial products, including soaps and detergents, lubricants, plastics and protective coatings, and printing inks, and as carriers of pesticides for aerial spraying, for controlling grain dust, and as feedstocks for manufacturing chemicals. Considerable public interest has developed in the last two decades in replenishable biodegradable carbon sources and in liquid fuels such as biodiesel.

Recent Fats and Oils Industry Changes

Although many chemistry and processing principles have long been established, the industry has undergone major changes in the last three decades, with many starting in the mid-1980s. Rising costs of energy have led to more efficient equipment designs and to the installation of heat recapture systems throughout modern extraction plants and refineries. In the United States, Occupational Safety and Health Protection Agency (OSHA) regulations to prevent injuries and protect the health of workers have led to increased use of safety guards, dust collection systems with shrouding of equipment and improved ventilation, oversight of workers entering dangerous areas such as bins, and periodic checks for hearing loss. Redesigning and retrofitting the equipment was expensive. Process control computers became available concurrently in the early 1990s, and many companies chose instead to install robots in health- and safety-risk areas and to automate processes for operation from control rooms. The few people now seen on extraction plant and refinery floors are mainly repair and cleaning personnel, with hardly any “operators.” With computers making process adjustments, product quality, defined as “uniformity,” typically has improved.

Environmental Protection Agency (EPA) air emissions regulations have led to increased dust controls and to reduced solvent losses in extraction plants. Regulations on the discharge of polluting process streams into public waterways have led to containment and treatment where required of rain runoff from grain storage and processing properties and even from employee parking lots. Silica gel adsorption (modified caustic refining) processes have been developed to avoid production of waste waters in refining oils and problems associated with their disposal. Similar changes are being adopted at various rates throughout the world.

The effects of economic development loans and private investments in developing countries in the 1960–1990 era have matured and are changing the global fats and oils industry. World production of soybean has increased 7.5 times since 1960, with approximately 44% of the world’s crop now entering global trade as soybean, meal, and oil [20].

The growth of palm oil production has been even more spectacular. Hardly known as a crop after World War II, palm oil production increased nearly 21 times since 1960. World production of palm oil has surpassed that of soybean oil. Currently, the total world production of soybean oil is 42.00 million metric tons, whereas palm oil is 47.47 million metric tons (2010/2011). This has brought equatorial countries such as Malaysia and Indonesia into the group of leading edible oil producing and exporting nations. Establishment of trading rules, product definitions and standards, and analytical procedures for the erupting world market also has been part of the technical progress.

Strong competition in the world's oilseeds market, and concerns about the environment, have refocused interest in the United States on *nonfood-nonfeed uses* of crops, including biodegradable applications, renewable hydrocarbon sources, including liquid fuels, and *chemurgy,* the use of agricultural crops as chemicals feed stocks [21]. Nor is the United States alone in the current movement. Germany, Hungary, France, and other European countries, whose climates are too cold for raising soybean or oil palm but can raise rapeseed, recently have installed biodiesel production facilities.

AOCS, an Information Source

This chapter cannot summarize advances in all related technologies and concentrates on current major fats and oils extraction, refining, and utilization practices. The reader may need to browse the Internet for commercial information and computer-based technical abstract services for research reports. Much of the world's research on fats and oils processing and utilization is reported in the *Journal of the American Oil Chemists' Society* (*JAOCS*) published by the AOCS Press, an activity of AOCS (An International Society for the Science and Technology of Fats, Oils and Related Materials), Champaign, Illinois. The AOCS Press also publishes proceedings of selected Society conferences, a broad variety of related books, and the following journals: *INFORM*, a monthly business, news, and scientific publication addressed to professionals interested in the science and technology of fats and oils, surfactants, detergents, proteins, oleochemicals, and related substances; *Oil Mill Gazetteer,* a monthly news magazine for oil extraction and refining plants; *Lipids,* a monthly journal on basic chemistry and nutrition of lipids; *Journal of Surfactants and Detergents* (*JDS*)*,* a quarterly science and news journal on the practical and theoretical aspects of oleochemical and petrochemical surfactants, soaps, and detergents.

The Technical Services function of AOCS establishes, revises, and annually updates "AOCS Methods," the *Official Methods and Recommended Practices of the American Oil Chemists Society* [22] for fats, oils, and soap technology; *Spanish AOCS Methods,* a Spanish translation of the more commonly used AOCS Methods; and *Physical and Chemical Characteristics of Oils, Fats and Waxes.* Leaders of the methods development committees coordinate closely with AOAC International (formerly the Association of Official Analytical Chemists). AOCS Methods are recognized as "Official Methods" in US FDA activities and when litigation becomes necessary in industry trade. Additionally, the Technical Services function operates a Laboratory Proficiency Program (formerly the *Smalley Check Sample Program*) and oversees distribution and statistical analysis of 30 different series of basic laboratory quality assurance/quality control test samples. Certification as AOCS Approved Chemists, or as AOCS Certified Laboratories, and successful participation in the Laboratory Proficiency Program, is expected for industry arbitrators and referees.

Nutrition and Health Implications

Comments on fats and oils in nutrition and human and animal health are limited in this chapter, primarily because of frequent changes in advice given, but regulations are summarized. Three classes of foods exist: generally recognized as safe (GRAS) foods and ingredients, food additives, and food supplements. All are monitored by the FDA, but at different levels of review; meat products are monitored by the Federal Safety Inspection Service (FSIS) of the USDA. Within the FDA, foods are under the jurisdiction of the Center for Food Safety and Applied Nutrition (CFSAN), and animal feeds under the Center for Veterinary Medicine (CVM). Readers are referred to other authorities [23, 24], and FDA and USDA Web-sites, for more details about US food and feed regulatory systems.

The US government does not guarantee the safety of foods; it is the responsibility of respective manufacturers, and ultimately is enforced by torts litigation. USFDA-CFSAN decides whether a food or ingredient is listed as GRAS (useable without restriction), or as a food additive (useable in limited applications and amounts), on the basis of thorough review of applications and may request additional tests for proposed additives. *Food supplements* (vitamins, minerals, botanicals, plant extracts, etc.) may be marketed directly by the manufacturer, but the FDA has the power to check on underlying safety data and to remove unsafe products from the market. Food supplements also include *nutraceuticals, functional foods,* and *designer foods*—groups of products consumed for expected health benefits aside from nutrition and medication. These products became a major growth market in the last decade, but at times have been launched with limited documentation, and possibly knowledge about effectiveness. The FDA earlier held that food and drugs were separate classifications and did not allow health claims to be made for food products until the 1990s. Congressional legislation, including the Nutrition Labeling and Education Act of 1990, Dietary Supplement Health and Education Act of 1994, and the Food and Drug Administration Modernization Act of 1997, led to today's practices of allowing health claims, provided specific requirements are met [25].

A basic problem when considering food safety is the wide range of genetics, physiological ages, and individual health status in the United States population of nearly 300 million. Almost any product may prove adverse for an unpredictable portion of the population, especially if consumed in large quantities.

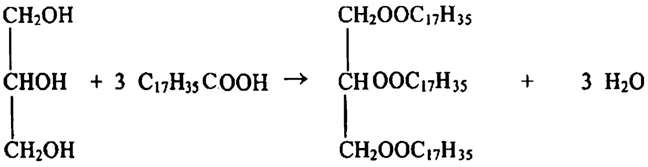
Many of the extraction, concentration, and purification techniques and equipment used in preparation of nutraceuticals originated in the vegetable oils and proteins processing industries. Byproduct streams often are further purified for this market.

Nomenclature and Molecular Structures

General

The choice of the term "oil" or "fat" usually is based on tradition and the physical state of the material. Generally, *oils* are liquid at ambient temperatures, and *fats* are semisolid mixtures of crystals in oil. *Fats* often are of animal origin (beef tallow, pork lard, and butter fat) or *hardened* (hydrogenated, interesterified, or thermally fractionated) vegetable oils, whereas oils are extracted from plant seeds or tissues or fish. In English-speaking countries outside the United States, oils liquid at room temperature sometimes are called *soft oils,* and those hard or pasty are called *hard oils*. Nutritionists generally use "fats" for solids or liquids.

Over 95% of the weight of most extracted/separated food fats are TAG (triacylglycerols or triglycerides) formed by the enzymatic combination of glycerol (a trihydric alcohol) with three fatty acids also yielding one molecule of water for each ester linkage:



$$\begin{array}{lcl} \begin{array}{l} CH_2OH \\ | \\ CHOH \\ | \\ CH_2OH \end{array} + 3\ C_{17}H_{35}COOH \rightarrow \begin{array}{l} CH_2OOC_{17}H_{35} \\ | \\ CHOOC_{17}H_{35} \\ | \\ CH_2OOC_{17}H_{35} \end{array} + 3\ H_2O \end{array}$$

Glycerol + 3 Stearic acids → Glycerol tristearate + 3 Water
(tristearin)

The reaction is reversible and favored by the presence of moisture and catalysts including lipases, alkalis, and alkaline metals. In the oleochemicals industry, TAG are split by high-pressure steam. Unassociated fatty acids are called free fatty acids or FFA.

Fatty Acids

Fatty acids are the building blocks of TAG. More than 90% of fatty acids have an even number of carbon atoms and are in aliphatic chains ranging from 4 to 22 carbons in length. The major fatty acid synthesis pathway is production of stearic acid (18 carbons) after which separate desaturase systems introduce 1, 2, or 3 unsaturated (double) bonds. Additional enzymes become active in elongating the chain as needed. Shorter fatty acids also are produced. Trace amounts of odd-number carbon fatty acids are found in most fats and also have been synthesized for research purposes. Microorganisms frequently produce odd-number carbon fatty acids, with heptadecenoic (17 carbon) acid a major component of *Candida tropicalis* yeast fat. Up to 8% C_{17} fatty acids have been found in milk and meat fats of ruminants (cattle, sheep, goats) and are of rumen microbe origin.

The names of common fatty acids under several conventions, carbon numbers, and selected properties are shown in Table 34.1. The common (trivial) names of some fatty acids are of long standing and often indicate the initial source studied. As examples: butyric acid is a major component of butter flavor; the 6, 8, and 10 saturated fatty acids have been called the *goaty acids* because they impart the characteristic flavors of goat and sheep milks and their cheeses. The terms *olein* and *stearin* were applied to the liquid and solid fractions, respectively, of tallow separated by pressing in early manufacture of oleomargarine and compounded shortenings. The iodine value (IV) is an indicator of the unsaturation of a fatty acid or fat/oil. It is determined by AOCS Method Tg 1a-64 or Cd 1–25, respectively: the higher the IV, the more unsaturated the fat and the lower the melting point.

Fatty acids sometimes are designated by the number of carbon atoms in the chain, followed by a colon with additional numbers indicating the number of double bonds. In the 18-carbon series, C18:0, C18:1, C18:2, and C18:3 represent *stearic, oleic, linoleic*, and *linolenic* acids, respectively. One- or two-letter abbreviations sometimes are used, with these acids designated as St, O, L, and Ln, respectively.

Under the most common convention, fatty acids are named on the basis of the number of carbon atoms, starting with the terminal carboxyl (-COOH) carbon as number "1."

Table 34.1 Names and characteristics of some important fatty acids

Carbon atoms and abbreviations	Common name	Symbol	Systematic name	Melting point (°C)	Iodine value	Common sources
Saturated fatty acids						
3:0	Propionic	–	Propanoic	−20.8	–	Bacterial fermentation
4:0	Butyric	B	Butanoic	−7.9	–	Milk fats
5:0	Valeric	–	Pentanoic	−33.8	–	Bacterial fermentation
5:0	Isovaleric	–	3-Methylbutanoic	−51.0	–	Dolphin and porpoise fats
6:0	Caproic	H	Hexanoic	−3.4	–	Milk fats, some seed oils
8:0	Caprylic	OC	Octanoic	16.7	–	Milk fats, *Palmae* seed oils
10:0	Capric	D	Decanoic	31.6	–	Sheep and goat milk, palm seed oils, sperm head oil
12:0	Lauric	La	Dodecanoic	44.2	–	Coconut oil
14:0	Myristic	M	Tetradecanoic	54.4	–	Palm and coconut oils
16:0	Palmitic	P	Hexadecanoic	62.9	–	Palm oil, most oilseeds and animal fats
18:0	Stearic	St	Octadecanoic	69.6	–	Animal fats
19:0	Tuberculostearic	–	10-Methylstearic	11.0	–	Tubercle bacillus lipids
20:0	Arachidic	Ad	Eicosanoic	75.4	–	Some animal fats
22:0	Behenic	–	Docosanoic	81.0	–	Peanut and various other seed oils
24:0	Lignoceric	–	Tetracosanoic	84.2	–	Minor amounts in some seed oils
26:0	Cerotic	–	Hexacosanoic	87.8	–	Plant waxes
28:0	Montanic	–	Octacosanoic	90.9	–	Beeswax and other waxes
30:0	Mellisic	–	Triacontanoic	93.6	–	Beeswax and other waxes
Unsaturated fatty acids						
10:1	Caproleic	–	9-Decenoic	–	149.1	Milk fats
10:2	Stillingic	–	2,4-Decadienoic	–	–	*Stillingia* oil
12:1	Lauroleic	–	2-Dodecenoic	–	128.0	Butterfat
14:1	Myristoleic	–	9-Tetradecenoic	18.5	112.1	Some feed fats, milk fats
16:1 (*n*–7)	Palmitoleic	–	9-Hexadecenoic	0.5	99.8	Many fats and marine oils
16:3	Hiragonic	–	6,10,14-Hexadecatrienoic	–	–	Sardine oil
17:1	–	–	9-Heptadecenoic	14.0	–	*Candida tropicallis* yeast
18:1 (t-oleic)	Elaidic	–	9-Octadecenoic	43.7	–	Butterfat
18:1 (*n*–9)	Oleic		9-Octadecenoic	16.3	89.9	Almost all fats and oils
18:1	Petroselinic	–	6-Octadecenoic	30–33	–	Parsley seed oil
18:1 (*n*–7)	Vaccenic	–	11-Octadecenoic	44.0	–	Butterfat, seed oils
18:2 (*n*–6)	Linoleic	Lo	9,12-Octadecadienoic	−6.5	181.0	Most vegetable oils
18:3 (*n*–6)	(gamma) Linolenic	–	6,9,12-Octadecatrienoic	–	–	(Omega-6); Evening primrose, borage, vegetable oils
18:3(*t*)	Eleostearic	–	9,11,13-Octadecatrienoic	–	–	Tung oil
18:3 (*n*-3)	(alpha)Linolenic	Ln	9,12,15-Octadecatrienoic	−12.8	273.5	(Omega-3); Linseed, soybean, canola, other vegetable oils
20:1	Gadoleic	–	11-Eicosenoic	23–24	81.8	Some fish oils
20:3 (*n*–9)	Eicosatrienoic		5,8,11-Eicosatrienoic	–	–	Brain phospholipids
20:3 (*n*–6)	Dihomo-γ linolenic		8,11,14-Eicosatrienoic	–	–	Shark liver oil
20:5 (*n*–3)	EPA		5,8,11,14,17-Eicosapentanoic	−33.5	75.0	Fish, plants
22:1	Erucic	E	13-Docosenoic	34	–	Rapeseed oil
22:2			13,16-Docosadienoic	–	–	
22:5 (*n*–3)	DPA		7,10,13,16,19-Docosapentaneoic	–	–	Fish oils
22:6 (*n*–3)	DHA		4,7,10,13,16,19-Docosahexaenoic	–	–	Cooler climate fish oils
Fatty acids of more unusual structure						
18	Chaulmoogric	–	13,(2-Cyclopentenyl) tridecanoic	68.5	90.5	*Chaulmoogra* oil
18	Malvalic	–	8,9-Methylene-8-heptadecenoic	–	–	*Malvaceae* seeds, cottonseed
18	Ricinoleic	–	12-Hydroxy-9-octadecenoic	5.5	86.0	Castor oil
18	Vernolic	–	12,13-Epoxy-9-octadecenoic	30–31	–	Some *Compositae* seeds
19	Sterculic	–	9,10-Methylene-9-octadecenoic	–	–	*Sterculiaceae* seeds, cottonseed

(continued)

Table 34.1 (continued)

Carbon atoms and abbreviations	Common name	Symbol	Systematic name	Melting point (°C)	Iodine value	Common sources
20	Arachidonic	–	5,8,11,14-Eicosatetraenoic	−49.5	333.5	Lard
20	–	–	5,8,11,14,17-Eicosapentaenoic	–	–	Some fish oil
20	Lesquerolic	–	14-Hydroxy-11-Eicosenoic	–	–	*Lesquerella* seed oil
22	–	–	4,7,10,13,16,19-Docosahexaenoic	–	–	Some fish oil

The terminal letter *e* of the respective alkane hydrocarbon is replaced with *oic* to indicate an acid; thus:

$CH_3CH_2CH_2CH_2CH_2CH_2CH_2CH_3$

Octane

8 7 6 5 4 3 2 1

$CH_3CH_2CH_2CH_2CH_2CH_2CH_2COOH$

Octanoic acid

The suffix *dioic is* used if the acid contains two carboxyl groups.

Occasionally in the literature, the carboxyl unit is regarded as a group substituted for hydrogen. In this case, the number 1 location is moved one position away from the reactive end, and the suffix *carboxylic* acid is added:

5 4 3 2 1

$CH_3CH_2CH_2CH_2CH_2COOH$

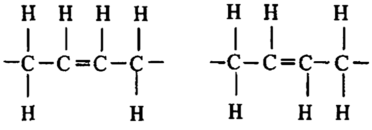
1 - Pentanecarboxylic acid

(hexanoic acid)

(caproic acid)

A double bond between two carbon atoms indicates the site, and possibly type, of hydrogen unsaturation. When double bonds are present, the suffix *anoic* is changed to *enoic, dienoic*, or *trienoic* to indicate the number of bonds. The location of the first carbon in the double bond is indicated by a number preceding the systemic name. Under International Union of Pure and Applied Chemistry (IUPAC) convention, stearic, oleic, linoleic, and linolenic acids are called octadecanoic, 9-octadecenoic, 9,12-octadecadienoic, and 9,12,15-octadecatrienoic acids, respectively.

The 3D geometric configuration of hydrogens at double bonds is indicated on paper by the Latin prefixes *cis* (both hydrogens on one side) and *trans* (hydrogens across from each other). Linoleic acid, with the *cis* configuration in both double bonds, is called *cis*-9, *cis*-12-octadecadienoic acid. Most fatty acids occur in nature in the *cis* form. Oleic acid is in the *cis* configuration and the corresponding *trans* form is called elaidic acid.



cis ***trans***

Increasing the number of double bonds lowers the melting point of the fatty acid from its fully or partially saturated form. Double bonds also are the sites of oxidation initiation on free fatty acids and within triacylglycerols. Creating a *trans* bond by hydrogenation increases the melting point, but not as much as full saturation.

Locations of the double bonds, and especially the last double bond in long-chain polyunsaturated fatty acids, are of special interest. Whereas chemists traditionally count with the carboxyl carbon (-COOH) assigned number “1,” biochemists and nutritionists assign number “1” to the methyl carbon (CH_3^-). Thus, linoleic acid (9,12-octadecadienoic acid), known as C18:2 to a chemist, carries the same trivial name for biochemists but is known as C 18:2 ω-6 or C 18:2 *n*-6 with *omega* or *n* signifying “count from the methyl carbon.” From a nutrition viewpoint, four families of fatty acids (*n*-7, *n*-9, *n*-6, and *n*-3) exist. Members of the *n*-7 and *n*-9 families generally are synthesized by each species as needed, but members of the *n*-6 and *n*-3 families may be dietary essential either because the species is unable to synthesize the fatty acid, or metabolic mechanisms are impaired in specific individuals. Plants, including plankton, are the ultimate source of dietary essential fatty acids (EFAs). Humans and most animals are considered to require linoleic and linolenic acids (C 18:2*n*-6 and C 18:3*n*-3, respectively). Fish vary by species, with carnivorous members like the salmonoids (salmon and trout) requiring EPA (C20:5*n*-3;5,8,11,14,17-eicosapentaenoic acid) and DHA (C22:6*n*-3;4,7,10,13,16,19-docosahexaenoic acid) when raised in captivity [26].

Contents of individual fatty acids in the fat of a species can vary by over 100%. Table 34.2 summarizes general fatty acids contents of the major edible fats and oils, and Table 34.3 does likewise for industrial fats and oils.

On a global basis for all species, oils produced in the tropics are more completely saturated and have the highest

Table 34.2 Fatty acid composition of some edible oils and fats[a]

Source	<14:0	14:0	16:0	16:1	18:0	18:1	18:2	18:3	20:0	20:1	22:0	22:1	24:0	24:1
Almond oil	–	0.0	6.5	0.6	1.7	69.4	17.4	–	–	–	–	–	–	–
Avocado oil	–	–	11.0	3.4	0.7	71.5	12.0	1.5	–	–	–	–	–	–
Barley bran oil	–	0.5	10.8	0.2	1.0	17.8	55.3	4.4	–	–	–	–	–	–
Borage oil	–	–	11.3	–	3.7	16.3	38.1	23.0	0.2	3.9	–	2.4	–	1.4
Buffalo gourd seed oil	–	–	11.8	–	3.5	21.9	60.6	0.0	0.0	–	–	–	–	–
Butter fat	23.8	8.2	21.3	1.8	9.8	20.4	1.8	1.2	–	–	–	–	–	–
Canola oil[b]	–	–	4.8	0.5	1.6	53.8	22.1	11.1	1.1	1.5	0.3	0.1	0.1	–
Cherry pit oil	–	–	7.8	0.4	2.4	43.9	44.8	0.5	0.7	–	–	–	–	–
Cocoa butter	–	0.1	25.4	0.2	33.2	32.6	2.8	0.1	–	0.0	–	–	–	–
Coconut oil	58.7	16.8	8.2	–	2.8	5.8	1.8	–	–	–	–	–	–	–
Corn oil	0.0	0.0	10.9	–	1.8	24.2	58.0	0.7	–	–	–	–	–	–
Cottonseed oil	–	0.8	22.7	0.8	2.3	17.0	51.5	0.2	–	–	–	–	–	–
Evening primrose oil	–	–	8.5	–	2.5	8.5	72.5	11.0	–	–	–	–	–	–
Fish (manhaden) oil	–	9.6	20.5	12.6	3.3	11.0	0.7	1.6	0.3	–	–	0.8	–	–
Grapeseed oil	–	0.1	6.7	0.3	2.7	15.8	69.6	0.1	–	–	–	–	–	–
Illipe butter	–	–	23.7	–	19.3	43.3	13.7	–	–	–	–	–	–	–
Lard	0.5	1.3	23.8	2.7	13.5	41.2	10.2	1.0	–	1.0	–	–	–	–
Lupine oil	–	–	8.3	–	2.5	55.0	17.7	9.3	–	–	–	–	–	–
Macademia nut oil	–	0.6	8.5	21.7	3.7	56.0	1.7	–	–	1.4	–	–	–	–
Mango kernel oil	–	–	7.6	–	36.0	49.4	5.0	0.5	1.4	–	–	–	–	–
Mustard seed oil	–	0.1	1.9	0.3	0.1	17.7	9.1	0.5	0.6	3.91	1.8	55.1	0.2	1.9
Okra seed oil	–	0.2	33.7	0.6	3.3	17.9	42.2	0.2	0.1	–	0.2	–	–	–
Olive oil	–	0.0	11.0	0.8	2.2	72.5	7.9	0.6	–	–	–	–	–	–
Palm oil	0.1	1.0	43.5	0.3	4.3	36.6	9.1	0.2	–	0.1	–	–	–	–
Palm kernel oil	54.2	16.4	8.1	–	2.8	11.4	1.6	–	–	–	–	–	–	–
Peanut oil	–	0.1	9.5	0.1	2.2	44.8	32.0	–	–	1.3	–	–	1.8	–
Rapeseed oil[c]	–	–	1.7	–	0.9	12.3	12.7	7.6	1.2	5.8	0.9	59.4	0.5	1.6
Rice bran oil	–	0.7	16.9	0.2	1.6	39.1	33.4	1.6	–	–	–	–	–	–
Safflower oil	–	0.1	6.2	0.4	2.2	11.7	74.1	0.4	–	–	–	–	–	–
Safflower oil[d]	–	–	4.8	–	1.3	75.3	14.2	–	–	–	–	–	–	–
Sal seed oil	–	–	5.3	–	34.0	49.1	3.8	3.3	4.0	–	–	–	–	–
Sesame oil	–	–	8.9	0.2	4.8	39.3	41.3	0.3	–	0.2	–	–	–	–
Shea butter	1.7	0.1	4.4	0.1	38.8	43.5	4.9	0.3	–	0.0	–	–	–	–
Soybean oil	–	0.1	10.3	0.2	3.8	22.3	51.0	6.8	–	–	–	–	–	–
Sunflower oil[e]	–	–	7.0		5.0	19.0	68.0	1.0	–	–	–	–	–	–
Sunflower oil[f]	–		4.0	–	5.0	65.0	26.0	–	–	–	–	–	–	–
Tallow	0.9	3.7	24.9	4.2	18.9	36.0	3.1	0.6	–	0.3	–	–	–	–
Teaseed oil	0.1	0.1	17.5	0.5	3.1	49.9	22.2	0.7	–	1.0	–	–	–	–
Tomato seed oil	–	0.2	15.0	0.5	4.4	21.9	50.8	2.3	–	–	–	–	–	–
Walnut oil	–	–	7.0	0.1	2.0	22.2	0.4	52.9	10.4	–	–	–	–	–
Wild cucurbit oil	–	–	19.0	–	–	34.0	47.0	15.0	–	–	–	–	–	–

[a] These are average values from recent years' crops

[b] Low-erucic-acid variety rapeseed

[c] High-erucic-acid variety

[d] High-oleic variety

[e] Maturing in coller climates

[f] Mid-oleic sunflower seed

melting points, with melting points decreasing with distance from the equator (in north and south latitudes). Many dietary fatty acids are transposed from plant or plankton feed sources to body tissues, and fish oils from the Arctic and Antarctic species generally melt at lower temperatures than those from the tropics. Many oilseed species are *daylight determinate;* that is, they require a specific number of hours of sunlight to blossom and produce seed. Thus, a plant

Table 34.3 Fatty acid composition of some industrial oils and fats[a]

Source	<14:0	14:0	16:0	16:1	18:0	18:1	18:2	18:3	20:0	20:1	22:0	22:1	22:2	24:1
Caster oil[b]	–	–	1.1	0.2	1.0	3.3	3.6	0.32	0.4	–	–	–	–	–
Chinese tallow	1.3	2.1	65.0	–	4.4	22.5	0.8	–	–	–	–	–	–	–
Crambe oil	–	–	2.0	0.4	0.4	16.9	8.6	6.4	0.5	3.2	2.0	57.2	–	1.4
Crepsis foetida oil[c]	–	0.1	4.8	–	2.9	4.3	27.8	–	–	0.3	–	–	–	–
Croton oil	2.5	5.4	6.2	0.2	3.2	15.8	49.4	3.0	2.9	8.9	0.2	0.6	–	–
Cuphea oil	–	76.4	7.8	2.4	–	0.7	5.9	6.9	0.1	–	–	–	–	–
Jojoba oil	–	–	1.0	–	–	9.0	–	–	–	70.7	–	16.3	–	3.0
Lesquerella seed oil	–	1.4	1.4	2.0	17.6	7.9	13.2	–	52.5	–	–	–	2.7	–
Linseed oil	–	–	5.3	–	4.1	20.2	12.7	53.3	–	–	–	–	–	–
Meadowfoam oil	–	–	–	–	–	–	–	–	–	64.5	–	18.5	13.5	–
Neatsfoot oil	–	0.7	16.9	–	2.7	64.4	2.3	0.7	0.1	–	–	–	–	–
Oitica oil[d]	–	–	7.0	–	5.0	6.0	–	–	–	–	–	–	–	–
Rapeseed[e]	–	0.1	2.6	0.3	0.9	11.2	12.8	8.6	–	7.5	–	48.1	–	–
Rubber seed oil	–	0.2	19.1	–	17.8	24.5	30.5	2.4	–	0.1	0.9	0.4	–	–
Stokes aster oil[f]	–	–	2.8	–	0.9	7.0	16.5	–	–	–	–	–	–	–
Tall oil	–	–	–	–	–	50.0	7.0	41.0	–	–	–	–	–	–
Tung oil	–	–	3.1	–	2.1	11.2	14.6	69.0	–	–	–	–	–	–
Veronia seed oil[g]	–	–	2.7	–	1.3	2.0	8.8	0.4	–	–	–	–	–	–
Whale oil	–	3.3	8.1	26.9	1.1	33.3	–	–	–	10.9	–	2.2	–	–

[a] These are average values from recent years' crops
[b] Contains 89.2% ricinoleic and 1.4% dihydroxystearic acids
[c] Contains 59.8% crepenynic acid
[d] Contains 78.0% licanic acid and 4.0% hydroxy acids
[e] High-erucic-acid variety
[f] Contains 71.3% vernolic acid
[g] Contains 78.5% vernolic acid and 5.8% hydroxy fatty acids

grown from seed adapted to another latitude may sprout and produce much foliage, but not flower and produce seed. Earlier soybean and sesame varieties were mainly light determinate, but indeterminate varieties also exist now.

Each oil has its unique properties and history, as demonstrated by sunflower seed. Sunflowers generally are indeterminate, with the same seed productive in the prairie provinces of Canada and in Mexico. When polyunsaturated oils were promoted in the 1960s and 1970s, United States processors of table oils (for salads and light cooking) purchased only sunflower seed grown north of the 39th parallel in the United States and Canada to maximize the polyunsaturated fatty acids (PUFA) content of their products. Fall seed maturation temperatures in the Northern climates are lower, resulting in higher PUFA and lower monounsaturates (oleic) contents than sunflower seed maturing at the same time in Texas. However, the higher oleic acid oil of Texas seed is less susceptible to oxidation and produces deep-fried foods and snacks with longer shelf lives. If growers in areas of Texas, capable of raising two crops annually, want to produce high-polyunsaturated oil, they merely have to time their second planting for seed to mature in the cooler winter months [27, 28].

Several important changes occurred in the early 1980s. First, oil processors realized that a significant market existed for high-stability oils and began seeking higher oleic acid content oils. Second, the medical community and nutritionists realized that, although PUFA do not cause cancer, they might promote growth of existing cancer cells more than monounsaturated fatty acid (oleic). Emphasis was changed from encouraging consumption of polyunsaturated vegetable oils to reducing fat intake in general, with recommendations that no more than 30% of dietary calories come from fats, of which no more than 10% are saturated (animal, tropical, or hydrogenated) fats [19], with some nutritionists believing that PUFA also be no more than 10%, essentially leaving at least 10% for monounsaturated fat.

The world's traditional monounsaturated fat is olive oil (~70% oleic acid), historically consumed in countries surrounding the Mediterranean Sea, thus the term *Mediterranean Diet*. However, olive oil is expensive and too limited in supply to satisfy the growing popularity of high-oleic acid oils. Thus, US table oil processors turned to importing high-oleic acid content canola oil from Canada. A mid-oleic sunflower seed oil (~65% oleic acid), grown in the northern states, was introduced to the fried foods industry in 2000. High oleic acid varieties of safflower and peanut have been introduced, and a transgenic high-oleic acid soybean was patented in late 2001. Two nontransgenic lines of high-oleic acid sunflowers, containing 80–82 and 90–92% oleic acid, were developed in the early 1980s, and their oils were marketed as feedstock for oleochemicals production. The venture was not commercially successful. Interest in nontraditional fats/oils sources, including newly domesticated crops, forest oilseeds, bacteria, yeasts, molds, and algae, has increased in recent years. The reader may consult the references and later publications for unusual fatty acids and their occurrence in various sources [29, 30].

Triacylglycerols

Glycerol esterified with one, two, or three fatty acids is found in nature and can be made commercially. The designations monoacylgly-cerol, diacylglycerol, and triacylglycerol (TAG), respectively, now are encouraged in the scientific literature, but the older mono-, di-, and triglyceride terminology is used in commerce. As the number of fatty esters on the glycerin "backbone" decreases, the compound becomes more polar and functionally effective as a surfactant. Mono- and diacylglycerols are further described throughout this chapter.

Triacylglycerols are named in various ways. For example, unsaturated fatty acids sometimes are indicated as U and the saturated as S. If glycerol is completely esterified with stearic acid, the resulting monoacid TAG may be designated as SSS, or, more descriptively as StStSt, tristearin, tristearoylglycerol, or glycerol tristearate.

If more than one species of fatty acid is present, its relative location on the glycerol may be important to its functionality, enzyme susceptibility, and oxidation stability of the fat/oil. Several conventions have been developed to specify arrangements of fatty acids on the glycerol molecule (if known). To avoid confusion from inversion of the 1 and 3 carbon positions, hierarchies have been established to designate the number 1 carbon under the R/S (*rectus-sinisturs*) system [31] and the *sn* (stereospecific numbering) system [32]. In the R/S system, the longest chain fatty acid is assigned to the 1 position, the second longest to 2, and the shortest to 3.

If the positions of fatty acids on the TAG molecule are known, the *sn* system is preferred for identifying their locations: *sn* immediately before the word glycerol, as in 1-stearoyl-2-oleoyl-3-myristoyl-*sn*-glycerol, identifies the respective fatty acids in the 1, 2, and 3 positions; the term *rac* (racemic mixture), as in *rac*-StOM, identifies the middle acid in the 2-glycerol position and the remaining fatty acids are equally divided between the *sn*-1 and *sn*-3 positions; and the term β, as in β-StOM, identifies the middle acid in the 2-glycerol position, but distribution of the other two acids is unknown [33]. Despite international efforts to standardize chemical terminology and abbreviations, a variety is still used, especially in *cis* and *trans* notations.

Table 34.4 Relative rates of oxidation and hydrogenation of fatty acid chains ([a]Modified from Beckman [34])

Fatty acid	Iodine value	Relative oxidation rate	Relative hydrogenation rate
Stearic (18:0)	0	1	0
Oleic (9*cis*-18:1)	90	10	1
Linoleic (9*cis*, 12*cis*-19:2)	181	100	20
Linolenic (9*cis*, 12*cis*, 15*cis*-18:3)	274	150	40

Oxidation

As shown in Table 34.4, oxidation and hydrogenation reactivity of a fatty acid, in free form or as part of a TAG, increases with the number of double bonds [34]. Oxidation of fatty acids and TAG (aldehyde formation, breakdown into shorter chains, and crosslinking to form polymers) is initiated at double-bond sites. However, linoleic acid, C18:2 *n*-6, does not decompose into a mixture of C9, C3, and C6 compounds. Instead, as the molecule starts degrading, positions of the double bonds migrate and provide many opportunities for splitting. Over 250 different breakdown compounds have been found [33, 35–38]. Part of the confusion about oxidation reactions is related to the type present, with light-induced photosensitized singlet oxygen oxidation the fastest [39]. For this reason, removal of photo-sensitizers such as chlorophylls and porphyrins during refining and use of oil is extremely important.

The extent of prior oxidative activity in a fat sample may be estimated by the following tests: Peroxide Value (PV), a titrametric method (AOCS Method Cd 8–53); thiobarbituric acid test (TBA), which measures the presence of malonaldehyde; Anisidine Value-Totox; Kreis test; oxirane test; total and volatile carbonyl compounds content; chromatographic analysis; ultraviolet spectroscopy; fluorescence; and especially organoleptic evaluation. However, the resulting peroxides are unstable and decompose. Peroxide values are not cumulative and, alone, are not always indicative of the extent of earlier oxidation. An oil sample may be starting to degrade or may already have passed through a serious oxidation cycle; thus, age and history of the sample should be considered in forming a conclusion.

Predisposition (susceptibility) to oxidation can be estimated for oils and fats by the Active Oxygen Method (AOM) and Oil Stability Index (OSI) for oils and fats and by the Schall oven test or the oxygen bomb method for fat-containing products. The stability of triglycerides formerly was determined by the AOM (AOCS Method, Cd 12–93), in which heated air was bubbled through a heated liquid sample of the oil or fat, and the number of hours for the sample to reach 100 milliequivalents (meq) of peroxide was recorded. The AOM procedure was put in *surplus status* (still legal but not preferred) in 1997 and the focus then shifted to the Oil Stability Index (OSI) (AOCS Method, Cd 12b-92). However, AOM is still used in product and purchase specifications and in reports. In the OSI procedure, heated air is bubbled through heated liquid triglycerides and is scrubbed on exiting in a bath of deionized water, whose conductivity is continuously monitored spectrophotometrically. Absorption of polar degradation products is noticed immediately. Whereas the AOM method determines the time for the triglyceride to reach a specific level of oxidation (100 meq peroxide value), OSI determines the *induction period* (time required to exhaust the antioxidant properties), but not oxidation progress in the oil. OSI values always are lower (less time) than AOM values.

Biohydrogenation and Conjugated Linoleic Acids

Rumen microorganisms, in cattle, sheep, and other ruminants, hydrolyze exposed TAG in feeds and metabolize the glycerol. The unsaturated FFA are especially toxic to microorganisms and are *biohydrogenated* by enzymes in microbial cell walls to prevent permeation into their cells, possibly by raising their melting points above rumen temperature. Although polyunsaturated fatty acids can be reduced to stearic acid, rumen hydrogenation often is incomplete. It yields products such as monounsaturated vaccenic acid (*trans*-11-octadecenoic acid; *trans*-11 C18:1) and conjugated linoleic acids (CLA) in which the methylene carbon in the C9 through C13 (—C═C—C—C═C—) sequence of linoleic acid is eliminated to form a *conjugated sequence* of double bonds (—C═C—C═C—). The conjugated sequence can appear in many positional isomers along the 18-carbon chain, usually between carbons C6 and C14, resulting in a variety of CLA. Two of the CLA, 9 *cis*-11 *trans*-18:2 (rumenic acid), and 10 *trans*-12 *cis*-octadecadienoic acid have shown physiological activity: the 9,11-*ct* CLA isomer displaying antimutagenic properties, and the 10,12-*tc* CLA isomer displaying antifat deposit and cholesterol-modulating properties [40, 41].

Although more research is needed to document consistent modes of action, capsules containing CLA produced by microbial fermentation are sold as food supplements. CLA in ruminant products are considered natural, and levels of up to 1.5% of the fat in beef and 6% of the fat in cheeses made from spring pasture milk have been reported.

Vaccenic acid and CLA conjugation were once thought not to exist in plant oils, but with improved analytical instruments and methods, are increasingly reported, especially in spices. Conjugation also is an early step in chemical hydrogenation and in initiation of oxidative degradation of fats and oils. Recent research further supports that the majority (78%) of c9,t11-CLA in cow's milk is produced exogenously from body fat vaccenic acid in the mammary gland by *delta*-9 desaturase [42].

Recently, in 2008, the United States Food and Drug Administration (FDA) approved CLA containing conjugated *cis* and *trans* double bonds as a food ingredient. The announcement states that CLA is generally regarded as safe for use in foods. This clears the way for CLA, to be used as an ingredient in foods and beverages sold in the United States.

Fatty Acid and Triacylglycerol Melting Characteristics

Body temperatures are approximately 37°C/98.6°F for humans; 38.5–39.5°C/101–103°F for domestic animals; and 40.5–41.5°C/105–107°F for poultry. As shown in Table 34.1, the longest saturated fatty acid that is fluid at these temperatures is caproic (C10:0). All longer saturated fatty acids must be accompanied by lower melting unsaturated fatty acids as in a TAG structure to be fluid. The C18 oleic acid (*c*-9-octadecenoic acid) has a melting point of 16.3°C, the *trans* isomer elaidic acid (*t*-9-octadecenoic acid) melts at 43.7°C, and the biohydrogenated product *trans*-vaccenic acid (*t*-11-octacedenoic acid) melts at 44°C [40].

Fat digestion and absorption in mammals occur in the small intestine and mainly consist of emulsification of TAG by bile salts, lecithin, and agitation, followed by pancreatic lipase cleavage at the 1,3 positions to produce free fatty acids and 2-monoglycerides. These are carried by the bile to surfaces of the microvilli, where they are absorbed through the membranes into the intestinal lymph. While passing through the intestinal epithelial cells, the fatty acids and monoglycerides are resynthesized into new TAG that are transported, mainly in lymph chylomicrons, through the thoracic lymph duct which empties into the circulatory blood [43]. (Although the free fatty acids may be rearranged onto new TAG, fatty acids in the 2-position remain as before.) Animals have various mechanisms for ferrying fatty acids during absorption at the small intestine, and moving fatty acids and triglycerides through their circulatory systems. Readers interested in details are referred to books on medical physiology.

The positioning of fatty acids on the triglyceride chain follows several patterns. If unsaturated fatty acids are limited in availability, nature tends to place them in the number 2 position to obtain the lowest melting point for a plant's triglycerides. If more plentiful, the preferred positions will be 2 and 1 or 3 or both [44]. In contrast, saturated fatty acids are more likely to be in the 2 position in fats of warm-blooded animals, with palmitic acid in the 2 position in (pork) lard as an example.

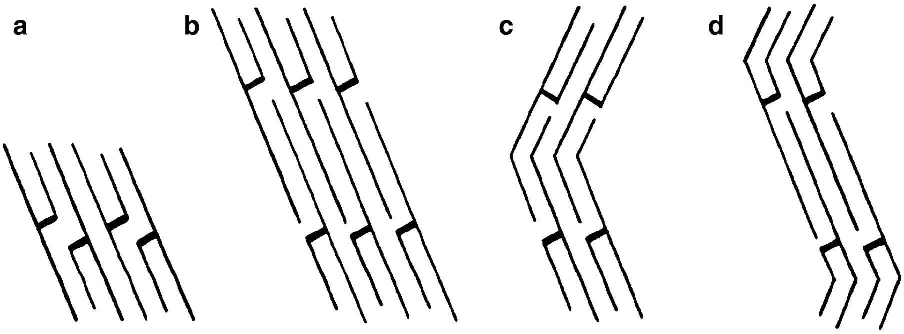


Fig. 34.1 Drawings of two- and three-chain triacylglycerol layers: (**a**) saturated monoacid SSS-type; (**b**) saturated symmetrical PSP-type where 2 chain differs from 1 and 3 chains in length; (**c**) symmetrical POP-type where 2 chain is unsaturated; and (**d**) symmetrical OPO-type where 1 and 3 chains are unsaturated. Crystals grow in bilayer units

Factors affecting the melting points of specific fat samples include: the types of fatty acids present (lengths of the fatty acids chains, number and location of *cis* and *trans* double bonds in the chains), location of specific fatty acids on the glycerol, compatibility of the different TAG in the mixture, and types of crystals present. Melting points increase with chain length. *Trans* fatty acids always have higher melting points than their *cis* counterparts for any chain length. Where only one double bond exists in a fatty acid, as in C18:1, the melting point is lower if it is located after an odd-number carbon than an even-number carbon, and also if the double bond is located near the middle of the chain as compared with a location at either end [45]. Compatibility of mixed fats from different sources can be a factor. For example, the melting points of multifatty acid TAG, consisting primarily of 16–18 carbon fatty acids, generally rise smoothly with increased content of higher melting fats. However, when TAG with 12–16 carbon fatty acid fats (from coconut and palm kernel oils) are added, concentration-related eutectic points and incompatibility (miscibility gaps) may occur [46–48]. Chemical interesterification of natural fat usually raises its melting point.

Polymorphism and Crystal Types

Polymorphism means "many bodies." Having determined the melting point of tristearin in 1849, Heintz continued to heat the capillary and witnessed: resolidification, a second melting point at a higher temperature, and resolidification and a third melting point for the same sample. Duffy confirmed the principle and reported three melting points, at approximately 52, 64, and 70°C in 1853. Later, five and even seven crystal forms were reported for some fatty acids and mono- and mixed-acid TAG, depending on the heating and cooling history of the sample [49].

As fatty acids or TAG cool, Gibbs free energy ($G = H - TS$) decreases by reduction of both enthalpy (H) and entropy (S, the degree of disorder). At decreased S, fatty acid chains assume pole-like structures that are less co-repulsive and pack more tightly into crystal lattices. It is widely accepted that, to participate in a crystal structure, TAG assume an "h" configuration (also called two-legged chair or tuning fork). This can be envisioned in our gravity-oriented world by assuming that each ester linkage acts as a hinge. If the glycerol number 2 carbon chain is held upright at the methyl end, the number 3 carbon chain hangs directly beneath it forming the back and one leg of the chair, and the number 1 carbon chain juts out at a right angle and curves downward to form the second leg [48, 50–54]. In order to save space, half of the chairs are packed upside down to form a palisade-like structure. A stack of two such structures forms a bilayer whose outer surfaces, consisting of methyl groups, display low inter-attraction. This helps explain why TAG crystals are relatively flat, grow rapidly in length and less rapidly in width by adding parallel chair structures, and grow slowly in thickness by adding additional bilayers.

Within each layer, the chair backs and legs can be further envisioned to act as a vertical loose palisade of knobby-surfaced posts. (Although the hydrocarbon chains have stiffened into pole shape, the carbon atoms are not positioned as beads on a tight string, but rather in sawtooth-like fashion with carbon–carbon bonds of 112°.) The poles can slide behind each other to obtain tighter packing with lowered free energy and can be tilted in two directions to allow the sawtooth-configured carbon atoms on adjacent chains to slip by each other and pack even more tightly. Furthermore, all the vertical units in one layer can be tilted at an opposing angle to the units in the other layer.

As shown in Fig. 34.1, if the three fatty acids in the TAG are saturated and approximately of the same length, each of the layers in the fat bilayer will be approximately two fatty acid chains in height. If the fatty acid on the number 2 glycerol carbon is appreciably shorter than those on the 1 and 3 carbons, each of the layers will be three chains thick. If the TAG is symmetrical, with the 2 position glycerol carbon or both the 1 and 3 carbons unsaturated, the layer will be three chains long but with a zigzag configuration to accommodate the *cis* configuration.

On rapid chilling, a glass (vitreous, γ) form occurs, which can change into a α or β' form as activation energy becomes available. The form showing the least amount of crystalline

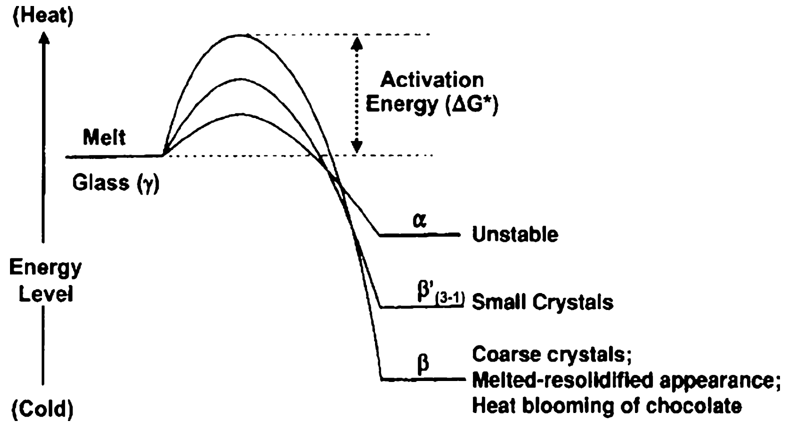


Fig. 34.2 Gibbs energy relationships of polymorphs of a triglyceride. (Modified from Sato [53] pp. 227–263)

order for a TAG, as determined by X-ray diffraction and infrared spectroscopy, is called the *alpha* (α) form. The most compact crystalline form, with the lowest free energy and the highest melting point, is called the *beta* (β) form. One or more intermediate *beta* prime (β') forms also may exist and are indicated as β'_3, β'_2, and β'_1 as the crystal progresses to tighter packing, a lower free energy state, and an increased melting point. When a TAG is cooled very slowly and without mixing, it preferentially assumes the β crystal form. But the tight packing of crystals requires time for alignment and may be thwarted by increased viscosity as the fat/oil mixture cools.

The free energy relationships between the different crystal forms are depicted in Fig. 34.2. Although the β crystal form has the lowest free energy G, induction of its formation requires the highest activation energy ΔG. Differential scanning calorimetry (DSC) often is used to follow free energy changes as fats melt and change between their polymorphic forms. Left to itself, a fat will seek its lowest thermodynamic free energy crystal state.

The formation of β crystals may intentionally be encouraged or hindered, depending on the processing application. The α crystals are relatively unstable, and commercial interest is primarily placed on the differences between the β' and β forms. Generally, β' crystals are smaller (about 0.5–2.0 μm diameter in shortening, and 5–10 μm in margarine [55]), whereas β crystals can grow as large as 20–30 μm. When the objective is to thermally fractionate fats by crystallization, production of the β form is encouraged by carefully controlling temperatures (to not shock the fat into a semi-stable β' form), gentle stirring, and nuclei seeding. The smaller β' crystals have smoother mouthfeel, minimize oiling from margarine, and entrap more air in creaming cake batters. Their production is intentionally encouraged by formulating mixtures of natural or preprocessed fats, by the inclusion of emulsifiers to interfere with crystal growth, and by rapid agitation during plasticizing of the margarine or shortening.

Figure 34.2 also helps explain the mechanisms of defects appearing in fatty products. If a fat that has been conditioned into a stable β' fine crystal form is suddenly exposed to thermal shock and then left unattended, the energy may activate it to settle into the lower energy, coarse β crystal form. Thus, chocolate bars, left to melt and cool several times in an automobile during the summer, turn coarse in texture and mousy in color.

An awareness of crystal packing characteristics and polymorphism helps one to understand incompatibility problems of different fats. Crystal formation has specific demands, and individual crystals in mixed systems each consists of only one species of TAG. However, surfactants and other molecules can act as impurities and interrupt crystal growth. Different TAG are considered compatible when they co-crystallize as separate crystals under the same conditions without the formation of a eutectic.

Other Lipids

Many minor lipids are extracted by nonpolar solvents along with FFA and TAG. Although present in small quantities, they must be removed in refining and sometimes are isolated and purified for medicinal and other uses.

Waxes

Waxes are fatty acid esters of alcohols and are formed by the general reaction:

$$\underset{\text{Alcohol}}{CH_3(CH_2)_nCH_2OH} + \underset{\text{Fatty acid}}{CH_3(CH_2)_nCOOH} \rightarrow$$

$$\underset{\text{Wax ester}}{CH_3(CH_2)_nCH_2COOCH_2(CH_2)CH_3} + \underset{\text{Water}}{H_2O}$$

Table 34.5 Sources and compositions of natural waxes

Type	Melting point (°C)	Main components
Animal waxes		
Beeswax	64	Myricyl palmitate
Chinese	82–84	Isoheptacosyl isoheptacosanoate, ceryl lignocerate
Shellac	81–82	Ceryl lignocerate, ceryl cerotate
Spermaceti	–	Cetyl palmitate
Wool (anhydrous lanolin)	36–42	Cholesteryl estolidic esters, alcohol esters of iso- and anteiso acids
Mineral waxes		
Montan	86	Tricontanyl esters of C_{28-30} acids
Petroleum waxes		
Microcrystalline	71–88	Hydrocarbons (490–800 molecular weights)
Paraffin	54–57	Hydrocarbons (350–420 molecular weights)
Vegetable waxes		
Bayberry	43–48	Trimyristin, tristearin
Candelilla	70–80	C_{29-33} hydrocarbons, simple esters and lactones
Carnauba	80–85	Esters of C_{26-30} alcohols and C_{26-30} ω-hydroxy acids
Esparto	69–81	Hydrocarbons, esters of C_{26-32} acids and alcohols
Japan	51–62	Tripalmitin
Jojoba (a liquid wax)	11–12	Docosenyl eicosanoate
Ouricury	79–85	Myricyl cerotate and hydroxycerotate
Sugarcane	79–81	Myricyl palmitate stigmasteryl palmitate

A major role of waxes in nature is the protection of plant tissues. Examples include coating upper surfaces of leaves to reduce dehydration by the sun and protecting seeds against moisture loss during storage. Minor quantities of waxes are always present in oils extracted by solvents. Waxes extracted from seed hulls cause cloudiness in refrigerated sunflower seed oil and failure of the 5.5-h cold test at 0°C. Sunflower seed oil may be dewaxed by first degumming or miscella refining to remove the natural emulsifier lecithin, which limits the growth of wax crystals, and then winterizing (chilling and filtering) the oil or its *miscella* (oil in its extraction solvent) [56, 57]. With the improvement of dehulling equipment, some sunflower seed processors remove the hulls before extraction. Extracted rice bran, a relatively recent development in the United States, contains waxes with potential commercial promise. Methyl and ethyl fatty acids esters, used as liquid fuels, solvents, and in other applications, also are "waxes" of short-chain alcohols; these are described later.

Waxes are common forms of high-energy storage, in the oils of fish and other marine animals. The major lipids of commercial whale oil consist of approximately 65% waxes and 35% TAG. The lipids of Australian orange roughy (*Hoplostethus atlanticus*) and dory fish oils are 97.1 and 90.9% wax esters, respectively [58]. Essentially all the oil in jojoba (*Simmondsia chiensis*) seed is in wax form. Whale and jojoba oils have been valued for stability in cosmetics and heavy-duty lubrication applications.

Extraction processes for waxes vary in sophistication, from: boiling crushed leaves and berries and skimming the oil (as in the regulated production of Candelilla wax a substitute for hard carnauba wax from *Euphorbia antisyphilitica* in the Big Bend area of Texas and Mexico); to screw pressing seeds such as jojoba; to direct solvent extraction. Cold-pressed jojoba oil is preferred for cosmetics because of concern about other lipids that might be extracted in solvent processes. Oil remaining in the meal has been recovered by secondary hexane extraction for industrial uses. Waxes are susceptible to hydrolysis by nonspecific lipases and are at least partially digestible. They may be hydrogenated or sulfurized into solid forms. The compositions of significant commercial waxes from natural sources are given in Table 34.5.

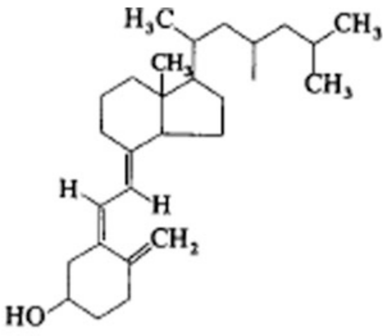
Terpenes

Terpenes are condensation products of the five-carbon *isoprene* (2-methyl-1,3-butadiene) and are extractable by nonpolar solvents. They are classified according to the number of isoprene units: two units, *monoterpenes*; three units, *sesquiterpenes*; four units, diterpenes; six units, triterpenes; eight units, tetraterpenes; and *polyterpenes*. Terpenes may be linear or cyclic. Taken together, this class of compounds includes major essential oils (used in perfume and flavorings), fat-soluble colors, fat-soluble vitamins, and steroids. A saw-toothlike shorthand form often is used to depict the longer chains.

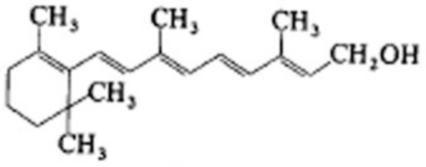
Examples of monoterpenes include the linear aldehyde citral, which is found in many essential oils, and the (*cis*) alcohol *geraniol,* a major component of oil of geranium. Cyclic monoterpenes include limonene, menthol, pinene, camphor, and carvone, major components of lemon oil, mint oil, turpentine, camphor oil, and caraway oil, respectively. Sesquiterpenes include farnesol, a component of rose oil, and bisabolene, a component of Bisabol myrrh. The diterpenes include phytol, a component of chlorophyll, and vitamin A, which is one half of the tetraterpene *β*-carotene. The triterpenes include squalene, a precursor of cholesterol. Examples of tetraterpenes are the oil-soluble: carotenoid plant pigments; xanthophylls, including the yellow pigments lutein in plant leaves and *zeaxanthin in* corn (*Zea mays*); capsanthin, the red pigment in red peppers (*Capsicum annuum*); lycopene, the red coloring of tomatoes (*Lycopersicum esculentum*); and *β*-carotene, a yellow-orange pigment that is the precursor of vitamin A. Bixin, from the seedpods of *Bixa orellana*, is the yellow-orange pigment in annatto food color. It is considered to be a form of carotene oxidized to remove both six-membered end rings. Polyterpenes include gutta, natural water repellant and electrical insulating material from *Palaquium gutta*, and natural rubbers with molecular weights of up to 1.2 million daltons obtained from the latex of the *Hevea brasiliensis* tree and from rubber-filled cells in the branches and roots of the guayule shrub (*Parthenium argentatum*).

Sterols

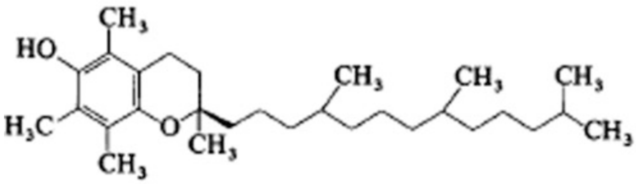
Structures of selected sterols are summarized in Fig. 34.3. Cholesterol is synthesized from lanosterol, which consists of six isoprene units. Cholesterol is a powerful emulsifier and intermediate for synthesizing other steroids and compounds in animals. It participates in multiple ways in fatty acids and



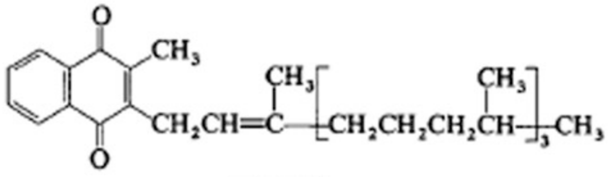
Vitamin D_3



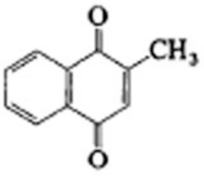
Vitamin A



Vitamin E



Vitamin K_1



Vitamin K_3

TGA transport and is monitored as a major index of arterial and general health. Other important (animal) zöosterols include ergosterol, a precursor of vitamin D required to avoid rickets in children. The ability of certain plant-origin phytosterols to reduce blood cholesterol levels, including low-density lipoproteins (LDL), was discovered in the 1950s and reconfirmed many times. Pine tree "tall" oil, a byproduct of papermaking, is a rich source for commercial preparations. In recent years, selected phytosterols have been approved as food additives and have been included in specialty margarines and spreads. Stanol homologues differ from sterols by saturation at the 5 position, to yield cholestanol instead of cholesterol, and sitostanol instead of *β*-sitosterol. Commercial phytosterol preparations are mixtures of *β*-sitosterol, sitostanol, campesterol, and campestanols. Free sterols have limited solubility in fat, and stanols are hardly soluble; thus, they are converted to fatty acid esters for use in formulated foods. Bennecol®, a table spread, developed by the Raisio Group in Finland and approved as GRAS by the FDA for the United States market in 1999, is an example of a plant sterol/stanol application. The FDA requires that a product provides at least 1.3 g/day of plant sterol esters or 3.4 g/day of plant stanol esters in use to qualify for a cholesterol-lowering health claim [59–61].

Frying studies have shown that phytosterols with double bonds at the 24–28 position in the ethylidene side chain (Δ^5-avenasterol, Δ^7-avenasterol, fucosterol, citrostadienol, and vernosterol) act as chemical antipolymerizers on the surface of frying oils and prolonged frying life. However, stigmasterol, *β*-sitosterol, spinasterol, lanosterol, ergosterol, and cholesterol are either ineffective or slightly increase polymerization [62–65].

Fat-Soluble Vitamins

The fat-soluble vitamins include: vitamin A (retinol), a colorless compound resulting from cleavage of *β*-carotene, vitamin D_3 (cholecalciferol), a steroid, and vitamins E and K (tocopherols and menadione and their derivatives, respectively), which consist of isoprene and phenolic-quinone components.

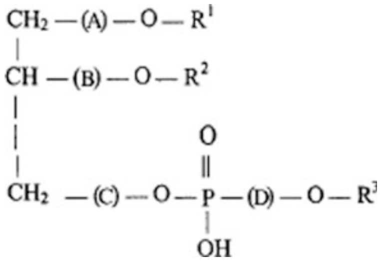
The carotenoids generally are highly susceptible to oxidation. Only selected members (xanthophylls, *β*-carotene, and bixin) are extracted for food coloring. They are used in products such as butter, margarine, and cheese, where exposure to oxygen and light is limited. The more stable yellow-orange pigments (lutein and xanthophylls) are extracted from alfalfa and marigold blossoms for use as poultry feed additives to increase the yellow color in egg yolks and in skins of broilers. Much of the world's *β*-carotene is produced synthetically and is increasingly obtained from (red) palm oil. There is current controversy whether lycopene, the red pigment of tomatoes, has antioxidant properties in the human body.

"Vitamin E" is a generic term for tocopherols and tocotrienols recovered from condensed deodorizer or physical refiner distillates. In earlier years, wheat germ oil was a major source. Structures of tocopherol and tocotrienol isomers are identical except for the presence of double bonds at positions 3′, 7′, and 11′ in the side chain of tocotrienols. The R_1, R_2, and R_3 positions are occupied by CH_3, CH_3, and CH_3, respectively, in the *alpha* (α) isomers; CH_3, H, and CH_3 in the *beta* (*β*) isomers; H, CH_3, CH_3 in the *gamma* (*γ*) isomers; and H, H, and CH_3 in the *delta* (δ) isomers. As shown in Table 34.6, tocopherols are the primary tocols in soybean, corn, olive, and sunflower seed oils, and tocotrienols are primary in wheat, rice bran, and palm oils. During chill fractionation, tocotrienols in palm oil concentrate in the olein (liquid) phase. The greatest vitamin E activity is shown by α-tocopherol—the commercial "vitamin E"—with lesser activities shown by other tocopherol and tocotrienol isomers. However, when used as antioxidants, the *gamma* and *delta* isomers often have proven more effective in protecting vegetable oils against oxidation. Intensive research continues on potentials of tocols for avoiding or curing various diseases, with considerable focus on diseases of aging.

Oils from the livers of cod and other fish species were early sources of vitamin D_3; also, ergosterol was irradiated by ultraviolet light to produce vitamin D_2. Currently, most of the commercial vitamin D is produced synthetically. Commercial vitamin E is recovered from deodorizer condensates from plants refining palm and other oils, including soybean.

Phosphatides

Phosphatides are natural emulsifiers with the following structures [66].



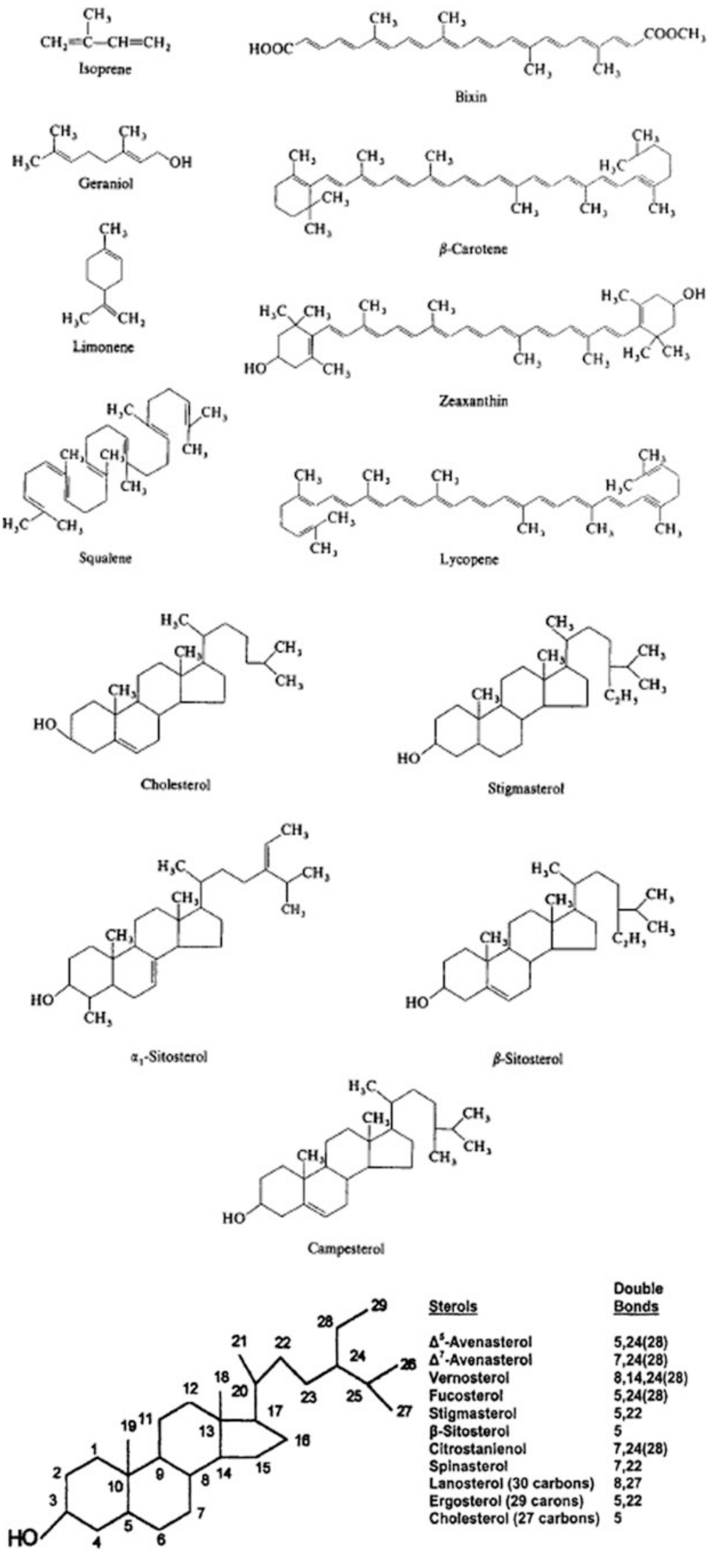
where

R^1 = fatty acid.

R^2 = fatty acid, or the phosphatidyl group in the 2 position.

When R^3 is one of the following, the phosphatide is named:

Fig. 34.3 Structures of selected sterols. Sources: animal—lanosterol, cholesterol, and ergosterol (also microbial); plant—all others. (From Warner [65], pp. 37–49, With permission)



Sterols	Double Bonds
Δ^5-Avenasterol	5,24(28)
Δ^7-Avenasterol	7,24(28)
Vernosterol	8,14,24(28)
Fucosterol	5,24(28)
Stigmasterol	5,22
β-Sitosterol	5
Citrostanienol	7,24(28)
Spinasterol	7,22
Lanosterol (30 carbons)	8,27
Ergosterol (29 carons)	5,22
Cholesterol (27 carbons)	5

Table 34.6 Tocopherol and tocotrienol contents (μ/g/g) of common refined edible oils (From: Sundram et al. [206], With permission)

Tocol isomers	Soybean oil	Corn oil	Olive oil	Sunflower oil	Milk fat (ghee)	Wheat germ oil	Rice bran oil	Palm oil	Palm olein	Palm stearin
α-Tocopherol	117.2	248.9	151.4	485.2	32.7	218.9	64.0	188.2	179.0	50.0
β-Tocopherol	19.8	10.1	13.3	3.0	n.d.	33.2	10.6	n.d.	n.d.	n.d.
γ-Tocopherol	560.7	464.1	10.9	51.0	n.d.	84.7	n.d.	n.d.	17.6	n.d.
δ-Tocopherol	178.2	58.2	n.d.	n.d.	33.8	n.d.	187.0	n.d.	n.d.	n.d.
α-Tocotrienol	n.d.b	n.d.	n.d.	n.d.	n.d.	n.d.	31.4	198.1	219.9	47.4
β-Tocotrienol	20.2	n.d.	n.d.	n.d.	n.d.	347.5	83.2	10.0	8.1	9.0
γ-Tocotrienol	6.2	n.d.	n.d.	n.d.	n.d.	n.d.	783.2	198.8	332.7	134.9
δ-Tocotrienol	n.d.	n.d.	n.d.	n.d.	n.d.	18.4	38.6	98.4	67.0	31.4
Total	902.2	781.4	175.6	547.5	66.5	702.7	1198.0	693.5	824.3	272.8

nd not detected

$R^3 = CH_2–CH_2–N^+(CH_3)_3$	diacylphosphatidyl choline
$R^3 = CH_2–CH_2–NH_3^+$	diacylphosphatidyl ethanolamine
$R^3 = CH_2–CH(NH_3^+)–CO_2H$	diacylphosphatidyl serine
$R^3 = C_6H_6–(OH)_6$	diacylphosphatidyl inositol
$R^3 = H$	diacylphosphatidic acid

and preceded by names of the two fatty acids if relevant.

In the pharmaceutical industry, the word *lecithin* is synonymous with phosphatidyl choline. However, the entire or modified mixture is sold as lecithin in food supplement capsules, and for food, feed, and industrial uses. Soybean oil phosphatides consist of 29–39% diacylphosphatidyl choline, 20–26% diacylphosphatidyl ethanolamine, 13–17% diacylphosphatidyl inositol, 5.9–6.3% diacylphosphatidyl serine, and 5–9% diacylphosphatidic acid [66].

The phosphatides have polar and nonpolar sites and generally act as water-in-oil emulsifiers. They are extracted by solvents with the oil, but preferentially will absorb available water, form gums, and precipitate. Commercial lecithin is produced by *water degumming* (precipitation from oil with ion exchange-treated water) as explained later.

Phosphatides precipitate on hydration during the storage of oils, foul bleaching earths, poison hydrogenation catalysts, and cause darkening of the oil during deodorization/physical refining and also if it is used in frying applications. Their removal is desirable, but requires close supervision to preserve oil yields. Four lipases are able to hydrolyze phosphatides. Phospholipase A cleaves fatty acids at the "A" position; phospholipase B (also called "A2") cleaves fatty acids at the "B" position; phospholipase C cleaves the phosphatidyl chain next to carbon "C;" and phospholipase D cleaves between the phosphate structure and choline, ethanolamine, serine, or inositol. Cleaving by phospholipase D and dissociation of the phosphorous group expose two negatively charged sites, which can complex with divalent cations (mainly calcium and magnesium, but also including iron, copper, and others) present in soybean tissue during solvent extraction. As a result, the phosphatide becomes nonhydratable.

The production of nonhydratable phosphatides (NHP) by phospholipase D can be minimized by heat inactivation of the enzyme by expanders/extruders while preparing seed for solvent extraction, by heating briefly to more than 112°C/235°F. However, some NHP are unavoidable, especially in wet fall seasons when high moisture seed may begin to sprout. Unless the NHP content is extremely high, they can be changed to the hydrated form by treating the crude oil with acidic chelating agents (*acid degumming*) to withdraw the divalent cations as described later. In the last two decades, several companies have introduced *enzyme degumming*. In these processes, usually a position 1 lipase is introduced to hydrolyze the R^1 fatty acid. This leaves the 1 position attractive to water and renders the phosphatide hydratable again.

The enzymatic degumming process is very robust and versatile. Using this process increases oil yields and reduces phosphorus content to a greater extent. The method is suitable for edible oil and biodiesel processing. Recently in 2010, Varenium has marketed their two Purifine PLC products for the degumming process.

Fats and Oils Sources and Consumption

Production

Global production and export estimates for fats and oils for 2003/2004 are shown in Table 34.7. Approximately 39% of the world's production of vegetable oils, and an additional 20% of oilseeds grown, enter international trade [20, 67]. Records of estimated production, imports, and exports are kept, country by country, by the Foreign Agriculture Service (FAS) of the US Department of Agriculture, and *Oilseeds World*, a publication in Hamburg, Germany.

The leading producers of the major oilseeds, in decreasing order, are *Soybean*: the United States, Brazil, Argentina, China, and Paraguay; *Rapeseed/Canola:* China, the

Table 34.7 Estimated production and exports of world's major edible-type vegetable oils (From: Foreign Agriculture Service Circular Series. Department of Agriculture, April 2010/2011)

	Seed/Copra		Meal		Oil	
Oil/Fat source	Production (MMT)	Exports (MMT)	Production (MMT)	Exports (MMT)	Production (MMT)	Exports (MMT)
Soybean	260.97	98.51	177.82	59.92	42.0	9.99
Palm	–	–	–	–	47.47	37.53
Canola/rapeseed	58.56	10.53	34.63	4.41	23.02	3.15
Sunflower seed	30.93	1.48	12.14	4.19	11.28	4.36
Peanut	34.71	2.49	5.93	0.11	4.89	0.17
Cottonseed	43.19	0.71	14.84	0.36	4.96	0.17
Copra/coconut	5.89	0.21	1.95	0.81	3.68	1.88
Palm kernel	12.73	0.02	6.70	5.03	5.65	3.27
Olive	–	–	–	–	3.01	0.65
Fish	–	–	4.97	2.87	–	–
Total	446.89	113.86	258.98	77.70	145.96	61.00

MMT million metric tons

European Union, Canada, India, and Eastern Europe; *Sunflower seed:* the former Soviet Union 12, Argentina, European Union, Eastern Europe, China, and the United States; *Peanut/Groundnut:* China, India, and the United States; and *Cottonseed:* China, the United States, India, Pakistan, the former Soviet Union 12, and Brazil.

Soybean oil is the world's largest supply of *visible* (separated) fats, accounting for approximately 29.8%. Palm, rapeseed/ canola, and sunflower seed follow it, and peanut oils in tonnages produced. The production of palm oil has been increasing and now accounts for 28.8% of the world's supply; added to the 3.4% palm kernel oil produced, the palm crop provides 32.2% of total oil supply. With oil palm plantings still to mature in various tropical countries, palm oil production alone is expected to bypass soybean oil within the next several years. However, palm oil generally is the lowest cost edible fat available and accounts for 54.7% of the world's exported oils. Thus, it is one of the more economical and easier acquired oils.

Currently, the United States is the world's leader in soybean production (35.2%), followed by Brazil (32.7%), Argentina (17.9%), and China (8%). Approximately 34% of the world's soybean meal (for animal feed), 29% of the oil (mainly in degummed form), and 34% of the world's soybean seed (mainly for overseas extraction) enter global trade. The United States exports 36% of its soybean crop in seed form, 12% of its processed meal, and 3% of its soybean oil. In contrast, Brazil exports 35, 65, and 48%, and Argentina exports 20, 98, and 98%, respectively. The latter two nations each surpass the United States in tonnage of meal and oil exports. This results partially from increasingly more of the domestic crop retained in the United States to supply the local population's needs. Also, larger tariffs were imposed on seed exports than on meal or oil in Brazil and Argentina in the 1970s and 1980s to encourage development of domestic oils extraction industries; these strategies appear to have been successful. Because of population growth, China now imports more soybean than is produced internally [20].

Rendered beef, pork, poultry, and other animal fats are not well reported internationally, and global statistics are unreliable. Total production of fats in the United States by the rendering industry for 2000 is estimated at 4.18 million metric tons [68]. Outputs of all rendering facilities captive to integrated broiler operations might not be included. Of the amount reported, 76% is inedible tallows and greases, 18% is edible beef or mutton tallows, and 6% is edible pork lard.

Fish oils are not well reported either, with annual estimates at about 1.3 million metric tons from sustainable (sea catch) fisheries. Production has ranged by as much as 50% between years, depending on availability of fish.

Changes in Sources

The maturing of nations as raw materials suppliers follows a sequence. Centuries ago, bands of marauders attacked villages to steal crops after harvest. Later, countries conquered neighboring lands for "a place in the sun" for their growing populations. With improvements in transportation and discovery of the New World, followed by Africa and Australia, it no longer was necessary to relocate large populations. Colonies could be established to supply the "mother country" with raw agricultural materials in addition to minerals and fossil fuels. As local education improved, colonies typically declared their independence, but the new countries needed something to trade for goods they were unable to produce themselves.

The oils of ancient times were olive and sesame oil in the Mediterranean basin, rapeseed and animal fats in Europe, and coconut (copra) oils in the tropics. Cottonseed was the world's first new oil of the Industrial Revolution age

Table 34.8 Gross composition of major undehulled oilseeds

Crop/Source	Moisture (%)	Protein (%)	Fat (EE) (%)	Crude fiber (%)	Ash (%)	Oil: co-products ratio
Soybean	8.5	36.5	19.5	5.8	4.9	1 : 4.1
Cottonseed	8.0	23.0	21.0	24.0	4.8	1 : 3.8
Peanut	6.5	25.7	49.2	4.9	2.3	1 : 1.0
Sunflower	6.0	21.1	42.0	17.4	3.3	1 : 1.4
Safflower	5.8	19.4	43.5	20.8	3.7	1 : 1.3
Coconut (copra)	4.0	7.5	67.3	5.0	1.9	1 : 0.5
Rapeseed/canola	8.0	22.0	41.2	11.5	5.1	1 : 1.4
Palm kernel	10.5	9.7	58.1	–	–	1 : 0.7
Sesame	8.0	24.2	47.6	11.2	6.1	1 : 10

EE ether extraction method.

(early 1800s) and became the dominant US oil after the Civil War (from the latter 1860s until the mid-1930s). But, cottonseed oil is a byproduct of growing cotton fiber, and edible oil and animal feed requirements of the world's growing populations soon exceeded supplies of this crop. Solvent extractors, invented in Germany in the early 1920s, maximized oil recovery and produced animal feed protein meals with less heat damage. As European demands grew for a closer and more reliable source of soybean than had been available from the Manchuria area of China since 1910, soybean export opportunities opened for the United States. Many of continental Europe's oil mills were demolished during World War II, placing the United States in the position of major soybean seed, oil, and meal supplier to the world. In 1960, the United States grew and traded about 60% of the world's soybean. China grew about 32% of the world soybean crop and supplied about 19% of the trade, but soon decreased as a world supplier because of its own population growth. The poultry broiler industry became global about 1960. Its large requirements for feed protein, best supplied by soybean meal, contributed to increasing world soybean production by 7.5 times in the succeeding 40 years.

Although slightly more than four tons of meal is produced for each ton of oil, the sheer volume of the meal business has co-produced enough soybean oil to keep it the world's major oil until now. Much of the production and processing technology was developed by the United States and European nations. In time, roads, canals, and port systems, crop production, and processing infrastructures in South America were funded as economic development programs by the World Bank, Regional Banks, and by private investors. The position of "lowest cost producer of soybean and soybean oil" has passed from the United States to South American countries. Brazil has opened its sub-Amazon basin to soybean growing, an area four to five times larger than available in the United States, and has developed varieties acclimated to the climate and day length. Its production of soybean is expected to surpass the United States in the near future. China also has developed salt-tolerant varieties that will grow in its coastal regions. At the same time, economic assistance and private investment in Southeast Asia and tropical countries have led to production of palm oils at prices lower than soybean oil.

The appearance of Asian soybean rust (*Phakopsora pachyrhizi)*, first of South America and in the United States in 2004, is a cause for concern. This fungus has potential for greatly reducing soybean yields per acre and could become a critical factor in world protein and oil supplies. No other crop is as capable of producing protein as soybean, essentially resulting in a monoculture. The Unites States has maintained a wheat rust response program for nearly half a century, continuously developing rust-resistant varieties for replacement as needed. Fortunately, biotechnology has provided even better tools for newer programs.

Many factors dictate which oilseed species will be grown and/or imported into a country. Climate and local demand for high-protein feed meals are leading factors. Cool weather and short growing seasons have essentially limited Canada, Northern Europe, and the former European Russian republics to growing rapeseed/canola or sunflower seed. Some European countries grow and export seed or oil of canola, but import soybean for their edible oil and animal feeding industries. Insect problems were important factors in selecting crops before modern insecticides and IPM systems became available. Devastation of the southeastern United States cotton crop by the boll weevil led to the introduction of large-scale peanut growing in the early 1900s.

As shown in Table 34.8, the oil content of row crop oilseeds varies from about 19% for soybean to 43% for sunflower seed, and 41–45% for rapeseed/canola. More feed co-products always are produced than oils, with a ratio 4:1 in the case of soybean. Soybean meal is the major feed protein source for production of poultry, currently the leading domestic and global meat source, and in the rapidly developing aquaculture industries.

The relative availability of fat- and oil-bearing byproducts of other current agri-businesses is an additional consideration in types and amounts of oilseeds grown. Because only about 12% of the return to cotton farmers comes from the seed, the domestic supply of cottonseed for crushing is dictated by world demand and the price of cotton and is hardly affected by price of cottonseed oil. Corn oil has become the second major oil in the United States due to large quantities of corn germ provided by rapid growth of the domestic corn sweetener and ethanol industries. There is sufficient processing of rice in the United States now to warrant two rice bran oil extraction plants, with quantities of stabilized rice bran also shipped to Japan. Beef, pork, and poultry packing operations always produce fatty tissues for rendering into inedible tallow, lard, chicken fat, and meals used in animal feeds. Generally, fats/oils from co-products of other local processing industries must clear the market first, at whatever price they can get, before growing or importing of significant quantities of high-oil content crops becomes economically attractive.

Consumption

The world's current production of edible oils is estimated at about 101 million metric tons, and an increase of ~2.5 million metric tons is needed annually to meet the needs of the growing population. Average per capita consumption of fats/oils is difficult to determine from gross disappearance figures because these materials also are used in animal feeds and industrial applications. Generally, consumption is related to personal income and local availability, but once fats are introduced into the diet, their priority among food expenditures remains high. The annual consumption of oils is estimated at 10.4 kg per person for the world and ranges from 27.2, 42.9, and 29.3 kg for the United States, Belgium, and West Germany, respectively, to 13.1, 4.7, and 5.9 kg, respectively, for Egypt, China, and India.

About 60% of the total fat consumed domestically is "invisible," in meat, poultry, fish, dairy products, eggs, and prepared foods. The visible 40% is used primarily in the form of salad and in-home cooking oils, shortenings, and margarine. From 1965 to 1990, average domestic consumption of fat decreased from 50.7 to 32.5 kg per capita for men 19–50 years of age and 30.3 to 23.4 kg/year for women.

Since 1990, consumption increased to 36.9 kg/year for men and 23.7 for women in 1995 [69]. Approximately two thirds of visible fats available per capita in 1940 were from animal sources. The use of vegetable source fats has grown significantly, accounting for about two thirds of the visible fats consumed in 1965, and 90% in 1985 [70]. More recent data have shown that fat consumption per capita has not deceased in grams per day intake, but has decreased as a percentage of total diet with the average caloric intake increasing.

At times, the public acts indifferently regarding caloric intake. In the mid-1990s, various low-fat or nonfat snack foods were introduced, requiring different processing machinery, techniques, and flavoring technologies. Although promising at first, sales of these products showed a decline in 1997 [71]. The Procter & Gamble Company developed olestra (Olean™) during a 28-year period (1968–1996) before the FDA approved its use as a frying oil in *savory snacks* (salty chips, crackers, and tortilla chips) and spent several hundred million dollars in the process. Olean™ is a polyester of sucrose and 6 to 8 fatty acids, which is nondigestible by human lipases, and thus noncaloric. However, public interest shifted from snacks promoted as "healthy" to indulgence in salty, traditional-flavor, higher fat content products [72]. Many of the newer snack products were withdrawn from the marketplace. Digestive tract upsets also found in earlier evaluations of other nondigestable fats were again reported among Olean consumers.

Extraction of Fats and Oils

Basic Processes

By trial and error over the centuries, man has learned five basic skills in handling oilseeds:

1. *Preservation of seed by natural or artificial drying* and cooling to a dormant state before storage, with protection from insects and rodents. Generally among oilseeds, sprouting enzymes become active once seeds rise above 75% relative humidity (RH). Rise in free fatty acid (FFA) content signals need for early processing.
2. *Removal of trash and hulls*, by stamping, threshing, disk hullers, cracking rolls, and other devices, followed by winnowing, sieving, or aspiration to separate kernels ("meats") from hulls and chaff.
3. *Freeing the oil*, by pounding seed with rocks, mortar, and pestle-type grinders (which later became human, beast, water, or electric-driven ghanis); use of vertical stone/iron wheels known as "edge rollers" to crush the seed; cracking and flaking rolls; and, more recently, expanders/extruders.
4. *Heating the seed to increase oil recovery*. This was first interpreted as denaturation, making the protein matrix brittle to surrender the oil on pressing. Later, it was recognized that concurrent inactivation of enzymes also arrests development of various types of degradations.
5. *Separation of oil* from crushed seed by: draining; squeezing cooked mash in cloths by lever, wedge, or hydraulic presses; continuous screw presses, or by solvent extractors.

Initially, various societies used plant oils for medicinal and cosmetic purposes. Later, they were used for lighting and, as extraction and refining techniques improved, for food. Crude techniques are still used in remote areas.

Screw Press Operations

Continuous screw presses are used: (1) for extracting fats and oils in small operations where investment capital or supplies of raw materials are limited and installation of a solvent extraction plant is impractical; (2) to partially defat high-oil content seeds for easier handling in subsequent solvent extraction or hard pressing; and (3) for extraction of animal flesh and bones, fish, and fleshy-type oilseeds such as palm fruit, olives, and copra (dried coconut "meat") and oilseeds. These machines have been generically referred to as "expellers," but the Expeller® trademark belongs to Anderson International Corporation, Cleveland, OH, successor to the company founded by Valerius D. Anderson who patented the first continuous screw press in 1899.

The main principle in *hard pressing* is to preheat (cook and dry) prepared seed and animal materials to the point where cell walls become brittle and rupture readily on pressing. Generally, seed is dehulled, tempered, flaked, and cooked (dehydrated) to low moisture (3–4%) before hard pressing. Screw presses used for hard pressing row crop oilseeds (soybean, cottonseed, rapeseed/canola, sunflower seed, and peanut) typically include two stages: the second stage, operating at higher pressure, further extracts oil remaining after first stage pressing (Fig. 34.4). Approximately 4–5% oil is left in oilseed meals. Single-stage continuous screw presses typically are used in rendering operations and in pressing palm fruit and olives, which leave bone or hard seed pieces in the press cake. Approximately 9–12% fat is left in meat and bone, and fish, meals, and 5–6% oil in oilseed meals. High temperatures partially destroy the dietary essential amino acid lysine and reduce nutritional quality and economic value of protein feed meals. Milder *wet processing* techniques have been developed for meat and fish meals.

Generally, oils from press operations go to a settling tank. A layer of foam may be skimmed off; the *midfraction* (oil) is filtered and sent to refining, and the settled solids (*foots*) are spread over the stock going to the screwpress. Pressed oils benefit from cooling to less than 42°C/11°F as soon as possible to slow oxidation and setting of color.

The traditional practice in processing high-oil row crop oilseeds (containing over 25% oil, such as canola/rapeseed, sunflower, peanut, safflower, and corn germ) was to reduce the oil content to less than 18% using lighter duty *prepresses*, break up the press cake, and finish with solvent extraction. Such operations are called *prepress-solvent extraction*. Cottonseed also was prepressed in earlier years. Expanders with oil removal cages (Fig. 34.11, to be described in more detail later) are used to reduce oil content to less than 20% and are replacing prepresses in high-oil content seed extraction plants. Modified discharge heads are used on expanders to enhance hard pressing.

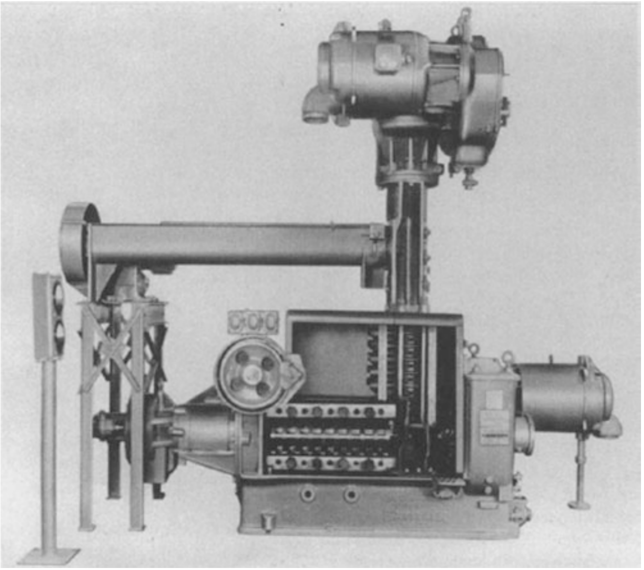


Fig. 34.4 Anderson Expeller® Press, model 55 in. TDMS.™ Note elevated conditioner carrying product from *left* to *right* across *top* of press; vertical screw first-stage press section (*right* side); higher pressure second-stage press screw section carrying product from *right* to *left*. Some hard presses have both stages on one shaft. (Courtesy of Anderson International Corp., Cleveland, OH)

Most oilseeds contain excessive fat for extensive animal feeding. "Scalping" oilseeds such as cottonseed and soybean by preheating and partially pressing to about 9–12% residual oil (which is sold) is becoming popular among large animal feeders, feed manufacturers, and farmer cooperatives. Toxic components, such as trypsin inhibitor (in soybean) and gossypol (in cottonseed), are partially inactivated, and the meals are improved for feeding selected animals.

Decanters and Centrifuges

Two major types of centrifugal separators are used in press-type oil/fat extractions and *wet rendering:* horizontal decanters and vertical stacked conical disk centrifuges. A drawing of a decanter, which separates solids and liquids, is shown in Fig. 34.5. Decanters are built with different internal designs depending on the solid–liquid ratio to be separated and operate at ~2,500–3,000 × gravity (*g*). When properly operated, a "dry" solids phase and a liquid phase of two immiscible liquids (aqueous matter and oil) are discharged.

A cut-away of a three-phase stacked-disk centrifuge is shown in Fig. 34.6 [16]. These are vertical separators, operating at ~6,000 × *g*, which employ a spinning bowl and set of conical stacked disks. Centrifuges are used

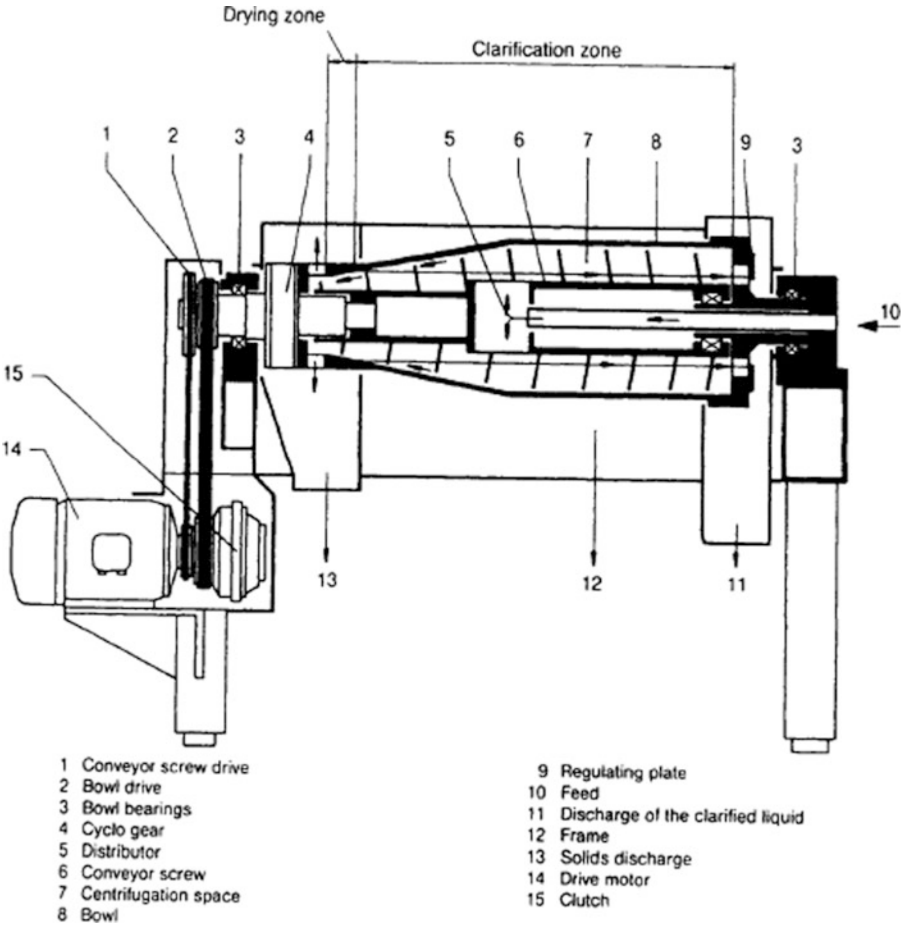


Fig. 34.5 Schematic drawing of Westfalia Model CA 450 Continuous Clarifier Decanter. (Courtesy of GEA Westfalia Separator Company, Northvale, NJ)

primarily in three-phase mode to separate the liquid portion into two immiscible phases and a fine solids (*sludge*) fraction. They are not meant to handle significant amounts of coarse solids. The material to be separated enters the spinning bowl at the bottom through a hollow spindle and is discharged under the lower plates to be thrown against the side of the bowl by centrifugal action. Continual arrival of fresh liquid forces the earlier liquid into the plates, with the heavier aqueous phase remaining at the outside of the bowl, and the oil phase moving into the center. In the past decade, both major manufacturers of centrifuges in the western world have introduced systems for adjusting the aqueous: oil phase ratios exiting their machines while running. The solids, collected at the outer extremities of the bowl, are allowed to escape as heavy slurries by momentarily lowering the bottom of the bowl on a timed *desludging cycle*.

Olive Oil

Traditionally, oil was extracted from fruit of the olive tree, *Olea europea*, by crushing cleaned, ripe olives, including the seed, using stone “edge rollers” or metal grinders; milling by mixing the paste while heating to coalesce the oil droplets; and shrouding it in press cloths or mats for squeezing in lever, wedge, mechanical screw, or hydraulic presses. Next, the *must* (oily liquids escaping the press) was centrifuged to separate the oil and aqueous phases. Lower quality secondary oils were obtained by reworking and repressing the pomace, or by drying it for solvent extraction. Modern olive oil extraction plants: crush the olives and seed using hammer mills; mix the paste while heating; and separate the must from pomace using decanters. The must is heated and separated by disk centrifuges. The solids may be reextracted several times. Some olive oil plants now use bladder presses, similar to those for crushing grapes in making wine, to remove the highest quality oil. Virgin oils are separated by pressing, and usually are usable after filtration. Pomaces are dried and their remaining oil extracted by solvent. Some operations use the desolventized extracted pomaces to fuel steam boilers for the plant [73–75].

Coconut Oil

Coconuts, from the *Cocas nucifera* palm, are dehusked before cracking the nut to drain away the coconut water. Then, the flesh is separated from the adhering shell and dried, raising the oil content from approximately 30% in

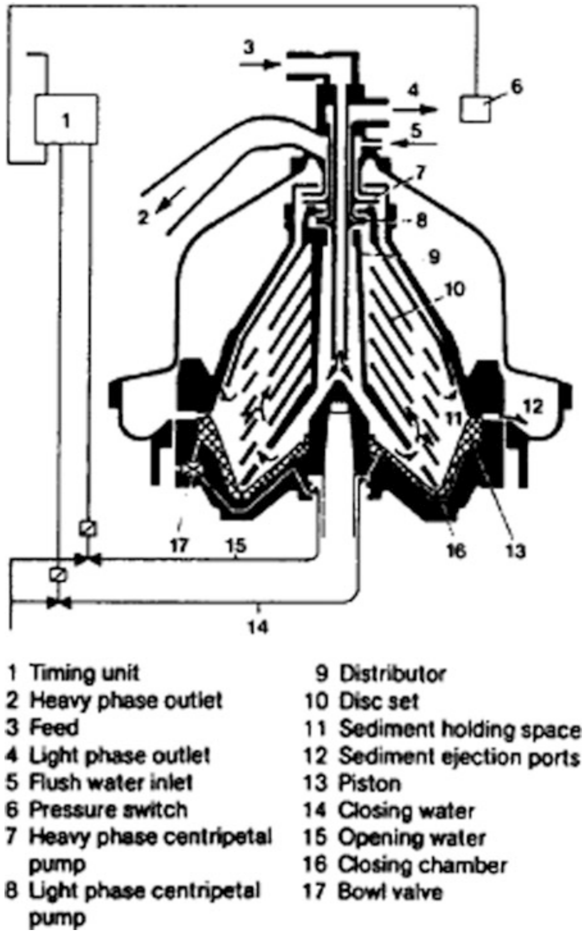


Fig. 34.6 Schematic drawing of Westfalia Model RSA refining separator with self-cleaning bowl. (Courtesy of GEA Westfalia Separator Company, Northvale, NJ)

the "meat" to 50–65% in the resulting copra. Considerable amounts of copra are prepared by hand and sun-dried to supplement family income in tropical countries. Centralized coconut mechanical husking and dehulling facilities, which dry the coconut meat using heat generated from burning the husks, offer advantages of improved moisture control and reduced mold spoilage and aflatoxin problems and are increasing in number. In the oil extraction process, copra is: (1) cleaned by shaker screens to remove trash; (2) size-reduced by hammer or attrition mills, or fluted roller mills; (3) optionally flaked; (4) additionally dried if necessary; and (5) hard pressed or prepress-solvent extracted [76]. Specially designed expanders have been applied to preparing copra for coconut oil extraction.

Palm Oil

Two distinctly different types of oils are produced from the fruit of the Southeast Asia and African oil palm, *Elaeis guineensis*, and its hybrids with the South and Central American palm, *Elaeis oleifera*. Palm oil is obtained from the red-orange fleshy part of the fruit, which resembles an oversized olive about the size of a small chicken egg. Palm kernel oil is derived from the kernel within the nut. Well over 98% of the fatty acids in palm oil belong to the C16 and C18 palmitic-stearic group, whereas approximately 64% of the fatty acids in palm kernel oil consist of the C12 and C14 lauric group.

A palm tree produces 10–15 fresh fruit bunches throughout the year, weighing 5–23 kg (10–50 lb) each. The bunches are cut from the tree with knives attached to long poles and are transported to the oil mill, sometimes by a small-gauge railroad. Although palm fruits somewhat resemble olives, they have a very strong lipase enzyme which is deactivated first by steaming the bunches at about 40 psig for 50–75 min. This also loosens the fruits on the stalk, which are freed in thresher drums and passed through a digester to convert the fleshy pulp to mash. Then, the mash is pressed by twin-screw or hydraulic presses to yield red crude oil. The nuts are dried, shells cracked, and the kernels separated and bagged or bulk-stored for sale and/or solvent extraction in a fashion similar to the processing of row crop oilseeds [77].

Animal Fats and Fish Oils

Extraction of Lards and Tallows

A variety of methods, old and new, is used for extracting animal fats [78]. The method used depends on the required properties of the edible or inedible nonfat product produced and the age of the facility since construction or its last modernization. In recent years, continuous processes generally have replaced batch processes for edible and inedible rendering to take advantage of heat recapture systems, economies and reduction of worker safety problems through automation, and the ability to minimize discharge of offensive odors. Modern edible and inedible rendering plants located near urban areas often are under negative pressure, and the building and processing equipment operate under slight vacuum. Steam vapors from processing are condensed, and air taken into the building is water-scrubbed before return to the atmosphere. Edible and inedible animal fats and fish oils may be subjected to refinery processes similar in principle to those of vegetable oils. Industrial fats often are split by high-pressure steam into fatty acids for further processing by the oleochemicals industry [79].

Edible products for human consumption are processed under federal or state inspection in facilities separate from inedible rendering plants. In the United States, construction of the processing area and designs of equipment selected are previewed and approved before startup by the Food Safety

Inspection Service of the USDA. Only raw materials passed by FSIS inspectors may be used. They may be chilled or frozen and shipped in refrigerated trucks or railroad cars under FSIS seal to the edible rendering facility. On-duty FSIS (or state) inspectors monitor the rendering operation. Edible rendering plants are limited to one species in the United States, and mixing fat of different species is not permitted [80].

Edible beef tallow is produced from the fat of cattle (*Bos taurus*). A counterpart, produced from sheep (*Ovis aries*), is known as mutton tallow. Lard is rendered from the fatty tissue of pigs (*Sus scrofa*). Edible chicken fat is rendered in some countries [81]. Essentially, three methods are used for separation of fat from animal tissues. In older *batch wet rendering* processes, an autoclave is filled with precut raw material, closed, and steam is injected to raise the temperature to about 140°C/280°F. After heating for 2–3 h, the pressure is slowly reduced to atmospheric to avoid emulsification. After a settling period, the free fat is drained from the autoclave; the cracklings are pressed by single-screw machines and sent to drying. Indirect heating is used in *batch dry rendering*. A reactor, equipped with a rotating agitator, and jacket at 6–7 bar (88–103 psi), liberates the fat in 1.5–2 h. On discharge, the fat is drained and the cracklings are pressed. Where batch-type cage presses are still used in other countries, the cake is formed in the shape of large wheels, sometimes called *greaves*.

Much of the edible rendering industry has adopted continuous *wet rendering*, in which the minced raw material is heated first to 60°C/140°F, and then to 90°C/195°F, in an airtight melting section within minutes. The heated material then is separated into solid and liquid fractions by mechanical horizontal decanters (Fig. 34.5) or dewatering screw presses, and the oil is separated from the aqueous portion (stick water) and fines by a three-phase disk-type centrifuge (Fig. 34.6). The pure fat is flash-dried and cooled to approximately 12°C/20°F above melting point by passing through a tube-in-shell or plate heat exchanger before storage and shipping [82].

Lower heating temperatures are used when producing partially defatted beef fatty tissue (PDBFT), or similar edible products of other species, used in making processed meats, including frankfurters. Fatty tissues and trimmings are ground and heated to 43°C/110°F in a mixing tank to melt the fat but still retain the heat-setting properties of the tissues for later use. The slurry is passed through a disintegrator to rupture the fat cells, and then through a horizontal decanter for separation into solid and liquid phases. The solids are cooled for packaging and used in making processed meats. The liquid is heated to 93°C/200°F to coagulate proteins and passed through a desludging disk-type centrifuge to "polish" the oil by separating the stick water and removing any solids [83].

Higher heat treatments, using open cooking kettles, produce darker and more strongly flavored fats and cracklings, also usable as edible food ingredients, which are preferred in specific applications. But, for the most part, bland edible animal fats are preferred and may later be neutralized (alkali refined), bleached, hydrogenated, deodorized, rearranged (interesterified), and fractionated in ways similar to vegetable oils before incorporation into margarines/spreads and shortenings, or used for frying [84].

Inedible Animal Products

Inedible animal products include viscera and inedible parts, carcass parts condemned by inspectors, fatty trimmings from butcher shops, outdated fresh and processed meats, and dead animals that have been skinned. They may be rendered by wet or dry batch processes. In modern rendering, the flesh and bones are ground or chipped, heated to denature the protein and release the aqueous and fat fractions. Next, the slurry is passed through a decanter or dewatering press to separate the solids from the liquid which is further heated to reduce viscosity in separating fat, stick water, and sludge. The meat/bone fraction is dried, reaching high temperatures specified by the federal government to destroy pathogenic bacteria. The solids then are pulverized by hammer mills and sieved for animal feed use, with the separated large pieces recycled to the grinders. Inedible renders, sometimes called *the original recyclers*, play an increasingly significant function as the world wrestles with problems of biodegradable waste disposal.

Restaurant Greases

Used frying oils typically are processed in separate inedible systems. Restaurants and large commercial fryers dump their spent oils into on-site bins, equipped with steam heating coils provided by the grease processor. Pilferage of grease has been a problem in the industry, and covers on the bins are kept locked. The collector may attach a steam line to heating tubes in the bin to melt the fat if required; then, it is pumped to a tank on the collection truck. An exchange of fats occurs between the oil and the product during frying. For example, the used grease may contain substantial quantities of chicken fat from frying. At the processing plant, the liquid grease is heated to 100°C/212°F and flashed into an evaporator flash chamber operating at 82–85°C/180–185°F and 21 in. mercury vacuum to remove the moisture. Next, the dried grease is passed through a decanter to remove any solids and partially cooled before storage [85]. Some inedible renderers also provide services for pumping restaurant and grocery store grease traps. These fats are segregated for handling in separate processing facilities and sold in separate markets.

Table 34.9 Specifications or typical analyses of edible and industrial/feeding animal fats (Compiled from: NRA [80])

Grades	Titre min (°C)	FFA max (%)	FAC[a] color max	Moist. max (%)	Iodine value max	Saturated (%)	Fatty acids distribution unsaturated (%)	Linoleic (%)
Edible[b]								
Beef tallow—USDA certified	41.0	0.75	3	0.10[c]	40–45			
Lard—USDA certified	38.0	0.50	[d]	0.20[c]				
Inedible—industrial[b]								
Top white tallow	41.0	2	5	1 MIU[e]				
All-beef packer tallow	42.0	2	None	1 MIU				
Extra fancy tallow	41.0	3	5	1 MIU				
Fancy tallow	40.5	4	7	1 MIU				
Bleachable fancy tallow	40.5	4	None	1 MIU				
Prime tallow	40.5	6	13–11B	1 MIU				
Special tallow	40.0	10	21	1 MIU				
No. 2 tallow	40.0	35	None	2 MIU				
"A" tallow	39.0	15	39	2 MIU				
Choice white grease	36.0	4	13–11B	1 MIU				
Yellow grease	g	15	39	2 MIU				
Typical analyses—feed grade fats								
FGF (for all feeds)	29–45	40	37	2–4 MIU	40–100	25–50	50–75	4–40
FGF (for milk replacers)	38–41	5	9	1 MIU	47	50	50	4
All-beef tallow	38–43	5	7	1 MIU	47	50	50	4
All-pork fat	32–37	15	37	2 MIU	68	38	62	12
All-poultry fat	28–33	15	19	2 MIU	85	28	72	20
Acidulated vegetable oil soapstock	18–31	70		4–6 MIU	90–140	6–31	69–94	20–75
Palm oil	28–36	5		2	53	42	58	10

[a] *FAC* Fat Analysis Committee, AOCS

[b] American Fats and Oils Association specifications for tallows and greases

[c] Insoluble impurities 0.05% maximum

[d] Lard color maximum = 1.5 red Lovibond color (5.25-in. cell); Lard peroxide value 4.0 meq/kg max

[e] *MIU* moisture, insoluble impurities and unsaponifiables combined

[f] When required, titer to be negotiated between buyer and seller on contract-by-contract basis

Animal Fat Specifications, Production, and Utilization

Specifications, or typical analyses, of edible, industrial, and feed animal fats are shown in Table 34.9 [80, 86]. Much emphasis is placed on Titer, the solidification temperature of fatty acids in a saponified sample of the fat or oil (AOCS Method Cc 12–59). Feed ingredients in the United States may not be as aesthetically attractive as food ingredients, but are required to pass the same toxicology standards.

Rendering produced an estimated 4.18 million metric tons of animal fats in the United States in 2000 [87]. Of this amount, approximately 18 and 6% were edible tallow and lard, respectively, and 41 and 35% were inedible tallow and grease. Approximately 15 and 34% of the edible tallow and lard, respectively, and 37% of the inedible tallow and grease were exported. Of the inedible tallow and grease used in the United States, an estimated 75% was used as animal feed, 16% was converted to fatty acids by the oleochemicals industry, 4% was used in soaps, and 3% in lubricants. Inedible animal fats are the lowest cost domestic fat sources. Their market price per pound sometimes is less than fuel oil, and rendering plants have chosen to burn them as fuels. In 2001, animal fats were included with vegetable oils for federally supported trials of biodiesel fuel.

Fish Oils

Raw materials for producing fish oils include: (1) *pelagic-type* (surface feeding) fish pursued for reduction to meal and oil; (2) waste products produced at facilities that process edible fish; and occasionally (3) *by-catch species* also netted with the primary catch. The type of processing used depends on geographic location, species of fish normally taken in the area, and whether done at an on-shore plant or a factory ship. As in continuous wet rendering, whole fish or trimmings are ground, cooked, pressed, or decanted to yield solids and liquid, and the liquid fraction then centrifuged into stick water and oil [88, 89]. A significant fish protein industry, using trawlers for netting and mother ships for processing, has developed in Alaska. Local fish species have low oil content, and most of the oil produced is used on board to power steam boilers and engines [90]. Dried animal meals and fish-meals do not contain natural antioxidants, and it is common practice to preserve their fat by using synthetic antioxidants. This is especially important in fish meals, where rapid polymerization can generate sufficient heat to cause spontaneous combustion of stored meals.

Fish oils may be alkali-refined, bleached, hydrogenated, deodorized, and used in making margarine/spreads, other food products, and nutraceuticals. Direct food uses are approved, but seldom made in the United States. Because fish oil contains fatty acids with three or more double bonds, it readily polymerizes in the presence of air and is a major drying oil used for coatings. This property is further enhanced by *kettle bodying* (heating and mixing while bubbling in air). The fatty acid compositions of different fish oils, processing, and uses are presented by Bimbo [91], and Bimbo and Crowther [88]. Nonfood uses include animal feeds, fish attractants and lures, automotive gaskets, caulking compounds, ceramic deflocculants, core oils, fatty acids, fatty acid chemicals, fermentation substrates, fire retardants, fuel oil, illuminating oil, insecticidal preparations, leather tanning, lubricants and greases, mold-release agents, mushroom culture, oilfield chemicals, oiled fabrics, ore flotation, plasticizers, polyurethane lures, pressed wood fiber boards, printing inks, protective coatings, refractory compounds, rubber compounds, rust proofing, soaps, specialty chemicals, and tin-plating oils.

Feeding Animal and Marine Fats

Animal and marine fats provide approximately 2.5 times more calories per unit dry weight than carbohydrates or proteins, but are lower priced. Digestive tract capacities limit growth of broilers and turkeys and productivity of laying hens, and these animals respond well to high-energy diets, another name for fat-containing feeds. This also is true, to a lesser degree, for pigs and fish. However, high intakes of oil can disrupt normal function of rumens, and various bypass techniques are used in feeding cattle. The most common has been to hydrogenate fatty acids to melting points above rumen temperatures. Because of concerns about bovine spongiform encephalitis (BSE or "mad cow disease"), feeding of mammalian meat meals to ruminant animals has been outlawed in the United States and much of Europe, but these restrictions do not apply to feeding of tallows and greases. Concerns also exist about potential crossover of BSE to cats, and ruminant meat and bone meals are avoided in formulating dry cat foods.

Special needs must be addressed in pet foods. The lower melting point animal fats (choice white greases) are more appealing in odor than the tallows. However, lower melting fats can wick into and disfigure paper packaging. Solutions have included using the higher titer fats, including nonpermeable plastic layers in multiwall bags, laminating fat barriers onto paper, and filling into plastic bags. Some foods for guard duty and sled working dogs, and some fish feeds, require 30–40% fat content. However, high-fat formulas do not extrude easily even on modern machinery. Solutions have included using twin-screw extruders, which convey better than single-screw machines, the inclusion of full-fat soybean meal where the oil still is bound within its natural matrix, and enrobing the products with fat after extrusion and drying.

In addition to increasing caloric density and feed palatability, and improving appearance, feed efficiency, and reducing feed costs, feeding of animal, marine, and vegetable fats can:

1. Provide needed molecular structures through dietary EFAs and phospholipids
2. Increase blood glycogen levels and endurance in working animals such as horses and sled dogs.
3. Lower the heat of reaction during digestion and metabolism, thus increasing tolerance of heat by large animals during summertime or in tropical areas.
4. Prevent dermatitis and improve the appearance of skin and hair–an important effect of polyunsaturated fatty acids.
5. Carry fat-soluble vitamins and natural color compounds, for example, yellow pigments to improve the color of egg yolks; red and orange colors for feeding salmon.
6. Prevent segregation of mixed feeds.
7. Lubricate feed-processing machinery.
8. Bind heat-sensitive flavorings, vitamins, medications, "instant gravy mixes" to pet foods and feeds after extrusion and drying.
9. Reduce dustiness of feeds and improving animal health [86].

Row Crop Oilseeds Processing

Extraction Plants

The term *row crop* in this chapter generally means annual crops that are grown in rows. Their seeds have many similarities in extraction and oils processing and often can be handled in the same facilities if provision is made for differences in dehulling requirements. As a group, these oils contain appreciable amounts of phosphatides, which must be

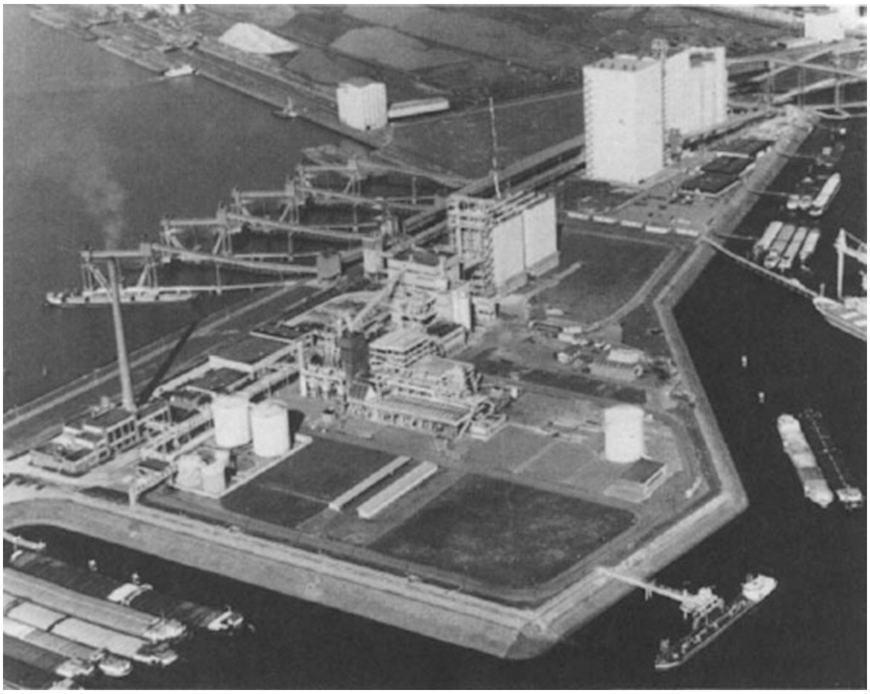


Fig. 34.7 Photograph of an operating soybean extraction plant. (Courtesy of Archer Daniels Midland Company, Decatur, IL)

removed, and are not as readily processed by physical refining as tree seed pulp crops such as palm and olive oils; instead, they must be prepared for lower temperature deodorization.

Concerns about marketing the meal (accounting for 65–75% of returns per bushel of soybean grown), and disposal of refinery byproducts, have reshaped the domestic soybean processing industry in the last 20 years. Soybean extraction plants now are located close to domestic meal feeding markets, and/or on major barge waterways and railways to reduce transportation costs to export facilities. Byproducts of refining are difficult to handle, and their production exceeds market demands. The major oil processors now generally locate extraction plants and oil refineries on the same property to enable spraying and drying surplus gums (crude lecithins) and soapstock on the meal for sale as animal feed. Independent edible oil refineries have almost disappeared, although refineries already associated with an extraction operation may purchase crude oils from other sources.

Economies of scale have led to minimum capacities of about 2,000–3,000 metric tons per day for new United States regional soybean extraction plants, and larger facilities (4,000–6,000 t per day) for soybean plants that pool some of their output for export. At the United States average yield of 38 bushels/acre, each 1,000 metric tons per day solvent plant capacity requires the output of 966.65 acres/day, or 338,328 acres/year (136,975 ha). A 2750 metric tons per day plant would use the soybean crop of 930,402 acres/year (376,681 ha). An extraction plant of this size will support a 500 t per/day refinery. Regional plants are smaller than export plants because accumulating this much soybean, in competition with other processors, can be difficult in some areas of the United States.

The largest solvent extractors can process 8,000 and 10,000 metric tons of soybean per day and are used primarily for the international trade. Two large extractors are used in the world's largest oilseed extraction plant in Argentina, reported to process 16,000 metric tons of soybean per day. Installations typically include: (1) facilities for unloading railroad cars, barges, or ships; (2) storage for at least several weeks' supply of seed, solvent, meal, and crude oil; (3) seed cleaning, preparation, oil extraction, and meal desolventizing equipment; (4) an on-site oil refinery; (5) repair and maintenance shops; (6) a quality control laboratory; and (7) offices and locker rooms for supervisors and workers. A photo of an operating soybean extraction seaport plant is shown in Fig. 34.7.

Seed Preparation for Extraction

A general flow sheet for direct solvent extraction of many row crop oilseeds is shown in Fig. 34.8. Initial quality of the seed and its preparation for extraction have the most effect on yield of extracted oil, subsequent required refinery operations, and yields of (saleable) *neutral oil*. Freshly harvested seed should be cleaned of trash, which may become ignited during drying, or harbor moisture that accelerates seed heating in storage. Oils of most good quality, dry, row crop seeds contain 0.5–0.75% FFA (AOCS Method Aa 6–38, Ac 5–41) with up to twice this range often accepted in trading without discounts. Rising pile temperature and FFA signals seed deterioration. The maximum moisture content for holding seed for long periods without spoilage varies with storage temperature and ranges from 8 to 13% among species. It is inversely related to the oil content, because less protein and carbohydrate is available to compete for water in high-oil seeds; however, the optimum relative humidity for storage is constant at about 65–70% for all seeds. If not adequately cleaned before storage, the seed should be completely cleaned before further processing to prevent clogging and damage of equipment. The processing system should be well equipped with magnets to arrest tramp metal that arrives with the seed or that is shed by equipment. Electronic metal detectors now are being installed as occurrence of stainless steel tramp metal increases.

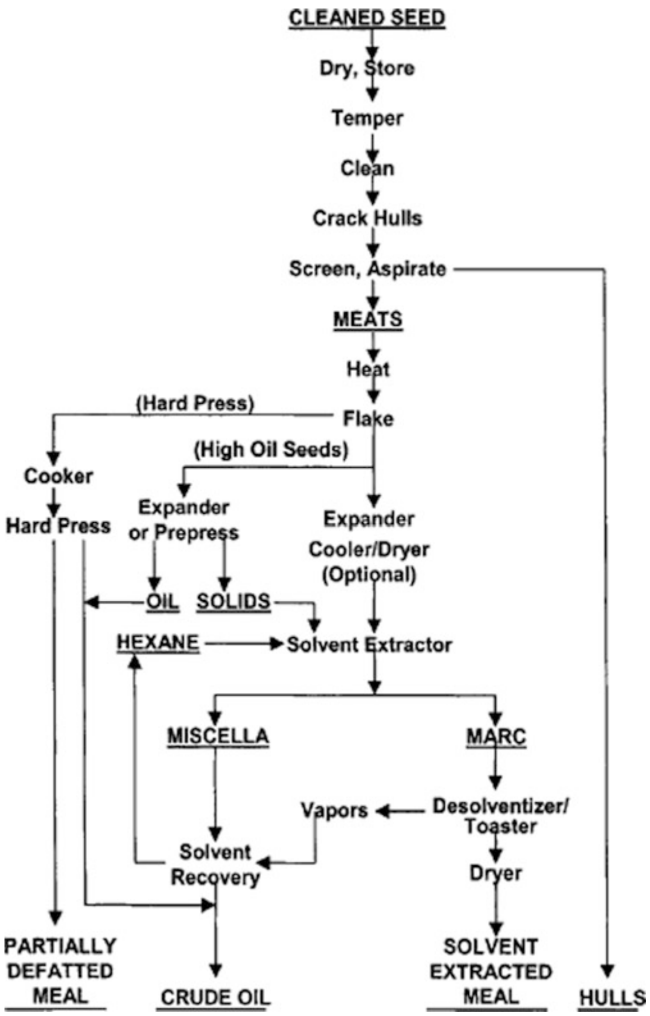


Fig. 34.8 General flow sheet for extraction of row, crop oilseeds

Oilseeds do not have fat cells like those of animals for storing fats. Instead, oil is stored in microscopic globules throughout the cells of dicotyledonous oilseeds, or in corn germ or rice bran. Yields and processing costs are highly dependent on the effectiveness of preextraction operations to disrupt cells and free the oil for recovery. Operations differ among various oilseeds, mainly in techniques of dehulling. Traditional processes have heated soybean and let it stand (temper) for several days to loosen the hull (Fig. 34.9) [92]. They are being replaced by hot dehulling systems which loosen the hull and crack the seed in one operation immediately before flaking and extraction (Fig. 34.10). Avoiding a second heating step saves energy.

Hulls are removed in a two-step process, called *dehulling* or *decortication*, in which seed is first cracked, and the hulls removed by screening and/or aspiration to obtain "meats" for processing. The major objective of removing the hulls is reduction of fiber content in meals for feeding poultry and swine, resulting in increased protein content and reduction of volume of material sent to the extractor. Partial retention of hulls in the meats was required earlier to improve traction in screwpressing operations, but hulls are no longer required with availability of expanders equipped with oil drainage cages. It is common practice either to leave sufficient hulls with the meats to just surpass minimum protein guarantees of meals, or to adjust high protein meals to industry trading standards

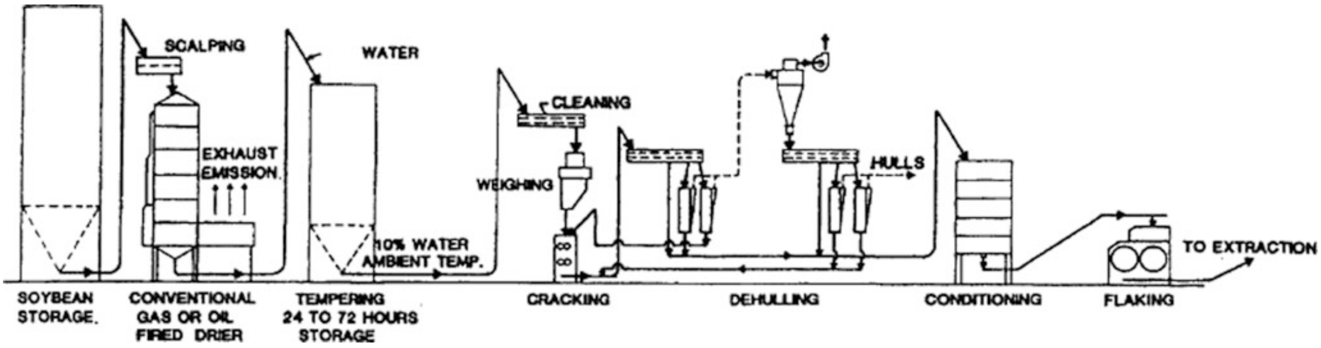


Fig. 34.9 Preparation system for conventional dehulling and flaking of soybeans. (From Moore [92], 141A–144A, With permission)

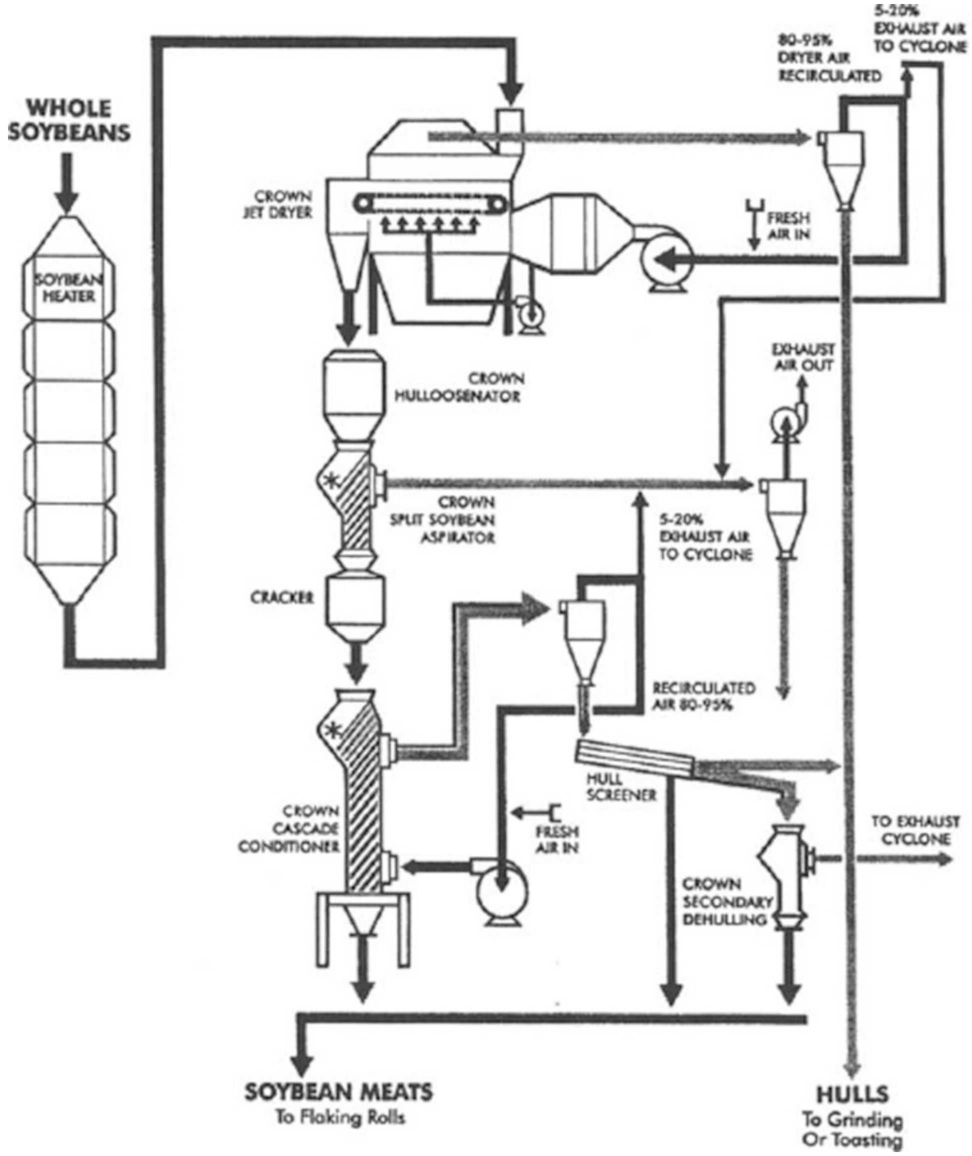


Fig. 34.10 Patented crown iron works hot dehulling system for soybean. (Courtesy of Crown Iron Works Company, Minneapolis, MN)

by adding back hulls after extraction. Currently, dehulled (low fiber) soybean meal trades at 48% protein, nondehulled soybean meal at 44%, cottonseed meal at 41%, and dehulled sunflower seed meal at 42% protein.

The meats are heated by steam injection to soften and increase the moisture content for plasticizing if needed in flaking. Cookers and heaters, used in the oilseeds processing industry, often are shallow circular ring-cooker pans with sweep arms in a multi-stack design, as shown in Fig. 34.16. In earlier processes, seed was heated to about 74°C/165°F before flaking to about 0.3-mm (0.012-in.) thickness. Now, it is realized that *phospholipases*, enzymes that make the phospholipids nonhydratable and more difficult to remove from the oil by water degumming, are highly active at this temperature, and seeds preferably are heated to less than 57°C/135°F or higher than 85°C/185 F to avoid the range of maximum phospholipase activity. Often, the flakes next pass through an expander for rapid heating to 105–121°C/220–250°F for homo-genization and shaping into collets.

Before the mid-1980s, processing concepts were based on classification of oilseeds into two groups. Meats containing over 30% oil on a dehulled or as-processed basis, including rapeseed/canola, oil-type sunflower seed, peanut, safflower seed, and copra, were considered *high-oil seeds*. Typically,

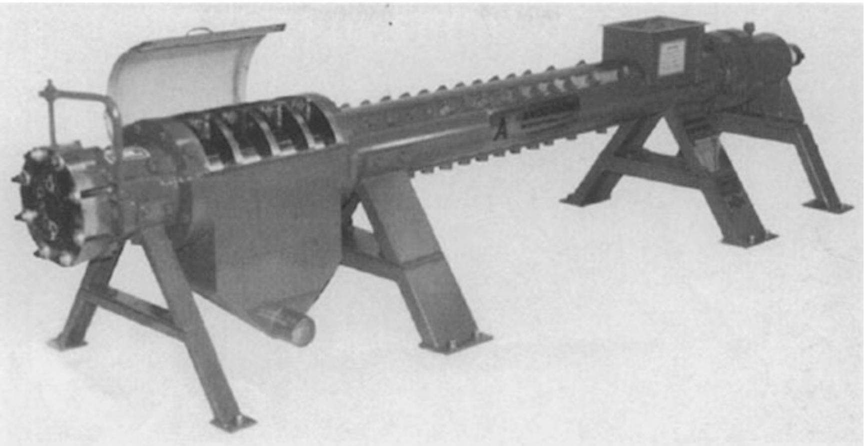


Fig. 34.11 Anderson International Corp. Hivex-Series-Expander™ with oil drainage section for preparing extraction collets from high-oil-content seeds. (Courtesy of Anderson International Corp, Cleveland, OH)

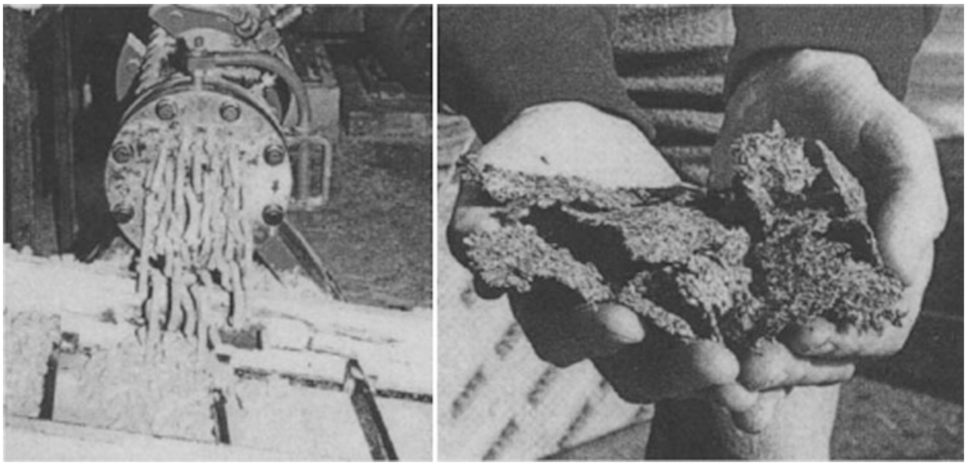


Fig. 34.12 *Left*: Soybean collets for extraction exiting die plate head, Solvex™ expander; strands break into random lengths. *Right*: Puff sheets made by hydraulically positioned cone choke head on expander. (Courtesy of Anderson International Corp., Cleveland, OH)

these seeds were dehulled (shredded in the case of copra); heated, flaked, and hard pressed, leaving 4–6% residual oil in the meal. Processors desiring to recover additional oil would first prepress high-oil seeds to 15–18% oil content using lighter-duty screw presses, and then solvent extract the press cake to less than 1% residual oil content; this process is called *prepress-solvent extraction*. Hard pressing was considered too inefficient for low-oil content seeds such as soybean unless nearby markets were available for the oil-rich meal. Soybean typically has been direct solvent extracted to less than 0.75% residual oil content. Cottonseed was extracted originally by hard press, later by prepress-solvent extraction, and now mainly by expander-direct solvent extraction techniques.

Introduction of the expander, a high-shear extruder with an interrupted-flight screw, in the mid-1980s revolutionized oilseed extraction practices. Essentially all solvent-extracted cottonseed and approximately 70% of domestic soybean tonnage now are processed with expanders. The expander heats, homogenizes, and shapes seeds or flakes into porous collets (pellets) that are more dense (weigh more per unit volume) than flakes, but are more rapidly extracted, approximately doubling the throughput of continuous solvent extractors. In effect, solvent extraction is changed from a diffusion process to a leaching process. Even though the expander homogenizes the seed, prior flaking still enhances oil recovery, but can be done at 0.5-mm (0.020-in.) thickness instead of the typical 0.3 mm (0.012 in.). The flakes or collets are cooled, with some drying occurring, to about 6°C/10°F below the boiling point of the solvent before entering the extractor. Collets also drain more completely than flakes, greatly reducing steam costs for desolventizing the extracted meal.

The Anderson International Company of Cleveland, Ohio, has patented an expander with a drainage cage (Fig. 34.11) to reduce the oil content of high-oil seeds to less than 20%, thus enabling the production of collets for direct solvent extraction from completely dehulled seeds such as sunflower seed and peanut. Replacement of the die plate (Fig. 34.12) with a hydraulically operated cone discharge head in the mid-1990s solved many of the problems first experienced in using oil-drainage cage-equipped expanders.

Introduction of the expander has enabled extraction plants to handle additional seed species, with purchase of only minimal cleaning and dehulling equipment where needed.

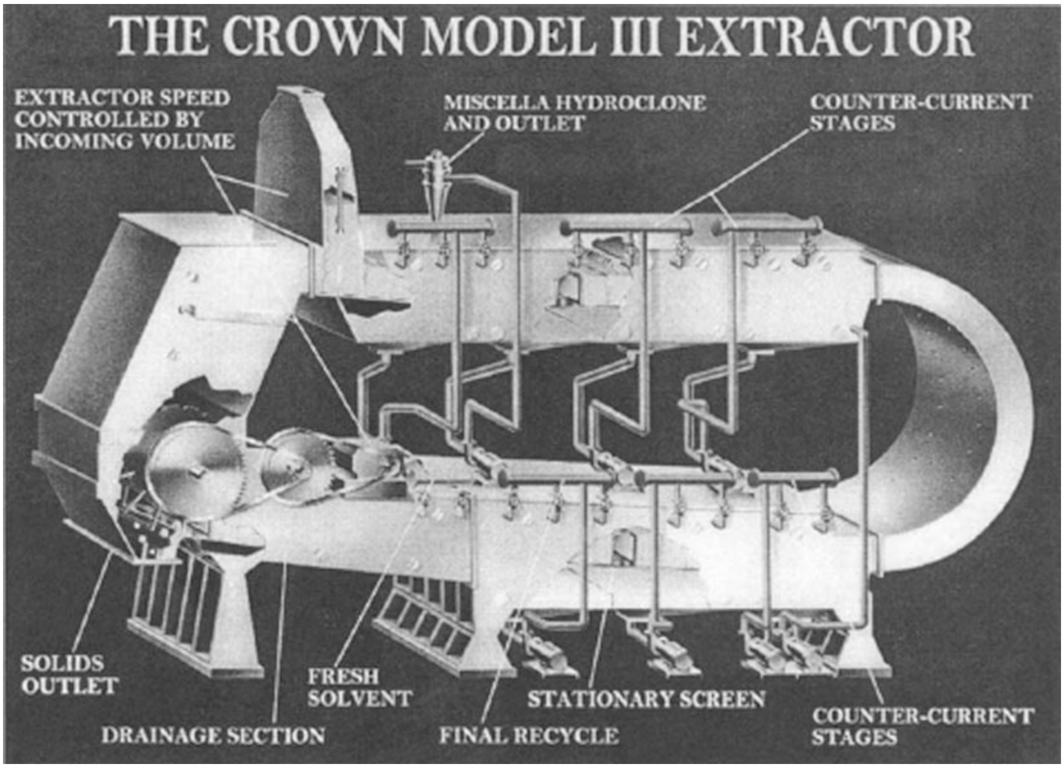


Fig. 34.13 Rectangular loop-type continuous countercurrent solvent extractor. (Courtesy of Crown Iron Works Company, Minneapolis, MN)

Prepress-solvent extraction facilities are being replaced by expander-direct solvent extraction equipment, leaving two basic extraction processes in modern large volume oilseed extraction plants: expander-direct solvent extraction, and hard press for applications where seed supplies are limited, or other considerations do not warrant construction of solvent extraction plants or the expense of skilled personnel and additional safety precautions for their operation. Hydraulic cage presses still are used in processing industrial crops such as castor seed, and for edible oils in developing countries.

Solvent Extractors

Hardly any batch-type oilseed solvent extractors remain. Three of the more popular types currently manufactured include: (1) *shallow bed-type* extractors, where a 0.5–1.5 m thick layer of collets or flakes is pulled across a linear screen and extracted by drenching with a countercurrent flowing miscella consisting of solvent and solubilized oil (Fig. 34.13); (2) *diffusion belt type*, where deeper beds of collets or flakes are conveyed on a woven mesh or folding-pan belt while drenched in countercurrent fashion with miscella (Fig. 34.14); and (3) *deep bed-type* which are constructed as carousels with pie-shaped cells that are filled with collets or flakes (~3 m deep), and extracted in countercurrent flowing fashion by drenching with miscella. *Marc* (wet extracted flakes/collets) dropping doors and moving parts have been eliminated in the newer carousel models, and the cells are revolved across a fixed screen (Fig. 34.15). Shallow bed extractors are built in capacities of up to 8,000 metric tons per day, and carousel-type extractors at up to 10,000 metric tons per day.

Solvents

Many solvents have been proposed for extracting oilseeds, but later found ineffective; others were used for a period, but disallowed because of health concerns about residues in food and feed products, or worker exposure [93]. All extraction solvents approved currently are flammable. Most commercial oil extraction currently is with hexane, a mixture of petroleum refinery fractions with a boiling point of 65–68°C/145–155°F, that consists of at least 60% *n*-hexane, with the balance being short-chain homologues and branched compounds. Some plants are using isohexane, which boils at a lower temperature. The US-EPA raised many concerns about extraction solvents in the late 1980s and 1990s. The FDA's position that *n*-hexane is a neurotoxin was put aside after industry-sponsored research showed the problem does not exist in mixed solvent systems [94]. The EPA also raised concerns that discharged volatile organic compounds (VOC) are contributors to ozone production [95]. A major solvent

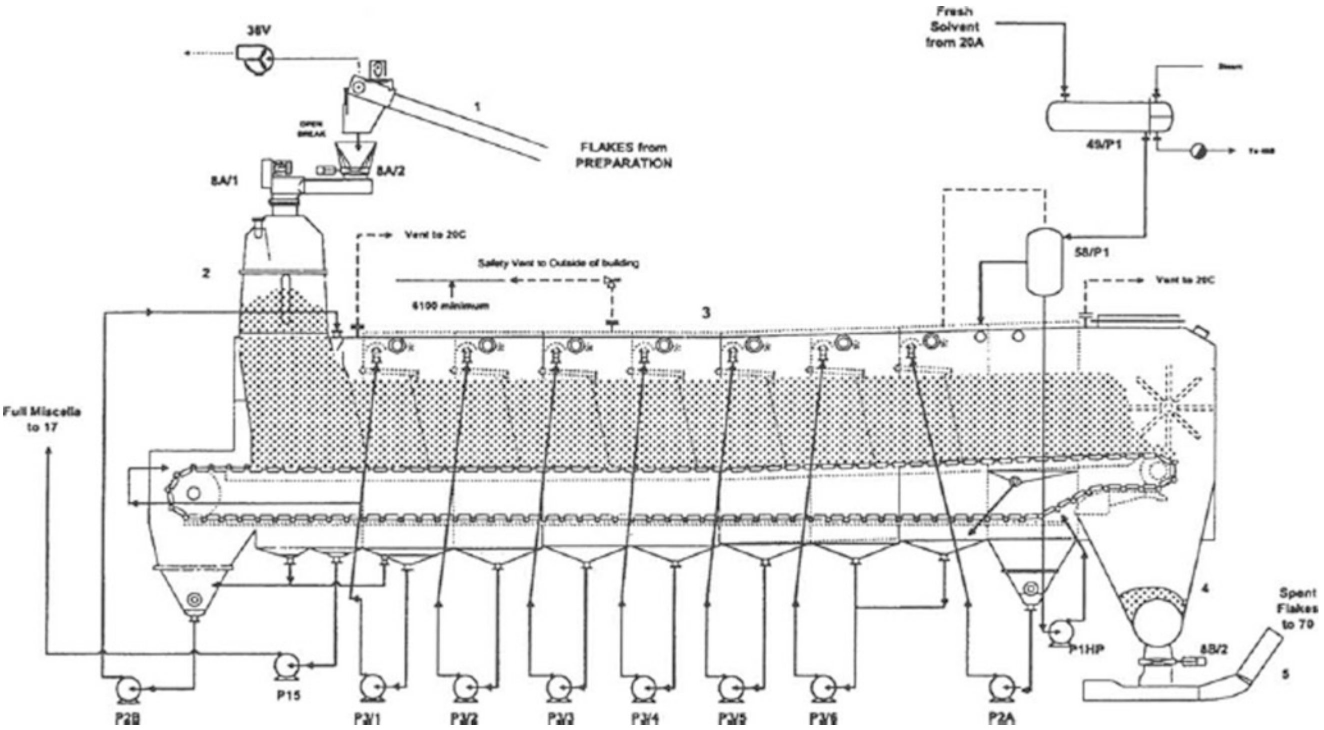
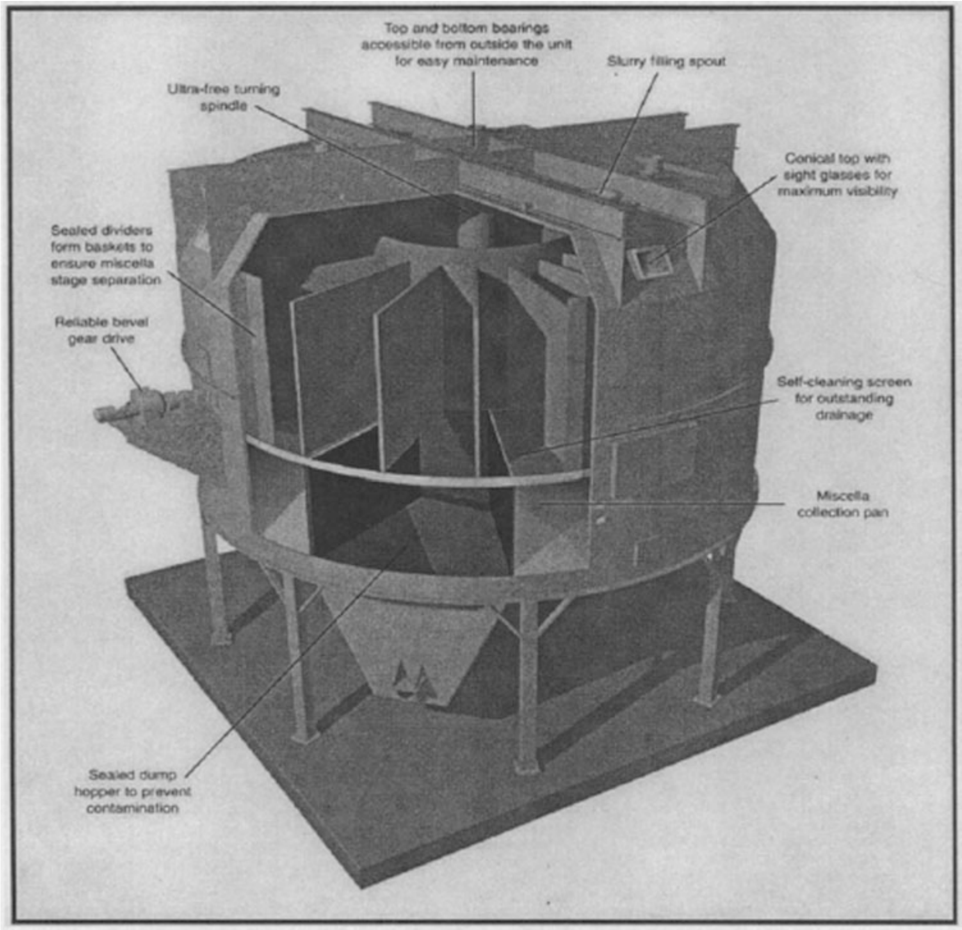


Fig. 34.14 Drawing of DeSmet LM perforated belt diffusion-type extractor. (Courtesy of Desmet Ballestra Oils and Fats, Brussels, Belgium)

Fig. 34.15 Reflex™ "basket or circular type" 10,000 t/day extractor. *Note*: Basket revolves within shell. (Courtesy of Desmet Ballestra Oils and Fats, Brussels, Belgium)



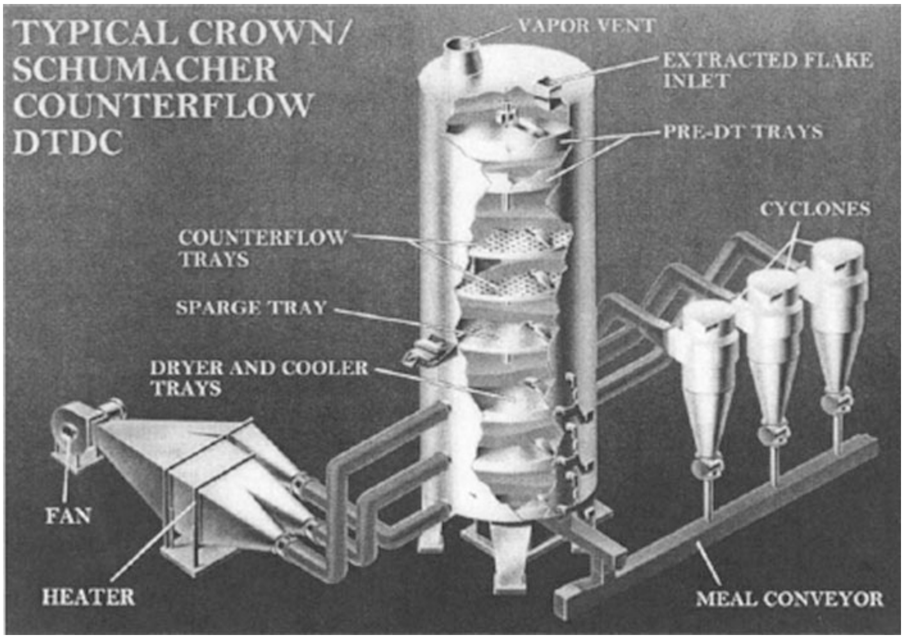


Fig. 34.16 Schumacher type desolventizer–toaster–dryer–cooler (DTDC). (Courtesy of Crown Iron Works Company, Minneapolis, MN)

containment program was instituted and resulted in domestic extraction plants reducing hexane losses from as high as more than 1 gal per ton of seed extracted to 0.25 gal, with some plants reducing losses to less than 0.16 gal per ton seed processed. Extraction solvents are highly flammable, and plants are built, equipped, and operated under Standard 36 of NFPA International, formerly known as the National Fire Prevention Association [96]. Additional worker exposure limits, safety practices, and training are established by the Occupational Safety & Health Administration (OSHA) of the US Department of Labor.

A major project added to the information about isopropyl alcohol (IPA) from earlier trials [97, 98] and showed it can be as effective an extraction solvent as hexane [99]. However, because of high retrofitting costs, IPA is not likely to be implemented while hexane-type hydrocarbons are allowed.

Batch extraction of vegetable oils with high-pressure CO_2 (carbon dioxide) in a critical state (at pressures required for maintaining a liquid phase) was heavily researched in the 1980s [100]. A continuous process has recently been commercialized in which CO_2 is injected into the barrel of a modified screw press (HIPLEX®, High pressure liquid extraction, Crown Iron Works, Harburg-Freudenberger, personal communication). Soy meals produced using this equipment maintain high protein solubility, with PDI's (protein dispersibility index) >70%, and a residual oil context <4%. This technology is being implemented where markets for food application demand "organic" labeling of ingredients and high protein solubility is a functional necessity for the product. Eighteen 400 metric ton/day critical propane extraction plants were processing soybeans in China in 2002. But, domestic critical propane demonstration lines have experienced safety problems. A critical CO_2 extraction laboratory instrument, for rapid analysis of fat content in seed and meals, has been marketed domestically since the mid-1990s.

Desolventizing-Toasting

The extracted, drained marc contains approximately 25% hexane holdup in soybean collets and 33% in flakes, which is vaporized in a desolventizer-toaster (DT) under vacuum. Some DTs have cooling sections, but separate dryer-coolers (DCs) often are used in large installations. Steam is sparged into the marc as the heat source for volatilizing the solvent. The condensate must be removed subsequently by drying. The moist "toasting" operation destroys enzymes and antigrowth factors such as trypsin inhibitors and hemagglutinins in soybean [101] and reduces meal protein solubility and digestibility by rumen microorganisms, thus improving rumen by-pass or escape in feeding cattle and sheep. Figure 34.16 shows a cutaway drawing of a Crown/Schumacher design desolventizer-toaster-dryer-cooler (DTDC) line. The drained marc enters at the top of a stack of circular trays and is mixed by sweep arms and pushed to fall through slots to lower trays. The bottoms of the initial trays also are steam jacketed. As the marc progresses downwards, it encounters rising steam and solvent vapors and is "toasted" by moist heat. The steam provides heat for vaporizing the solvent, but leaves condensate. The solvent-water vapors, drawn off at the top of the DT, are condensed and the solvent recovered. The desolventized flakes or collets continue to work down through the trays where they are dried by hot air and then cooled by ambient air.

Miscella Refining

In processing most oilseeds, hexane is stripped from the miscella by distillation to produce a crude oil that subsequently is alkali or physically refined. However, gossypol and other pigments become extremely difficult to *bleach* if left in warm cottonseed oil for more than a few days. It is normal practice for cottonseed oil mills to send their crude oil immediately to an alkali refinery or to operate an on-site miscella refinery, where phosphatides, FFA, and color pigments are removed by alkali treatment of the oil-extraction solvent mixture. Cooling the crude oil as produced, until refining, also slows *fixing* of color.

In the process, miscella leaves the extractor at about 30–35% oil and is concentrated to approximately 65% oil by evaporation. The FFA in the concentrate then is reacted with alkali (sodium hydroxide solution) to produce soaps that are removed with other water-soluble compounds by centrifugation. Next, the solvent is removed from the miscella-refined oil by further evaporation, and the soapstock is spread on the meal in the DT to recover its solvent. Hexane vapors from the miscella and the DT are condensed, and the solvent is recycled to the extractor for reuse. The non-condensable gases are passed through a mineral oil stripper to recover the last traces of hexane.

Refining of Vegetable Fats and Oils

Technically, *refining* means alkali neutralization of FFA in the oil. But over time, all postextraction processing of oils has become known as *refining*, and the facility in which it is conducted a *refinery*. Conversions of the resulting ingredients into margarines and spreads, and bottling oils, often are done at different locations or companies. The objectives in refining and processing edible fats and oils include: removal of FFA, phospholipids (*gums*), oxidation products, color and off flavor/odor compounds, and toxic substances to produce light-colored, bland products with long shelf lives; obtaining a mixture of TAG with the desired solids content profiles over the temperature range of product use; and preparation and storage of semi-solid products with desired textures. A flow sheet for refining and processing fats and oils is shown in Fig. 34.17. Refining procedures reliably purify oils extracted from cottonseed, peanut, and corn germ that have been contaminated with (water-soluble) mycotoxins or pesticides, but the resulting meals may have to be destroyed, used as fertilizer, or further treated to inactivate aflatoxins.

Oil Receiving and Handling

Maximizing yields of saleable oil requires even more detailed analyses and attention to lot-to-lot differences during refining than in preparing oilseed lots for extraction. Using soybean oil as an example, the first priority on receiving a shipment, or the output of an adjoining extraction plant, is to characterize the overall quality of the oil and determine what needs to be done to prepare it for market. The responsibilities of the refinery may include preparing freshly extracted oil ready for sale as *Crude Degummed Soybean Oil, Once Refined Soybean Oil,* or *Fully Refined Soybean Oil* for export under National Oil Processors Association (NOPA) Trading Rules [102], or other agreements made with the buyer.

To prevent hydration and precipitation of phosphatides during storage and shipping, the phosphorous content of crude soybean oil must be reduced to less than 0.02% (200 ppm) before entering the trade. This usually is accomplished by water degumming.

If the refinery purchases soybean oils for processing, one of the first tasks is to check composition of the received oil against the contract (usually NOPA Trading Rules), in as much as this determines the final price paid:

- *Flash point* (indicator of residual extraction solvent)—AOCS Method Cc 9c-95(97) (>250°F)
- *Unsaponifiable matter*—Method Ca 6a-40(97) (<1.5%)
- *Free fatty acids* (FFA), as oleic—AOCS Method Ca 5a-40(97) (<0.75%)
- *Moisture and volatile matter* (M&V)—AOCS Method Ca 2d-25(97) and insoluble impurities—AOCS Method Ca 3a-46(97) (<0.3%)
- *Phosphorous*—AOCS Method Ca 12-55(97) (<0.02%)

Additionally, the analytical laboratory estimates how much saleable oil can be produced from the lot, usually by neutral oil and loss (NOL) analysis (AOCS Method Ca 9f-57). In this procedure, a solvent-diluted sample of the oil is poured over a column packed with activated alumina (aluminum oxide). After evaporation of the solvent, the weight percentage of the oil that passed through the column is considered *neutral oil,* and the weight of oil retained is the "loss." The analyst may run a *bleach test* (AOCS Method Cc 83–63) or a *refining test* (AOCS Method Ca 9d-52, seldom used currently) if concerns exist about reducing the color of the oil to an acceptable range.

Phosphatides Degumming, Lecithin Uses

Phosphatides are essentially removed in modern oil processing to minimize fouling the bleaching earth, poisoning hydrogenation catalysts, and darkening the oil color by heat during deodorization and deep fat frying. Phosphatide contents

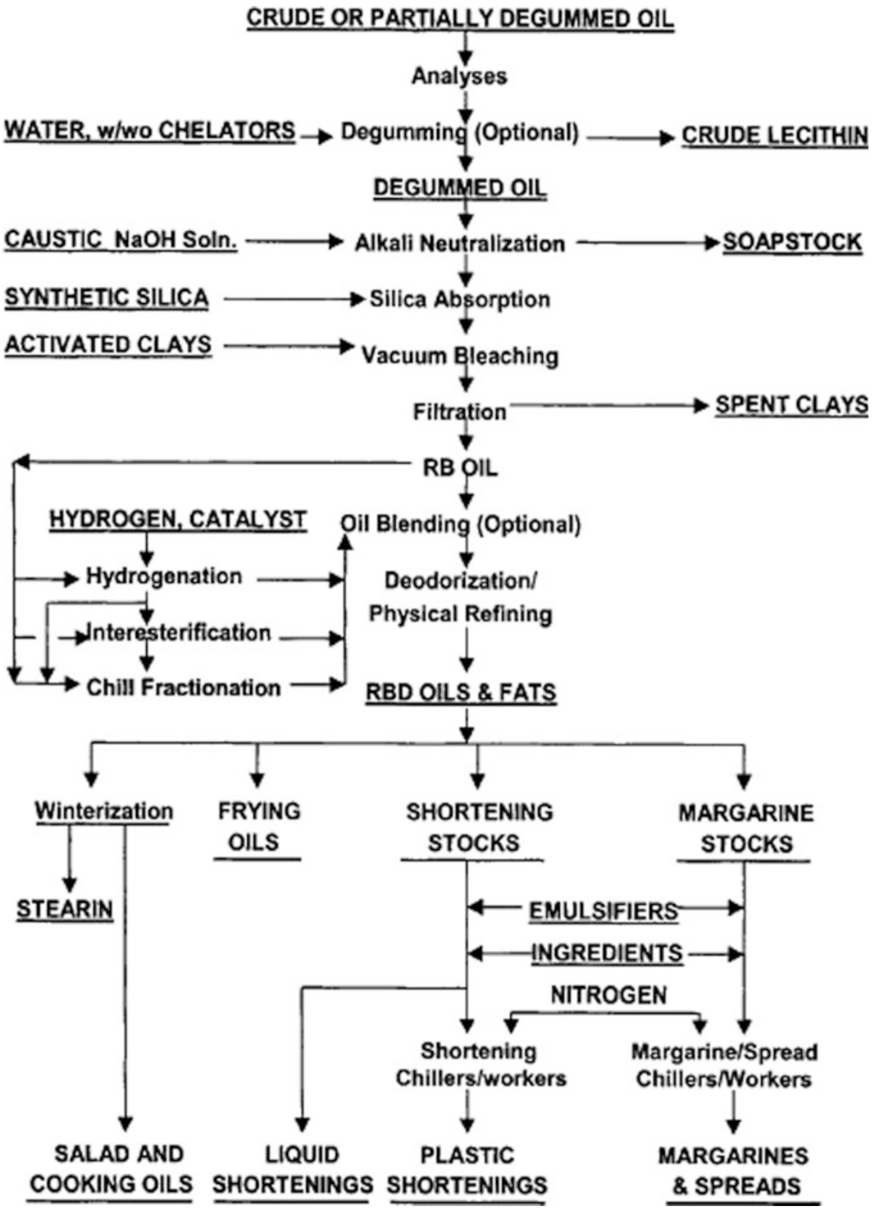


Fig. 34.17 Composite flow sheet of oils and fats refining and processing

of common vegetable oils are shown in Table 34.10. Soy, corn, and canola phosphatides are separated in North America, and some of the former Soviet Union countries separate sunflower seed phosphatides. For soybean oil, the relationship between phosphatide and phosphorous content is:

$$[\text{phosphatide } (\%) \times 10^4]/31.7 = \text{phosphorous ppm}$$

If lecithin is saved at the refinery, the hydratable phosphatides are separated by simple *water degumming* (hydration with deionized water, followed by centrifugation). However, some of the phosphatides will have been converted to NHP by enzyme action. In this case, a chelating agent (phosphoric acid usually because of lower cost) is added to the alkali in the FFA neutralization step to return the NHP to hydratable form. Palm oil has very low phosphatide levels and is physically refined after acid degumming and bleaching. The addition of chelating acids to crude row crop oils before alkali refining results in removal of hydrated phosphatides with the centrifuged soapstock, but viscosity is high and appreciable losses of neutral oil can occur by occlusion. Thus, many soybean oil refineries, without markets for crude lecithin, still run preliminary degumming operations and spread the gums with the later removed soapstock over the desolventized marc for drying in the dryer-cooler and sale as part of the soybean meal [103].

Commercial lecithin is produced by water degumming (precipitation from oil with ion-exchange treated water), separation by stacked disk centrifuge, and vacuum drying to less than 1% moisture content. Crude lecithins contain

Table 34.10 Phosphatide contents of common vegetable oil (From: Farr [103], With permission)

Type of oil	Phosphatide content As (%)	As Phosphorous[a] (ppm)
Soybean	1.0–3.0	311–940
Corn	0.7–0.9	220–280
Safflower	0.4–0.6	130–290
Sunflower	0.5–0.9	160–290
Peanut	0.3–0.4	95–190
Canola (super degummed)	0.16	50
Canola (crude)	1.0–3.0	311–940

[a] Phosphorous calculated as: [phosphatede (%) × 10^4]/31.7 = phosphorous (ppm)

70–72% acetone insolubles (AI) and are standardized to 62–64% and an acid value of 30 by addition of oil and fatty acids before sale. Crude lecithins may be treated with acetone to obtain free-flowing powders with 95–98% AI. Lecithin can be additionally purified, bleached, fractionated, hydrogenated, hydroxylated, acetylated, sulfonated, and halogenated [104]. One domestic company makes 13 kinds of lecithin for food uses alone.

Food applications of lecithins include: emulsification, wetting and dispersing agents, modification of baking properties, pan release agents, viscosity reduction of melted chocolate, anti-spattering agents in margarine, antioxidant effects, and nutritional supplements. Feed uses include wetting and dispersing agents in calf starters, and nutritional supplementation. Choline is a recognized vitamin, and inositol has been found essential in feeding some species of fish. Industrial uses include: emulsification and dispersion of active agents in pesticides, dispersing agents and stabilizers in paints and magnetic tapes, softening agents and penetrants for leather, and softening and lubrication of textiles. Cosmetic uses include: foam stabilizers and emollients in hair care, and emulsification, emollient, refitting and wetting agents in skin care. Pharmaceutical applications include: emulsifiers in parentoral nutrition, carriers and softening agents in suppositories, and emulsification and penetration improvement in cremes and lotions [105].

Some natural NHP always are present in crude oil, but development of additional NHP during seed extraction can be minimized by heat inactivation of phospholipases as explained earlier. In preparation for degumming, a tank large enough to supply the refinery with uniform oil feed stock for a suitable period of time is filled and mixed. Samples are taken for FFA, calcium, and magnesium analyses. Earlier practices of adding an amount of water equal to the weight of the phosphatides have been defined more precisely, and:

$$\text{added water} = (\text{ppm } P \times 3.17 \times 10^{-4}) \times 0.7$$

is recommended. The amount of phosphoric acid used is:

$$H_3PO_4 = [(Ca + Mg)/2] \times 10$$

with all components expressed in ppm.

The phosphoric acid and water are added to the warmed (65°C/150°F) crude oil stream, pass through a high-shear mixer, and are pumped to the stirred hydration tank. Details of the process, using an Alfa Laval (Lund, Sweden) PX-90 centrifuge rated at 33,000 kg/h for degumming, are shown in Fig. 34.18 [7]. After degumming, the crude oil is vacuum dried to <0.3% moisture and volatiles content and cooled to 50°C/120°F for storage or shipment. But, this step can be omitted if the oil is refined immediately [103].

The objective in acid degumming is to chelate the calcium and magnesium ions and render the nonhydratable phosphatide forms hydratable. In addition to phosphoric acid, citric and malic acids are effective, as well as ethylenediaminetetraacetic acid (EDTA). Acid-treated phosphatides are not used for production of commercial lecithins. Extensive reviews on oil degumming have been prepared [106, 107]. Lurgi, a German equipment manufacturer, has developed an EnzyMax™ process that cleaves the NHP with a phospholipase B at the triglyceride's second carbon to produce a lysophosphatide that is insoluble in oil and is removed by centrifuging [108].

In earlier times, the NHP content was determined by analyzing a water degummed sample for phosphorous, but the procedure took too long for commercial use. In reality, phosphorous is only a marker and loss of hydration actually results from the presence of divalent cations, primarily calcium and magnesium. Refineries now use induction coupled plasma (ICP) spectrographs for analyzing divalent cations content rapidly in aspirated crude oil and adjust the amount of phosphoric acid used for each "day tank lot" of analyzed oil. ICP units cost in the range of $75,000–125,000, but, reportedly, pay for themselves through increased oil yields in as little as 6 months in mid-size (500 t/day) refineries.

The crude oil from which gums are taken for lecithin production still contains non-hydratable phosphatides, but can be treated with a chelating agent before alkali neutralization and will be removed with the soapstock by centrifugation. Provision must be made for the added acid in calculating the amount of neutralizing alkali added.

Alkali Neutralization

The elimination of neutralized oil wash water is the major breakthrough of the decade in oils refining, with *modified caustic refining* or *silica refining* processes still being optimized. Phosphatides are removed by degumming, as already described, and FFA in the crude oil is still

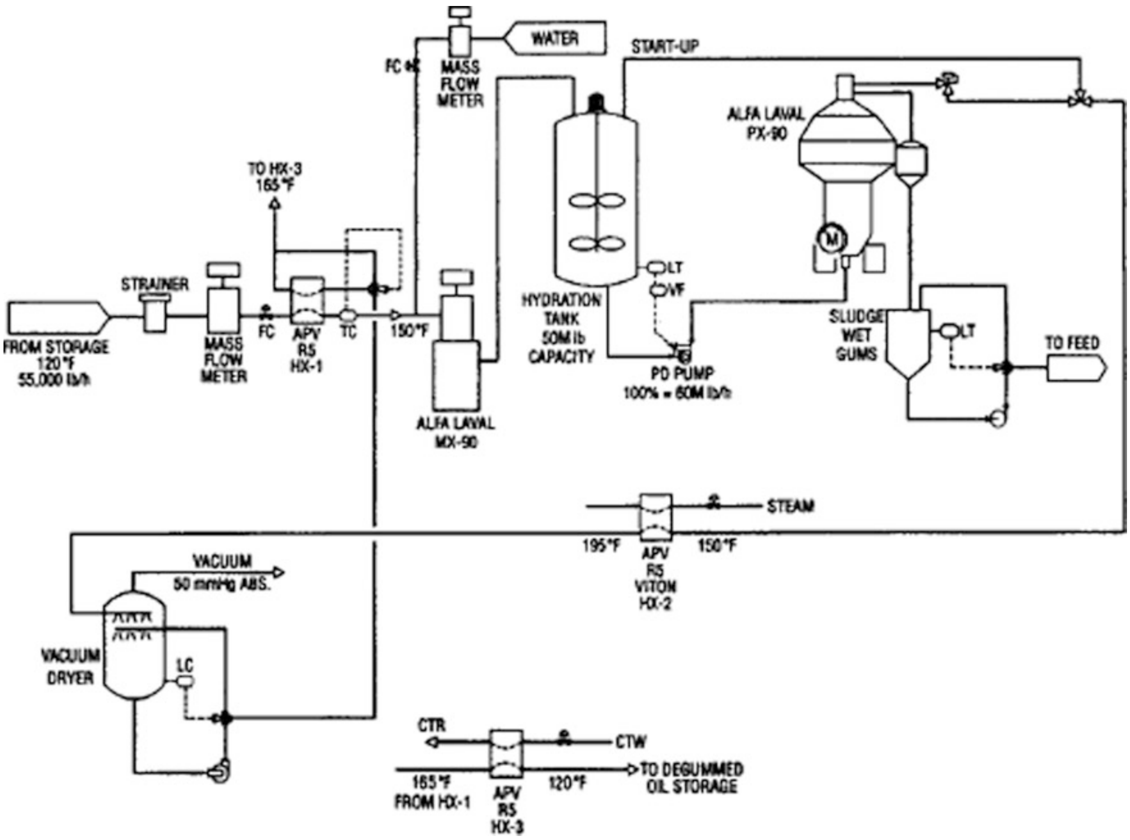


Fig. 34.18 low sheet of a modern water or acid degumming line. (From Farr [103], With permission)

neutralized with sodium hydroxide (*caustic*) solution (20–50°Be for cottonseed oil, and 16–24°Be for soybean, sunflower, safflower, peanut, and corn *oils*). But, the refinery then has the choice of water-washing the oil after removal of soapstock, or adsorbing the residual soaps on silica hydrogel before bleaching, thus eliminating problems of disposing the wash water. Process demonstration by the W. R. Grace Company began in 1986, with increasing growth in the commercial use of Trisyl™ in the 1990s. Several silica hydrogel suppliers now exist [103].

Two major processes have been used for alkali neutralization of FFA in row crop oils: the *long mix* and the *short mix*. The short mix process evolved in Europe, runs at a higher temperature, and reportedly is effective with a number of oils. The long mix process was developed in the United States and has been championed for refining soybean oil. It respects the principle that chemical reactions occur more rapidly and are harder to control, at higher temperatures (van't Hoff rule that the speed of reaction doubles with each 10°C rise in temperature). Traditionally, the long-mix process starts with crude soybean oil at ambient temperature, uses a low-concentration caustic solution, and has a mixer retention time of 15 min after which the oil–caustic mixture is heated to 70°C/160°F to reduce its viscosity before centrifuging. In the short-mix process, crude oil is heated to 90°C/194°F, mixed with high-concentration caustic solution for 1 min and centrifuged [109]. The throughput per hour of both systems is the same because the additional volume for the holding time is built into the long mix line.

A flow chart of a long mix neutralization process is shown in Fig. 34.19. The recommended retention time for soybean oil in this system is 6 min and was shortened by a major improvement in caustic–oil retention mixers developed by the Alfa Laval Company, Tumba, Sweden, in the mid-1990s [103].

The amount of caustic treat to be added for neutralization of soybean oil is calculated as:

$$\begin{aligned}\text{percent treat} = {} & [(\text{factor} \times \text{percent FFA}) \\ & + (\text{percent excess})/\text{percent NaOH}] \times 100\end{aligned}$$

where factor = 0.142, NaOH is determined from the 20°Be of the caustic solution, and percent excess is selected from the following ranges based on experience in the specific refinery: degummed soybean oil, 0.01–0.05; nondegummed soybean oil, 0.15–0.25. Continuing with Fig. 34.19, soybean oil from storage is adjusted to 38°C/100°F, passed through strainers, mixed with the treat in the rapid mixer, held in the retention mixers for 6 min, heated to 60°C/140°F, and passed through the primary (first) centrifuge [103]. The soapstock is returned to the extraction plant with surplus

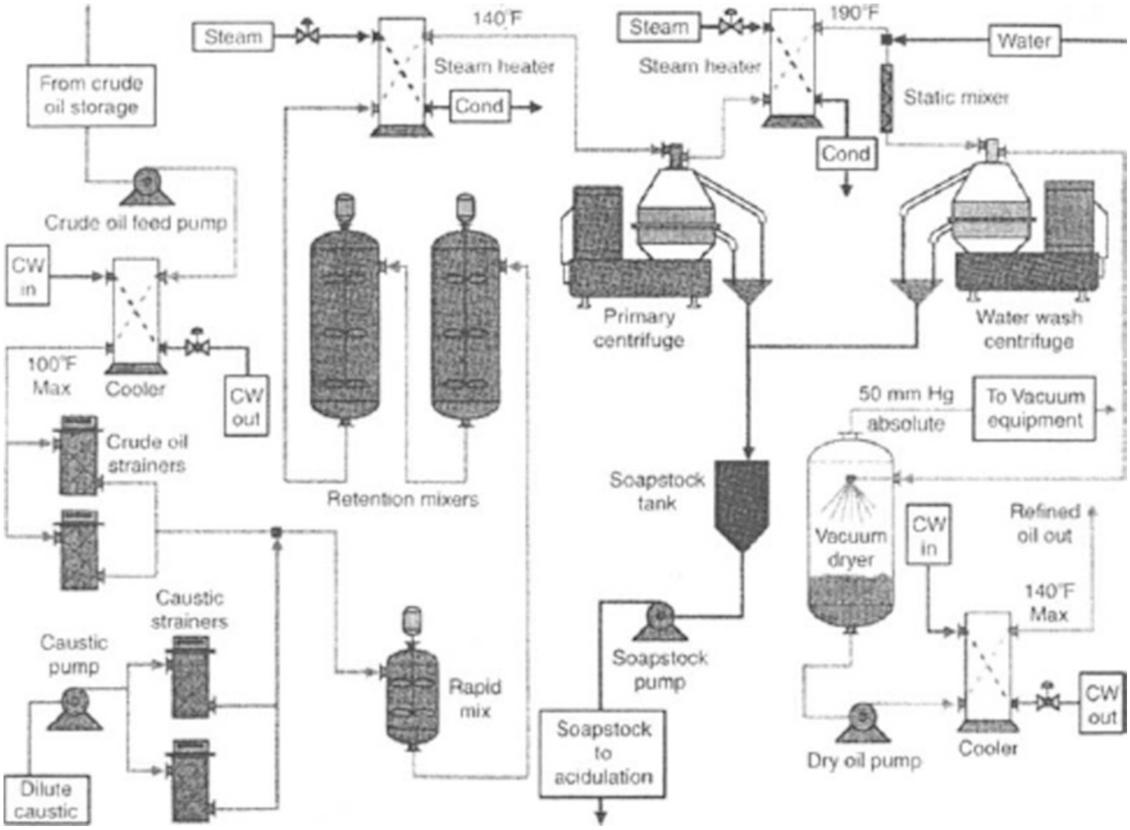


Fig. 34.19 Flow sheet of continuous refining of soybean oil (long mix process) with water wash option. (From Farr [103], With permission)

or acid-degummed phosphatides to be spread over the meal before the dryer-cooler. The refinery then can water wash and vacuum dry the oil by traditional methods, or treat it by silica gel adsorption. If a water wash is used, the oil is heated to 88°C/190°F with 12–15% soft water, held in an agitated tank for ~0.5 h, and centrifuged to produce an oil with <20 ppm soap. The soap can be reduced to "0" by addition of a small amount of phosphoric acid in the water-wash retention tank. The oil is next sprayed into a drying tank at 50 mmHg absolute [103].

Sodium silicate neutralization also has shown promise as a potential commercial method. The resulting soaps form a granular agglomerate, which can be removed by filtration to produce oil containing less than 100-ppm soaps. Thus, the costs of purchasing and maintaining primary and water-wash centrifuges and wash water disposal are avoided. The filtered oil can be treated with neutral or activated clays to remove color, peroxides, residual phosphatides, and soaps to produce oils with FFA fatty acids contents of less than 0.05% and "0" PV (*peroxide value*) [110]. The method has been patented and is being evaluated commercially.

Various researchers have reported on attempts to remove fatty acids from oils by ultrafiltration membranes. In the absence of membranes that can withstand extraction solvents, success generally has been limited. However, progress has been made on membrane degumming [111].

Silica Gel Adsorption

Silica hydrogels are very effective in removing phosphatides, residual soaps, and metal ions (all poisonous to hydrogenation catalysts), thus reducing the amount of bleaching clay required (by 50–75%) and leaving its function primarily to remove chlorophyll and secondary oxidation products. By eliminating soapy wash water, water discharged from refineries is reduced by ~50% in volume and has much lower biological oxidative demand (BOD). Additionally, cost and expenses of a second washing centrifuge are avoided [103].

The method of using silica hydrogel has changed during perfection of the process. Variations in procedure exist, but silica hydrogel can accept oil from the centrifuge at 0.2–0.4% water content. The currently recommended process consists of blending the silica hydrogel with oil directly from the soap removal centrifuge, with minimum, if any, drying of the oil. Silica hydrogel then is removed by filtration before mixing the oil with the bleaching earth [103].

Bleaching

Bleaching originally was a process for reducing color in oils, but the name has become a misnomer of this industry. Although limited color reduction occurs during the process, the major reduction of red and yellow colors occurs during the high heat of the deodorization process. The current practical function of bleaching is to remove chlorophyll and oxidation-degraded compounds and prepare the oil for hydrogenation or interesterification by scavenging the remaining soaps, phosphatides, and minerals that would poison the catalyst. Several authors have reviewed the theory and practical aspects of bleaching [112–114].

Bleaching earths are made from naturally occurring minerals, including palygorskite—also known as attapulgite, sepiolite, bentonite, and montmorillonite—and other minerals belonging to the aluminum silicate family. They may be used as such, but typically are preactivated by treatment with hydrochloric or sulfuric acids which: increase absorption by increasing surface area several fold; provide acid centers with catalytic properties; and impart ion-exchange properties to the clay. These properties are important in adsorbing various undesirable impurities in the oil, and in rendering complex organic structures adsorbable. In addition to adsorbing phosphatides and soaps, the process also removes pesticides, polycyclic aromatic hydrocarbons, *trans* and conjugated fatty acids, dimers, and polymers. Activated earths also provide a catalytic surface for breaking down peroxides. Decomposition is an exothermic reaction, with the heat apparently enhancing the *press effect* in color reduction of carotenoids in the filter press. The cation exchange property of the activated earth is credited with removing magnesium from the center of the chlorophyll complex and arresting its activity as a pro-oxidant. Cation exchange also is used for the removal of heavy trace metal pro-oxidants such as iron and copper, and for removing trace nickel in postbleaching of hydrogenated oil [114]. Although peroxides content is reduced, *p*-Anisidine Value (AV; AOCS Method Cs 18–90) increases. The AV is believed to estimate aldehydes (2-alkenals and 2,4-dienals) in animal and vegetable oils with the potential for later breakdown.

Bleaching clay load (typically 0.1–2.0%) and operating temperatures depend on the type and quality of oil processed [114]. Modern bleaching processes are conducted under vacuum (50 mmHg) to minimize later oil oxidation, and subsequent nitrogen blanketing of the oil during shipping and storage is recommended.

Close coupling of the refining and bleaching operations is highly recommended, especially when using the modified caustic refining or silica refining processes [103]. Because of the high level of unsaturated oils and peroxides in spent bleaching earth, it is very susceptible to spontaneous combustion unless quenched with water. Disposal in landfills is becoming increasingly difficult. Spreading spent bleaching clay on soybean meal for animal feeding is done in limited quantities, but caution should be taken because, by absorbing pesticides and mycotoxins, bleaching is one of the two safety valves in processing oils. The other is collection of volatile pesticides in the condensed deodorizer distillate.

Oils and Fats Modification

If processing proceeded next through deodorization, the resulting product would be liquid and known as “RBD oil” (refined, bleached, deodorized), and suitable for cooking and table use. However, consumers often want: oils that remain crystal clear when stored in household refrigerators; fat solids for an array of products such as butter-, ghee-, and lard-replacement spreads; air-entrapping shortenings for making tender, light-textured cakes; brittle or soft fats for toppings and frostings; cocoa butter replacements and substitutes; prolonged and prolonged bottled oil and frying lives; and fresh flavor in fried cereal-based snacks and nuts. Various techniques are used to craft fats with temperature-solids profiles such as the selected shortenings shown in Fig. 34.20 [115] from liquid oils. Modification processes include thermal fractionation, hydrogenation, and interesterification usually done before deodorization. Thermal fractionation and hydrogenation can use well-purified RB oil. Interesterification requires RBD oil with low FFA content. Hydrogenated and interesterified oils are postbleached and deodorized.

Solid Fat Index/Solid Fat Content

Two systems are used to characterize solids content of temperature-profiled fats. The Solid Fat Index (SFI) (AOCS Method AOCS Method Cd 10–57) uses dilatometers and was developed in the United States. It is the older method and is effective for fats containing up to 50% solids at 10°C/50°F. The Solid Fat Content (SFC) (AOCS Method Cd 16–81) uses pulsed nuclear magnetic resonance (NMR) techniques to quantify crystalization. It was developed more recently to accommodate palm oil and its products and is used in most other countries. SFC is effective for solids contents up to 95%; however, additional care in sample tempering may be required at the upper solids limits for either SFI or SFC. In both methods, a sample of the fat is completely melted to destroy its crystal memory, then chilled to 0°C to set the crystals. Readings are then taken at selected temperatures (10°C/50°F, 21.1°C/70°F, 26.7°C/80°F, 33.3°C/92°F, 37.8°C/100°F, and 40°C/104°F) to develop SFI or SFC curves. In some industries, five-point or three-point readings are taken; readings also may be made

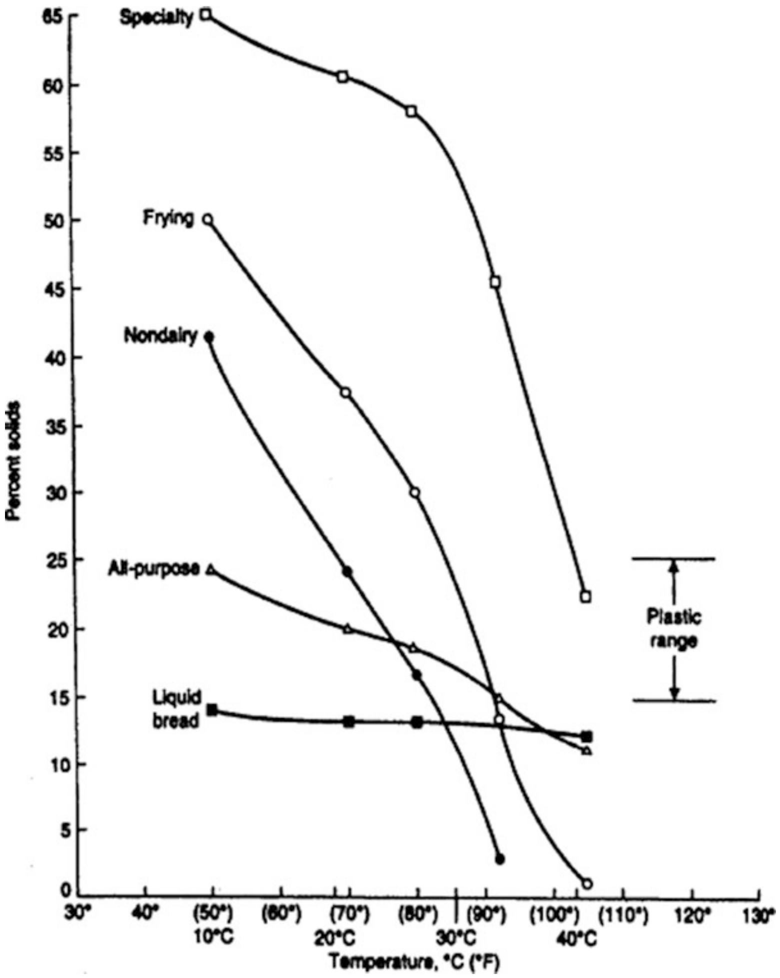


Fig. 34.20 Solids–temperature curves for various types of shortenings. (From O'Brien [115], With permission)

at higher temperatures. SFI and SFC give similar, but not identical results. A recent comparison of the two methods concluded that SFC reads higher at lower temperatures (10 and 21.1°C), but similar to SFI at higher temperatures [116]. The following equations for converting SFC of fats to SFI, with correlation coefficients, R^2, of 0.98–0.99 were offered: Commercial spreads:

$$\mathrm{SFI} = 1.98 + (0.72 \times \mathrm{SFC}) - (0.035 \times \mathrm{temp.})$$

Base stocks:

$$\mathrm{SFI} = 40.94 + (1.22 \times \mathrm{SFC}) + (1.03 \times \mathrm{temp.})$$

Bends (base stocks/liquid oil):

$$\mathrm{SFI} = 0.94 + (0.82 \times \mathrm{SFC}) + (0.02 \times \mathrm{temp.}) + (0.02 \times \mathrm{percent})$$

Profiles of margarine and table spread fats are shown later. Information available from SFI/SFC profiles includes solids contents at removal from most refrigerators (10°C/50°F), in typical kitchens (21.1–26.5°C/70–80°F), and at body temperature (37.8°C/100°F). The Steepness of the profile indicates relative effects of temperature change on fat firmness.

Melting properties of fats in the mouth are important. Generally, consumers can detect a greasy characteristic if more than 3% solids remain in spreads at mouth temperatures, or if more than 5–6% fat solids remain in baked, fried, and snack foods. Thus, doughnuts and bakery products are best eaten while warm to avoid greasy mouthfeel. The fried snack foods industry must choose between these alternatives: (1) snacks fried in oil may impart a cloudy appearance to see-through panels in the package, feel oily when picked up by the fingers, and stain clothing if dropped. (This is expected in long-time favorites such as corn chips, but may not be acceptable in newly introduced snacks.); and (2) Processors who want a "dry feeling" on pickup of the snack may select a fat that is solid at room temperature, but melts rapidly in the mouth to avoid the greasy sensation. Some snacks leave a greasy lining in the mouth if they are eaten at the same time the consumer drinks a cold soda pop. A fat, such as the one marked "Nondairy" in Fig. 34.20, could be used in coffee whiteners.

Another important factor is *plastic range*. Generally, 15–25% solids at ambient temperature are considered acceptable for working a fat-containing product (dough) without it becoming too oily to handle. Doughs containing the “all-purpose shortening” shown in Fig. 34.20 can be machined over a wide range of temperatures in a warm bakery. Stick margarines often are compounded to be softer and easier to spread than butter when taken from the refrigerator, to remain firmer at kitchen temperatures, and to melt completely without greasy mouthfeel when eaten. Soft (*tub*) margarine is ready to spread as taken from the refrigerator.

Thermal Fractionation

The simplest type of thermal fractionation is *winterization* for cosmetic reasons to obtain salad and cooking oils that remain clear when stored in the refrigerator. The oil is chilled in tanks with slow mixing to crystallize the higher melting point waxes, or TAG which are natural or produced by light hydrogenation to delay oxidation of the oil. A filter aid is added to assist filtration. After separation of the liquid (*olein*), the filtering apparatus is heated to melt and recover the fat solids (*stearin*), which can be used in compounding shortenings and other products [117]. The resulting oil is known as RBWD (refined, bleached, winterized, deodorized) and is expected to pass the “cold test” (AOCS Method CC 11–53) by resisting clouding for 5.5 h at 0°C. Shallow unstirred tanks in cool rooms were used for crystallization in earlier days. The rooms became warmer as crystallization progressed, demonstrating it as an exothermic reaction, and oils will readily assume the more compact crystal forms provided they are able to shed the energy as heat.

Thermal fractionation technology is most developed in the palm oil industry, where most oils are fractionated before sales. Solids profiles of stearins that have been fractionated from crude palm oil by chilling to different temperatures are shown in Fig. 34.21 [118]. Thermal fractionation and double fractionation can be useful tools in obtaining fat fractions with the specific desired characteristics (Fig. 34.22) [119].

Three major techniques are used in fractionation. In *dry fractionation,* oil is chilled with slow stirring and often seeded with crystals from an earlier batch. The resulting mixture can be quite viscous and present difficulties in draining oil from the crystals during filtration. In *solvent fractionation,* chilling and crystalization of oil occurs in a solvent, often acetone. The viscosity of the liquid phase is greatly reduced, resulting in easier separation and improved purity of crystals. Solvent fraction must be conducted in explosion-proof facilities. In *detergent fractionation,* a surfactant is added to the chill-crystallized oil–fat mixture and greatly improves drainage [77, 120]. Significant progress in recent years in dry fractionation technology has greatly simplified processing and reduced problems of handling solvents and byproduct streams [121]. A variation of the detergent fractionation principle has been used in dewaxing sunflower seed oil. Advantage is taken of the soap content (1,000–2,000 ppm) and 3–4% water remaining in un-washed alkalineutralized oil. After alkali neutralization, sunflower seed oil is pumped from the primary centrifuge directly to a heat exchanger for crystallization. The oil is held at 5–7°C/41–45°F for about 4 h for crystal growth, then heated by heat exchanger to 12–15°C/54–59°F, centrifuged to remove the wax crystals along with some water and soaps, washed with hot water, and centrifuged [122].

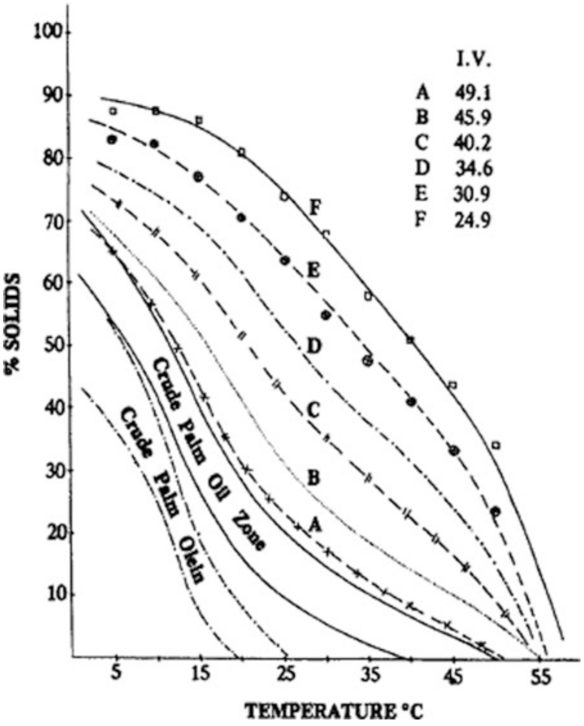


Fig. 34.21 Examples of stearins separated from crude palm oil by chill fractionation. (From Tan and Flingoh [118], With permission)

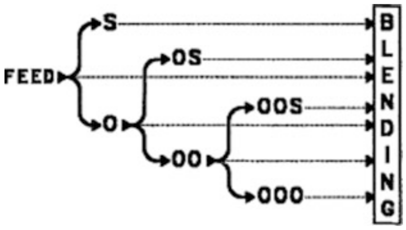


Fig. 34.22 Fat fractions from cascade fractionation. (From Tirtiaux [106], With permission)

Until the mid-1980s, continuous-belt vacuum filters mainly were used for separating oleins from stearins.

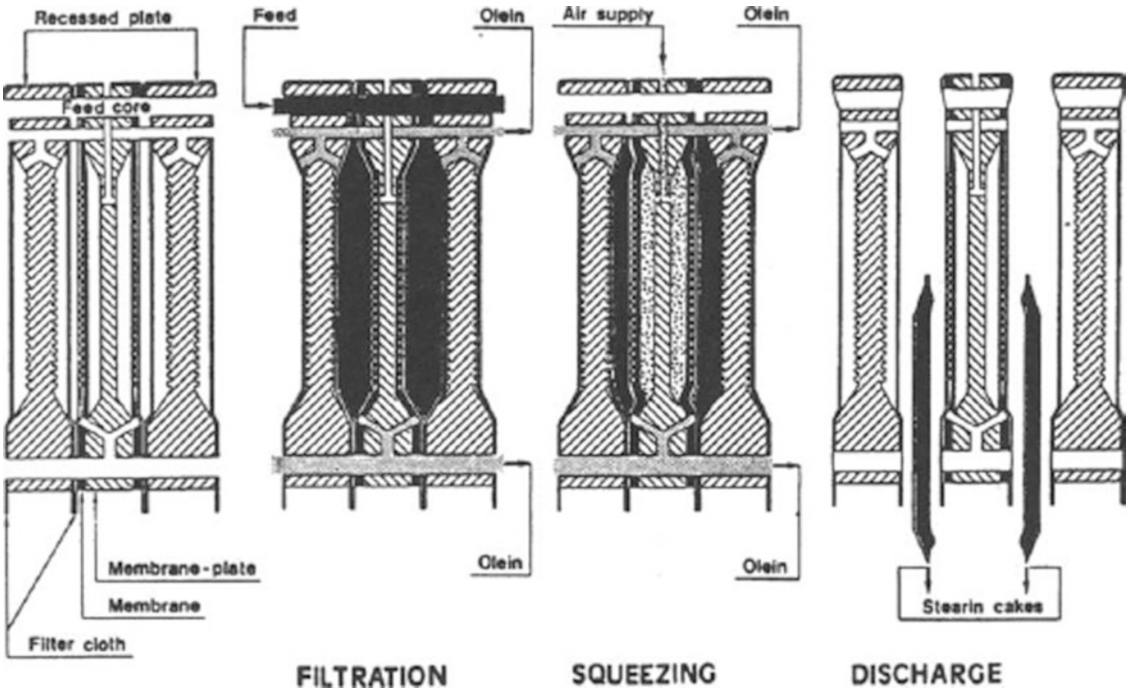


Fig. 34.23 Principle of membrane-type filtration of olein and stearin. (From Tirtiaux [106], With permission)

These are expensive and complex machines, requiring controlled temperatures in various sections and clearing of the belt for the subsequent filtration cycle. Simpler membrane filters have gained in popularity. As shown in Fig. 34.23, the active members resemble a plate and frame filter. Each cavity is lined on both sides with "membrane" filter stock, between which the crystal–oil mixture is pumped. Initially, free oil escapes through the filter stock, until the cavity is filled with crystals. Addition of the oil–crystals slurry is then stopped, and compressed air applied between the membrane and the frame to squeeze oil out from the enclosed crystal mass. When completed, the frames open, allowing the crystal cake to fall into a take-away screw. The frames then reclose, and the cycle is repeated. Stearins produced by membrane filters contain significantly less free oil than those from vacuum belt filters [119].

Oil is an effective solvent for higher temperature melting TAG, and results of thermal fractionation of common fats can be surprising. An example is milk fat (butterfat), which on extensive study has yielded 123 fractions melting as high at 54°C/129°F. Many tropical countries allow inclusion of up to 5% high melting hydrogenated fats in chocolate to raise the melting point and prevent blooming (loss of sheen, paling of color, and development of coarse crystals) resulting from repeated melting and resolidification. This also is a problem during the summer months in temperate countries. However, the latter countries often prohibit inclusion of fats other than cocoa butter in chocolate. The problem has been reduced in milk chocolate by adding high melting fractionated milk fat. Milk fat also has been used in higher-priced shortbread cookies, laminated pastries, and other bakery products where it provides a shortening effect, but doesn't melt and appear greasy.

Hydrogenation

Hydrogenation is the catalytically assisted addition of hydrogen to carbon–carbon double bonds. Its main uses are to increase fryer life of oils and shelf life of table oils and bakery products and to create solids for making shortenings, margarines and spreads, and various confectionery and specialty products. Many process reviews have been published [124–131].

Food technologists prefer to minimize the C18:3 content of oils to improve fryer and bottled oil shelf lives by reducing oxidation tendencies (Table 34.4). This has been partially accomplished by applying selective breeding and biotechnology to oilseeds, but the GMO products still are in the introductory phase. Because small amounts of linoleic (C18:2) and linolenic (C18:3) fatty acids are dietary essential, it is not desirable to eliminate them completely. For many applications, soybean and canola oils are partially stabilized against oxidation by (light) *brush hydrogenation*.

Selectivity, the tendency of the catalyst or process to remove one type of bond in favor of another, has several meanings in hydrogenation catalysts. *Preferential* or

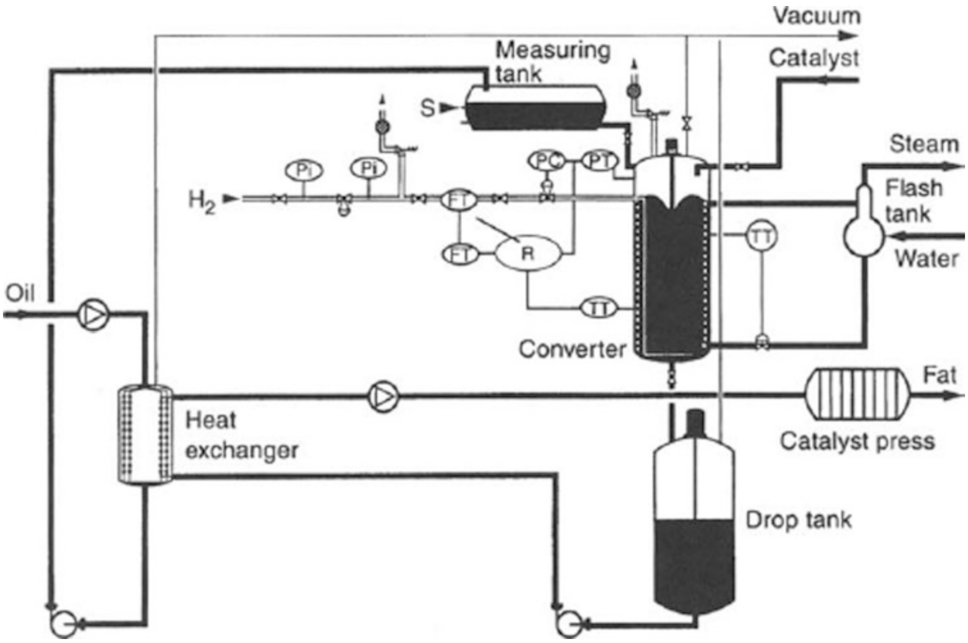


Fig. 34.24 Hydrogenation line. (From Hastert [130], With permission)

saturate selectivity indicates a focus on saturating a specific bond, for example, the C18:3 bond of linolenic acid to form C18:2 linoleic acid. Preferential selectivity is estimated from the following kinetic relationships:

$$C18:3 \xrightarrow{k_1} C18:2 \xrightarrow{k_2} C18:1 \xrightarrow{k_3} C18:0$$

with linolenic selectivity ratio (SR) defined as SR $1 = k_1/k_2$ and linoleic selectivity as SR $2 = k_2/k_3$. Linolenic acid selectivity for nickel catalysts varies from 2 to 3 and is as high as 6 for copper-based catalysts. Better control of linoleic selectivity, with SR 1 varying from 3 to 4 to 60, is achievable by the choice of catalysts, catalyst poisons, and reaction conditions. This gives the operator greater flexibility in establishing melting behavior and stability of the hydrogenated product [125].

Trans isomer selectivity is less directed and favors formation of *trans* bonds. Catalysts vary in selectivity, with performance greatly affected by catalyst dosage, effectiveness of mixing, temperature, and hydrogen pressure in the reactor. Bonds become conjugated during hydrogenation, but not all atoms fall back neatly into their former positions after the process. The new types of TAG produced increase diversity, which is desirable in controlling later crystallization processes. The production of *trans* (partially saturated) bonds is intentional and can be partially controlled [129].

Hydrogenation is conducted in hardening plants (Fig. 34.24). The *converter* is a pressurized reaction tank, equipped with a highspeed mixer and assisting baffles, means for adding and removing the oil, a gas distributor, a means to add catalyst, and heating and cooling coils. The hardening plant additionally has means for premeasuring and heating the oil, a drop tank, heat exchangers, and a catalyst filter. Purities of the oil and hydrogen affect the life of the catalyst, typically a thin film of nickel on an inert carrier [130].

In brush hydrogenation of soybean oil selective catalysts: (1) reduce IV by 15–25 to ~115 units; (2) produce ~15% *trans* isomers; (3) reduce C18:3 content to 3% maximum; and (4) increase C18:0 content by ~1% [131].

In preparing margarine bases, selective catalysts reduce IV to ~70 and produce about 50% *trans* isomers. Selective or nonselective catalysts used in preparing shortening bases reduce IV to ~75 and produce about 35% *trans* isomers. A high-activity catalyst is used in producing stearin flakes; the IV is reduced to ~5–10 with hardly any *trans* isomers remaining. The flakes are scraped off the surface of a chilled roll, or beadlets can be produced using a shot tower with chilled air.

Maintaining high levels of polyunsaturates is desirable in producing coating fats. A sulfided nickel catalyst is used to reduce IV to ~70 and produce about 65% *trans* isomers with production of saturates (C18:0) minimized to 2–4% increase [131].

Factors affecting the hydrogenation reaction include reactor design, purity of feedstock and hydrogen, operating conditions, and choice and efficiency of the catalyst. Operating conditions include hydrogen pressure, reaction temperature, catalyst dosage, and agitation. Because *trans* formation is indicative of incomplete saturation, conditions that favor keeping the catalyst covered with hydrogen favor saturate selectivity. Progress in the hydrogenation process typically is monitored online by refractive index, calibrated to IV for each process and product [131].

Production of temperature-profiled fat products, such as margarines/spreads and shortenings, often includes preparation of four to six base stocks hydrogenated to different IVs (Fig. 34.25), which are blended with oil and hardstocks to obtain the desired temperature-fat solids profile [132].

An example of broadening the plastic range (15–25% fat solids) of two base stocks by addition of hardstocks is shown in Fig. 34.26. By itself, the 80 IV base stock is in the 15–25% solids (machinable) range only between the temperatures of 10 and 20°C. But, by addition of 8% (5 IV) hardstock, the range was broadened to ~30°C. Addition of 12% hardstock to the 85 IV base stock broadened its working range from ~10–12 to 38°C, which is more typical of bakery operating temperatures [124].

Interesterification

Interesterification (INES) is the exchange of acyl radicals between an ester and an acid (*acidolysis*), an ester and an alcohol (*alcoholysis*), or an ester and an ester (*transesterification*) and can be random, directed, or enzymatic. The process has been called *intraesterification* if an exchange of positions occurs within the same molecule, and *randomization* if exchange occurs between molecules [44, 48]. The principles can be used to position fatty acids on molecules with hydroxyl sites to produce monoglycerides (emulsifiers), fatty acid methyl esters (FAME) for analytical purposes, liquid fuels such as methyl soyate (biodiesel), specialty fats such as medium chain triglycerides (MCT), and sugar-ester noncaloric fat substitutes such as Olean™.

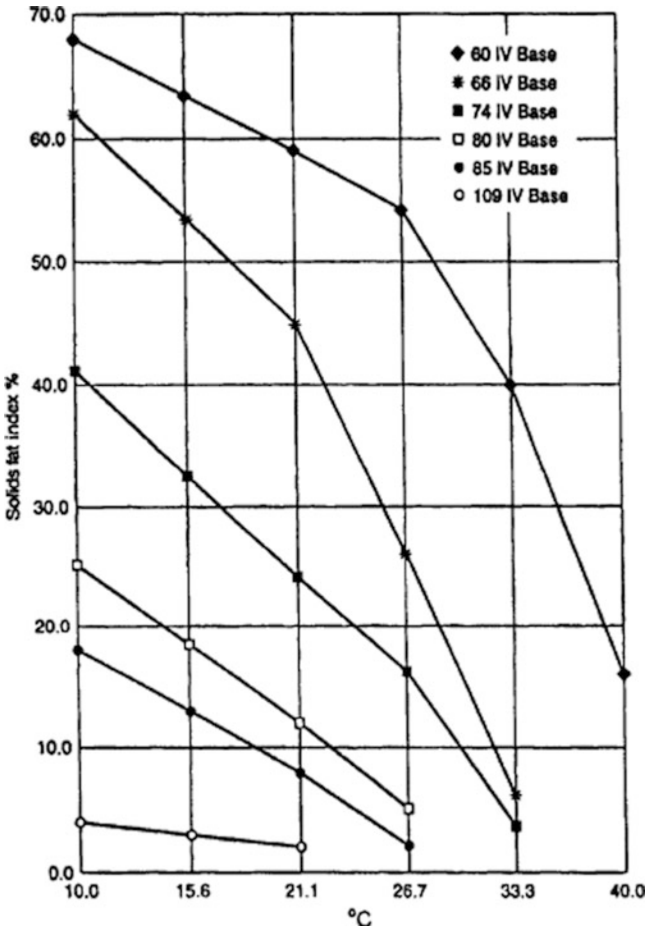
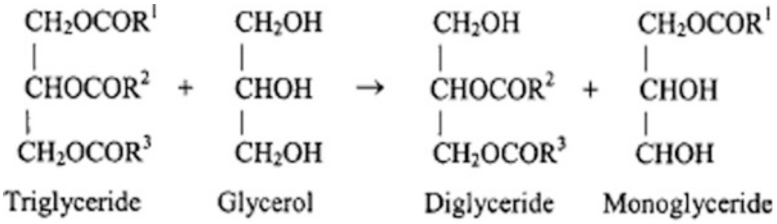


Fig. 34.25 Solid Fat Index (SFI) profiles for six hydrogenated soybean oil (H-SBO) base stocks. (From O'Brien [132], With permission)



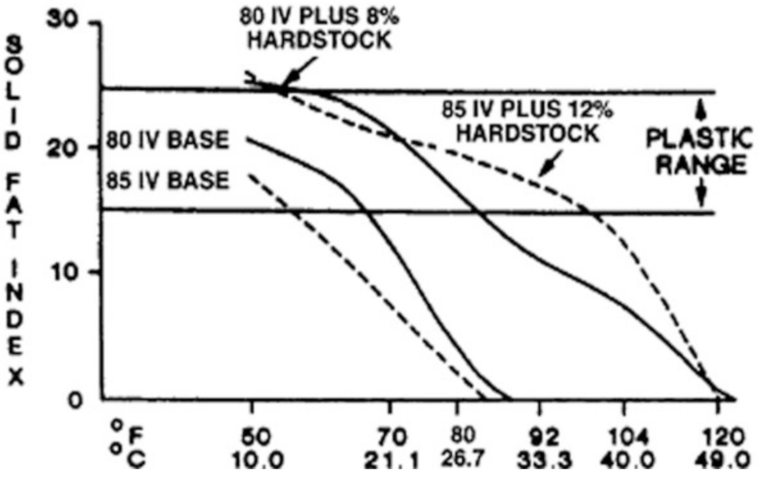


Fig. 34.26 Effects of adding hardstock (<5 IV) to broaden plastic range working temperatures of two base stocks. (From Hastert [124], With permission)

Table 34.11 Triglyceride classes of native and interesterified oils and fats and of noninteresterified and interesterified blends (1 : 1) of fully hydrogenated soybean oil with vegetable oils[a]

		Single oils					1: 1 Blends with hardened soybean oil			
		S_3	S_2U	SU_2	U_3	M.P.	S_3	S_2U	SU_2	U_3
		(%)	(%)	(%)	(%)	(°C)[b]	(%)	(%)	(%)	(%)
Palm oil	n	6	50	38	6	39.8	57	13	20	10
	r	13	39	37	11	47.0	41	43	14	2
Soybean oil	n	0	6	38	56	−7.0	50	2	17	31
	r	1	8	36	55	5.5	13	47	32	9
Cottonseed oil	n	<1	18	51	30	10.5	51	9	24	16
	r	3	18	44	35	34.0	25	34	31	10
Sunflower oil	n	0	1	24	75	–	51	0.3	11	38
	r	<0.2	4	27	69	–	11	47	34	8
Peanut oil	n	0	11	40	49	–	53	3	15	29
	r	1	10	38	51	–	16	47	31	6
Rapeseed oil	n	0	1	16	83	–	51	1	9	39
	r	0	1	17	82	–	10	44	37	9
Coconut oil	n	81	12	7	0	26.0	57	13	20	10
	r	74	24	2	<0.1	28.2	41	43	14	2
Palm kernel oil	n	76	15	9	0	–				
	r	53	37	9	0.7	–				
Cocoa butter	n	2	85	12	1	34.4				
	r	24	43	27	5	52.5				
Lard	n	8	30	50	12	43.0				
	r	10	32	40	18	43.0				
Beef tallow	n	22	60	18	0	46.2				
	r	13	38	17	12	44.6				

[a] From: Bookish [48], With permission
[b] From: Sonntag [136], With permission
n native; *r* randomized

Research publications exist from at least as early as 1852, and United States patents from 1939 [133].

The basic acidolysis reaction is:

$$R^1COOR^2 + R^3COOH \rightarrow R^3COOR^2 + R^1COOH$$

The alcoholysis reaction is:

$$RCOOR^1 + R^2OH \rightarrow RCOOR^2 + R^1OH$$

with methyl alcohol used for preparing FAME when determining the fatty acids components of TAG, and in making much of the biodiesel. Ethanol and other alcohols, up to four carbons in length (butanol), also have been used [44].

Glycerolysis has been used to prepare mono- and diglyceride emulsifiers by reacting TAG with an excess of glycerol [44].

Standard “mono- and diglyceride” products contain 40–50% monoglycerides, whose content can be raised to 53% without distillation. Distilled monoglycerides contain a minimum of 90% α-monoglycerides [44].

In the *sucrosolysis* preparation of Olean™, the generic name for Proctor & Gamble Company’s olestra, six to eight fatty acids, selected for functionality purposes, are positioned on sucrose. Although many noncaloric fats are limited or nondigestible, they are made from traditional vegetable oils. Akoh has summarized uses of these and similar products [134].

Except for enzyme-directed processes to place certain fatty acids in specific positions on TAG, such as production of coating fats, cocoa butter substitutes, or reduced-calorie fats [135], the majority of esterification processes are randomizations.

As mentioned earlier, nature arranges the fatty acids in plant oils to achieve the lowest melting point. The abbreviations S and U, for saturated and unsaturated fatty acids, respectively, also are used in indicating their prevalence on TAG: thus S_3, S_2U, SU_2, and U_3. Distributions of these combinations in nature and in randomized products are shown in Table 34.11 [48, 136]. The literature indicates that stability against oxidation often is reduced by withdrawing the unsaturated fatty acids from the *sn*-2 position [44]. Randomization always raises the melting point, more so for vegetable oils than for animal fats, with an increase of 8.5°C for soybean oil. S_3 TAG hardly exists in natural vegetable oils, with only low concentrations occurring in animal fats. However, some are formed during randomization.

Directed randomization is a specific technique that precipitates S_3 and S_2U TAG as formed, depending on operating temperatures chosen. Removal of saturated fatty acids from cottonseed oil has been demonstrated recently

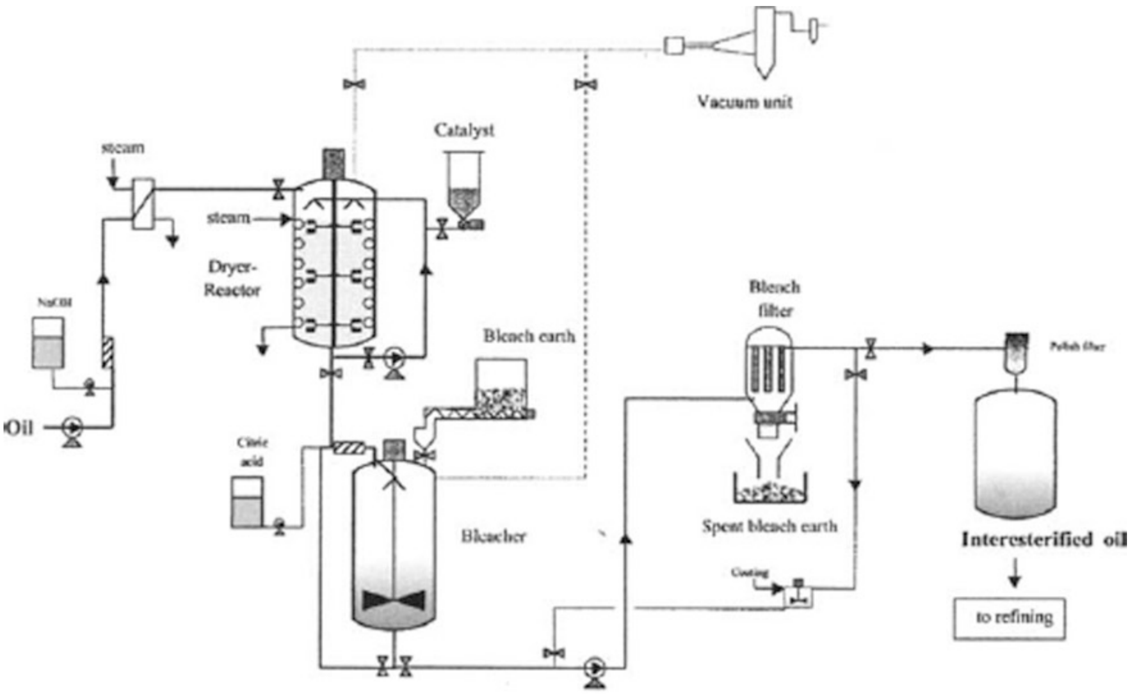


Fig. 34.27 Semi-continuous (chemical) interesterification line with postbleaching. (Courtesy of Desmet Ballestra Oils and Fats, Brussels, Belgium)

[137]. Directed randomization can be used to modify either the olein or stearin fractions.

A flow sheet for a semi-continuous interesterification line with postbleaching ability is shown in Fig. 34.27. The oil must be degummed, well-refined (<0.05% FFA), and free from peroxides and moisture that causes production of soaps. Palm oil stearin and completely saturated (<5 IV) C18 hydrogenated fats generally are *trans* free and may be used as part of the feed. Reactions are run under vacuum, and oil is best stored under nitrogen between processes [44].

A variety of catalysts is available, including alkali metals, alkoxides, alkali hydroxides, sodium hydroxide and glycerol mixtures, metal soaps, and metal hydrides [44]. Alkaline hydroxides (KOH or NaOH) in combination with glycerol were used in earlier days because of their low cost [138]. Currently, sodium methylate and sodium ethylate are popular because of their efficiency. Rozendaal has proposed that the actual interesterification catalyst is a sodium derivative of a diacylglycerol, activated during the process [139].

The oil is loaded into the reactor, shown with both an agitator and a pumped circulation-spray loop, and heated under vacuum (110–130°C) to reduce water and peroxide contents. Next, the oil is cooled to 70–90°C and the catalyst is added as dry powder at 0.05–0.15% or suspended in dry oil. Randomization requires about 30 min, with an additional 15–30 min allowed for completion. After the reaction is complete, the batch is transferred to a postbleacher where the process is arrested by inactivating the catalyst by addition of water or an (phosphoric or citric) acid solution. Bleaching earth added to absorb the inactivated catalyst and soaps removed by filtration and the oil sent to blending or deodorization. Losses from the formation of FFA and FAME are ~10 times the catalyst weight, with an additional 0.5–1.0% lost in the formation of mono- and diglycerides, for a total of 1.5–2.0%. Continuous processes also are available. Interesterification progress can be monitored by online ultraviolet (UV) spectrophotometry and completion of the process by melting point determination and other measurements [44].

The main objective in interesterification is to produce solids free from *trans* fatty acids for later use in compounded fat products. The final fat, or a series of base stocks, can be made for future blending. Randomization can improve the functionality of a fat, as shown for lard in.

Figure 34.28. Natural lard is within the 15–25% solids plastic range between 9 and 24°C; by randomizing, the temperature range is moved to 0–17°C; both conditions greatly restrict its use. But, by directed interesterification, the machinable range is moved to the more acceptable ranges of 0–32°C. In contrast, natural cocoa butter (Fig. 34.29) is almost brittle at 26°C, melts rapidly by taking energy from the tongue, causes a pleasant cooling sensation in the mouth, and is completely liquid with good palate "cleanup" at 35°C. Randomizing destroys these properties, changing the fat into a softer mass, which does not completely melt below 56°C [48].

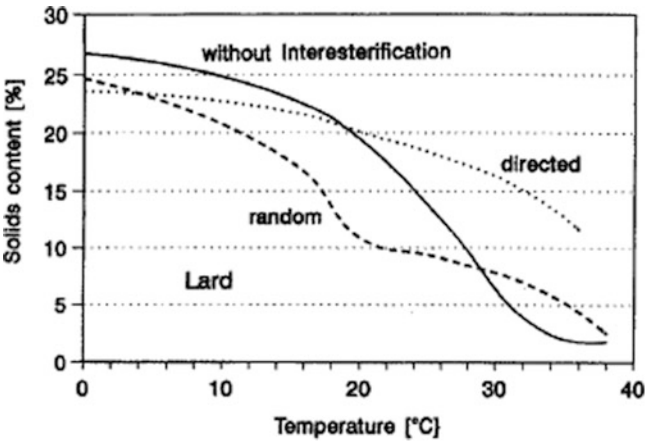


Fig. 34.28 Solids content of natural, random esterified, and directed interesterified lards. (From Bockish [48], With permission)

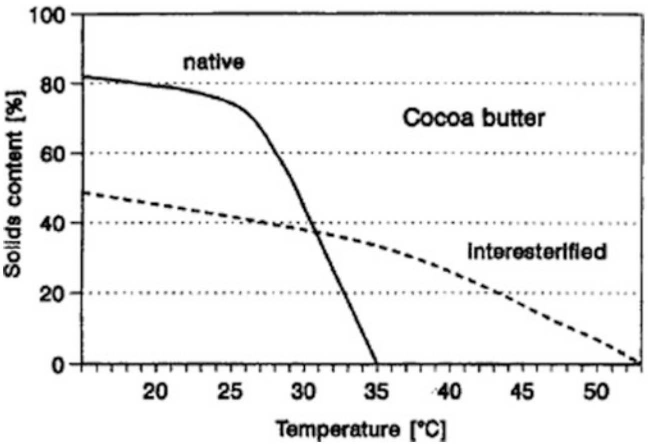


Fig. 34.29 Solids content of natural and randomized (interesterified) cocoa butter.(From Bockish [48], With permission)

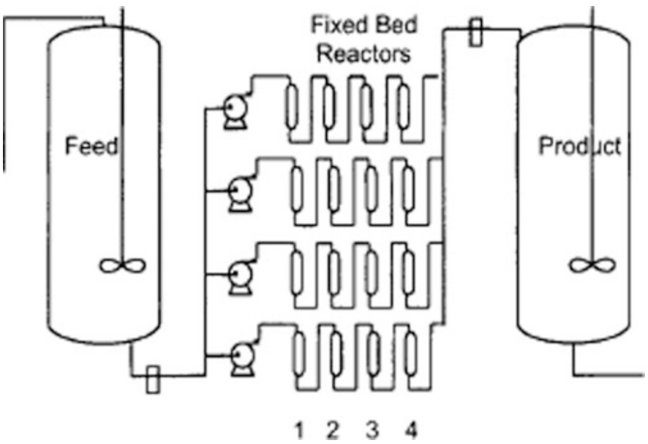


Fig. 34.30 Schematic drawing of multiple enzyme reactor system for enzymatic interesterification of *trans*-free margarine and shortening oils. (Courtesy of Novozymes A.S, Bagsvaerd, Denmark)

Interesterification is being pursued on the basis that *trans* production during hydrogenation increases the melting points of TAG and fatty acids, increases blood cholesterol levels, and increases atherosclerosis (plaque deposits in arteries) and coronary heart disease (CHD) incidence. However, considerable documentation also shows that the presence of saturated fatty acids (primarily palmitic, and to a lesser degree myristic and stearic) at the TAG *sn*-2 position strongly favors atherogenesis. Pork lard contains 70% of its palmitic fatty acid in the *sn*-2 position, whereas only 17% of the palmitic acid in beef tallow is in this position. Lard is considered more atherogenic to laboratory animals than tallow; randomization reduces this propety. However, atherogenicity of tallow is increased by randomization. Levels of saturated fatty acids at the *sn*-2 position are very low in natural vegetable oils, but are increased by randomization, with accompanying atherogenesis in feeding trials [140, 141]. Skepticism exists that switching to interesterification as a means of generating fat solids will reduce health concerns about hypercholesterolemic and atherogenic effects of fats.

Enzyme interesterification is rapidly becoming popular and has the advantage of selection of TAG positions at which fatty acids are interchanged. When 1,3 lipases are used, the current fatty acid at 2-position remains in place, avoiding transposition with saturated fatty acids. A simple way to conduct enzyme interesterification is by contracting for a fixed-bed reactor service. The supplier assists in installing the process and takes responsibility for providing portable reactor beds charged with specified 1,3-lipases to treat established quantities of oil. A schematic of a multiple enzyme reactor system is shown in Fig. 34.30. The freshest reactor is connected to the tank receiving the product and is preceded by a sequence of increasingly used reactors back to the most spent reactor connected to the feed tank. As spent reactors are retired, they are moved forward from positions “4” toward “1.” In this fashion, full use is made of remaining enzyme activity as feed flows through the system. The partially exhausted reactors filter the least processed feed and intercept competitive inhibitors that otherwise could reduce performance of fresh reactors. The reactors are shipped between supplier and user by truck. Large facilities can repack reactors in house [142].

Deodorization/Physical Refining

Deodorization is the final step in the production of RBD oils and modified fats. The process is called *deodorization* if most of the FFA is removed by alkali neutralization as with row crop (“soft”) oils, and *physical refining* if, by omitting alkali neutralization, the FFA are left in the oil for removal by steam distillation as with palm oil. Thorough degumming and bleaching must first be done before physical refining of all oils.

The volatile peroxides, other oxidation decomposition products, and odiferous compounds form reduced-boiling point azeotropes with water in the steam, at high temperatures, 250–260°C/482–500°F, and very low absolute pressures (~3 mbar). This is above the smoke point of soybean oil, but below the flash point, and oxygen must be excluded. Considerable heat bleaching of yellow–red carotenoids also occurs at this temperature. Typically: the deodorization process requires 20–40 min after come-up time, uses 0.5–2.0% sparged steam (the higher level if tocopherols are recovered), operates at between 2 and 4 mbar, and produces a product with about 0.03–0.05% FFA [143].

Historically, the standard deodorizer held 60,000 lb of oil (one railroad tank car). Except for refineries making only a few kinds of oil, as for export, building of continuous deodorizers slowed with the advent of just-in-time (JIT) delivery, supplier self-certification, and customers buying on the basis of their projected production schedules. This has led to development of improved batch-continuous systems, which are designed to handle many batches of different oil blends per day, with minimum cross-contamination and delays for process.

For many years, deodorizers (operating at above 270°C/520°F) were heated by several types of mineral oil-like thermal fluids which, in turn, could be heated in direct-fire furnaces at 315°C/600°F in plumbing and coils at 3.2 bar (46 psig). During the 1970s, some consumer health problems in Europe were ascribed to leakage of thermal fluids into oils during deodorization. As a result, the European market required that the local and imported oils industry shift to using high-pressure steam generators, operating at ~80 bar (1,150 psig) to provide a temperature of 295°C/560°F. Soon, other import countries also started adopting European standards. This essentially meant that heating coils and jackets of deodorizers, owned by suppliers wanting to sell oil in Europe, had to be rebuilt or new deodorizers and high-pressure steam generators purchased [144].

Deodorizers are built in many vertical and horizontal designs [143–145]. They typically are located outdoors and look like multistory plumbed silos or petroleum refinery reactors. A drawing of a modern deodorizer is shown in Fig. 34.31. The components have been gathered into one shell to take advantage of heat recovery opportunities, and the temperature gradient within the vessel gradually decreases from top to bottom. As with all manufacturers, improved designs are continuously evolving. The unit is operated at lower temperatures, 220–235°C/428–455°F, and a deeper vacuum (2 mbar) to minimize *trans* and polymer formation and loss of valuable minor components such as tocopherols. The supporting high-pressure steam boiler and vacuum system (typically a four-phase steam-jet ejector with barometric condensers, mechanical liquid ring vacuum pumps, or dry "ice" condensing system) is not shown. In the deodorization process, a well-prepared pre-warmed RB oil is passed through a deaerator to remove dissolved air, then heated in an exchanger by oil exiting from the deodorization stage. It is additionally heated by steam coils and passes into the deodorization tray, where it is deodorized by sparged steam. Next it passes through a heat recovery exchanger, heating the incoming oil. After partial cooling, citric acid (20–50 ppm) in solution is added to sequester iron or copper that may be picked up later by the oil. The oil is still hot and under vacuum, and the moisture flashes off. Then, the oil is cooled and passes through a polishing filter on its way to temporary storage under nitrogen before shipment.

The liquid used to *scrub* the vapors is previously condensed deodorizer distillate that is chilled, recycled, and drawn off as necessary to maintain a constant level in the scrubber. The condensate may entrap vapors of pesticides if they get this far in the refinery. Deodorizer distillate is rich in tocotrienols or tocopherols, some of which have vitamin E activity. During the time that consumption of vitamin E supplements in the United States was growing, and strong markets for natural antioxidants existed, deodorizer distillate was very much in demand and the domestic supply from refineries was committed by contracts with vitamin E producers. As much as 60% of the tocopherols in soybean oil could be extracted at the deodorizer, but in doing so, the refineries no longer are available to stabilize the shelf life of the oil or provide its full natural nutritional benefits. In recent years, tocopherols and tocotrienols from palm oil processing have eased supplies.

At this point in processing, the peroxide value in the oil should barely be detectable (<1 ppm), but will soon start increasing again. Thus, margarine and other profiled-temperature fats are blended before shipment. A rule of thumb in the deodorization department is: "Don't deodorize oil until it is sold and ready to be shipped."

After cooling to appropriate temperatures, processors may add additional oil-soluble ingredients, antioxidants, vitamins, colors, and others that customers might have difficulty dispersing in the oil/fat at their processing sites. Unless the fat needs to be texturized into a soft-plastic form by the addition of nitrogen and chilling to a shortening for specific handling requirements, it is shipped and kept at about 10°C above the melting point. Nitrogen purging of oil during pumping, and storage under a *nitrogen blanket,* are common if held for more than several days before use by the processor. Contact with copper- and iron-bearing materials and contamination with water should be avoided. Typically, one to two percent *trans* fats are produced during deodorization, with lesser quantities formed at lower temperatures. Ultra-high vacuum steam distillation, operating at microbars

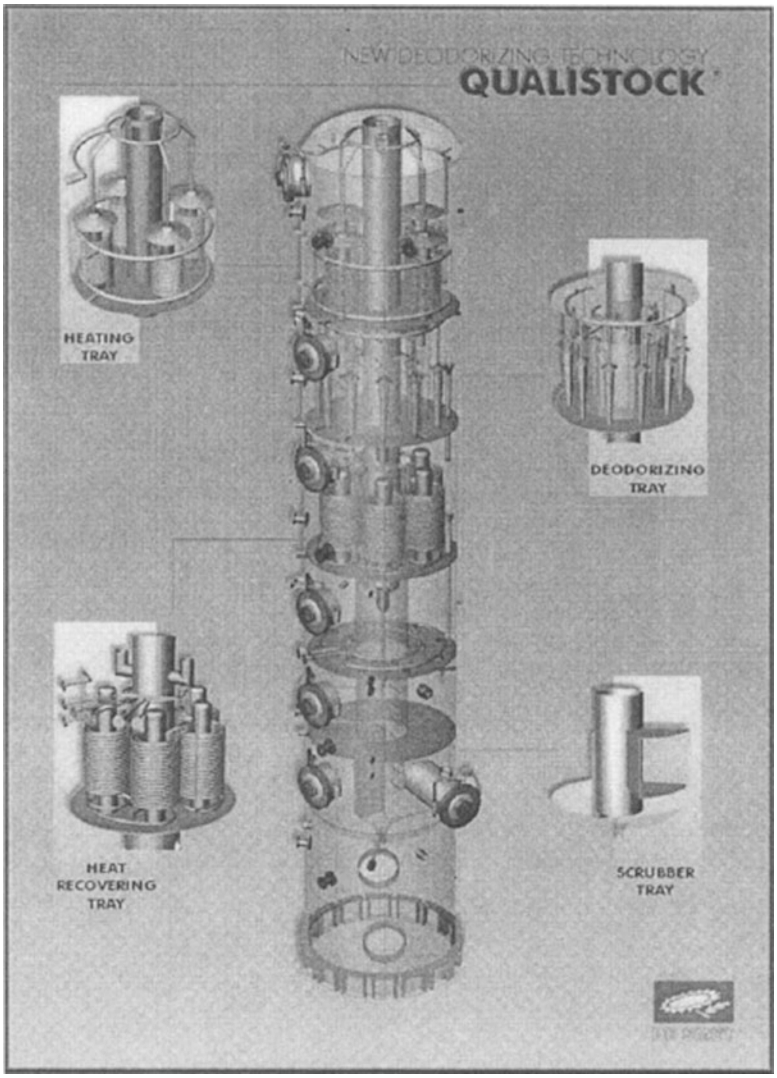


Fig 34.31 Cut-away drawing of DeSmet Qualistock® Continuous Deodorizer. (Courtesy of Desmet Ballestra Oils and Fats, Brussels, Belgium)

rather than millibars, has successfully deodorized oils and fats in laboratory and pilot plant scale and is used in the production of higher-priced products. However, fats and oils scientists have reported that the flavor of this oil is not as acceptable as from traditional deodorization/physical refining. Several steam distillation methods for the removal of cholesterol from animal fats were developed and patented after the fast foods industry switched from tallow to vegetable oils for frying French fries in the 1980s. Currently, the ability of tallow to regain this former market seems doubtful.

Notes on Major Row Crop Oilseeds

Soybean

Soybean (*Glycine max* L.) was domesticated in north China, probably during 1700–1100 BC in the Shang Dynasty or earlier [146]. Samples found their way to Europe during the 1700s, where it was grown as a curiosity in botanical gardens [147]. Documentation exists for soybean brought, grown, and made into soy sauce near Savannah, Georgia in 1765, by Samuel Bowen, an English seaman who had sailed to Canton, China earlier as an employee of the East India Company. Benjamin Franklin sent seed from London to a botanist in Philadelphia in 1770 [148]. A sample, from a salvaged Japanese junk, was brought from San Francisco to Alton, Illinois, in 1851, where it was grown and also distributed to botanists in other states.149 Additional samples were brought back by Commodore Matthew Perry's expedition to Japan in 1854, sent by missionaries in the Far East, and distributed in the United States before establishment of the Department of Agriculture in 1864 [149]. However, the plant mainly was grown for feeding cattle and as green manure to enrich soils.

The Treaty of Portsmouth (New Hampshire) to settle the 1904–1905 Russo–Japanese War gave Japan control of the Liaotung Peninsula in Manchuria. This area was leased previously from China by Russia during construction of the Trans-Siberian (Moscow to Vladivostok) Railroad for establishing all-weather naval and trading ports in Port Arthur (now Lüshun) and Dairen (Lütita, Talien, Dalian) [150]. The withdrawal of Japanese occupation troops from Manchuria resulted in a local surplus of the soybean crop, which had been expanded to feed the soldiers. Manchurian soybean was shipped to England in 1907 [151], and to Germany, other Northern European countries, and the US Pacific coast, by 1910 [152]. Dairen, Manchuria became the international soybean processing and export center for the early part of the 1900s. An English hydraulic press mill had been sold to China in 1868 [153], but Manchurian and Japanese soybean extraction continued mainly by crushing seed with stone edge runners and pressing with wedge presses, which produced press cake that could not survive shipment to Europe [152, 154, 155].

Europe wanted the soybean for oil and the meal for animal feed, but was dissatisfied with leaving 5–6% to six percent oil in the cake, characteristic of box presses and expellers (screw presses). England and Germany initiated work on batch solvent extraction processes, but World War I interrupted progress. The Bollman basket countercurrent extractor was patented in Germany in 1919, followed by the Hildebrandt U-tube extractor in 1934 [152].

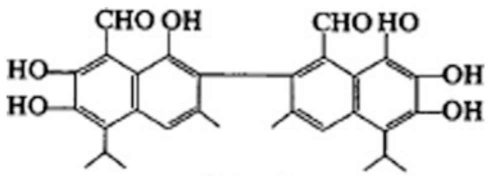
The first domestic crushing of soybean occurred in Seattle, Washington, in 1910, using seed imported from Manchuria [152]. Domestically produced soybean was crushed at cottonseed oil mills in the South and linseed (flaxseed) oil mills in the Midwest during the 1910–1920 era [152, 156]. The first commercially successful domestic soybean solvent continuous extraction plants were installed by Archer Daniels Midland Company (ADM) and the Glidden Company in Chicago in 1934, both using 100 tpd Hildebrandt U-tube extractors and commercial hexane as solvent. They also were equipped to produce phosphatides [153].

United States production of soybean, for domestic processing and exporting to Europe, grew slowly until acceleration by World War II, which devastated the German oilseed crushing industry [157]. This left the United States in the position of the world's largest soybean producer and processor. Soybean utilization increased during the reconstruction of Europe and Japan, by uses in famine abatement programs throughout the world, to meet growing world population food needs, and as feed to support world growth of the broiler industry which started about 1960. Eventually, land-rich countries in South America also became major soybean producers.

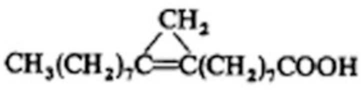
Cottonseed

Cotton (*Gossypium arboreum* and *G. herbaceum*) was grown in the Indus River Valley (modern Pakistan) as early as 3000 BC. "New World" cottons were grown in Peru (*G. hirsutum*) about 2500 BC, and Sea Island cottons (*G. barbadense*) in the Caribbean Islands. Most of the world's cotton today is *G. hirsutum*, with a fuzz-covered seed that requires removal of the fibers by *delinting* before dehulling and subsequent processing. Pima and Egyptian cottons (*G. Barbadense* type) have fuzz-free (*naked*) seed and can be processed directly. Cottonseed is among the first examples of a reclaimed byproduct in our industrial age. Invention of the cotton gin by Eli Whitney in 1793 led to increased domestic production of cotton and also to seed disposal problems. The state of Mississippi passed the first antipollution law in the nation in 1857, prohibiting throwing cottonseed into rivers and requiring its orderly disposal from gins located less than one-half mile from a town.

Cottonseed contains the polyphenolic yellow-red pigment gossypol, 1,1′,6,6′,7,7′-hexahydroxy-5,5′-diisopropyl-3,3′- dimethyl (2,2′-binapthalene)-8,8′-dicarboxaldehyde, in discrete bodies called *gossypol glands*. Processing challenges include removal of gossypol, or its 15 or more derivatives, from the oil to reduce color, and deactivate its toxic effects (through binding to the protein with moist heat) to enable feeding the meal to monogastric animals. Cottonseed oil also contains the cyclopropenoic acids, C18:CE malvalic acid (8,9-methylene-8-heptadecenoic acid) and C19:CE sterculic acid (9,10-methylene-9-octadecenoic acid), which form the pink color in the Halphen Test used to detect its adulteration of higher-priced olive oil [158]. Concerns about adulteration of olive oil appeared again when fractionated palm oil was introduced to Europe.

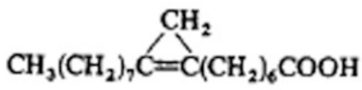


Gossypol



$$CH_3(CH_2)_7\overset{CH_2}{C{=}C}(CH_2)_7COOH$$

Sterculic acid



$$CH_3(CH_2)_7\overset{CH_2}{C{=}C}(CH_2)_6COOH$$

Malvalic acid

Cottonseed oil has long sold at a slight premium over soybean oil because of greater stability to oxidation and desirable flavor in fried snack foods such as potato chips. However, gossypol content and lower protein quality put the meal at a price disadvantage. Feeding whole cottonseed to dairy cattle, whose rumen microorganisms can detoxify limited amounts of gossypol, now utilizes over 70% of the supply and may portend eventual abandonment of oil extraction.

Rapeseed/Canola

Rapeseed/Canola belongs to the turnip rutabaga, cabbage, Brussels sprouts, and mustard family of crops that can be grown at low temperatures and moderate humidity. Three species have been grown as oilseeds: *Brassica napus*, known in Europe as rape, oil rape, Swede rape, and Argentine rape; *B. campestris*, known as rapeseed, oil turnip, turnip rape, and Polish rape; and *B. juncea*, known as leaf mustard, brown mustard, Oriental mustard, and Indian mustard. *B. campestris* was grown in India as early as 2000–1500 BC.

This crop is the classic example of an effective crop improvement program mobilized by a country in response to impending loss of international markets. When it was reported in the early 1960s that erucic acid (12-docosenoic, C22:1 *n*–9) in rapeseed oil could cause heart damage and other diseases, the Canadian government established a crash development program that led to the release of the first low-erucic acid rapeseed (LEAR) in 1969. Continuing efforts to develop varieties with even lower erucic acid content led to later release of double zero (low-erucic, low-glucosinolate) varieties, which were named canola in 1980. Because of reduced glucosinolate levels, considerably more of the meal could be used in animal feed than earlier. Canola oil, which contains less than 2% erucic acid compared with 20–40% in earlier rapeseeds, was granted GRAS status by the US FDA in 1985 [159].

However, the introduction of canola left unmet needs for erucic acid in industrial markets. High-erucic acid rapeseed then was imported from Northern Europe for extraction, followed by efforts to increase erucic acid contents in domestic industrial rapeseed as well as development of crambe (*Crambe abyssinica*) specifically for its erucic acid content. At the current state of development, equipment corrosion and poisoning of hydrogenation catalysts by sulfur in canola oil still are problems. Also, oil from the new canola varieties favors formation of β crystals, in contrast to β' formation in the earlier rapeseed varieties.

High *oleic acid canola* varieties were introduced recently to directly compete in applications formerly using olive oil.

Sunflower Seed

Sunflower (*Helianthus annus* var. *Marco-carpus*) is a New World crop, known to have been grown in Arizona–New Mexico in 3000 BC and in the Mississippi–Missouri Basin at least since 900 BC. A midsummer weed relative, with small multihead flowers, is common in the central United States. The crop was taken to Europe by early explorers and developed as a source of edible oil in Russia. The gray and white striped seed is known as *confectionery-type*. Its kernel is loose within the hull, and the seed usually is shelled for eating. No major antinutrition factors are known to exist in the raw seed. *Oil-type* varieties predominantly are black-hulled, with seeds about one third the size of confectionery types. The seed adheres tightly to the hull and provides a challenge in dehulling. It has been repeatedly shown that, although the setting of seed by sunflower plants is not as sensitive to day length as for soybean, the oleic acid content is inversely related to the temperature of seed maturation. Early users desiring polyunsaturated oil with high linoleic acid content specified northern-grown sunflower seed oil, whereas those desiring extended oil stability, as in the production of fried foods, have preferred oil from seeds matured during the summer in the southern states. Sunflower varieties with high-oleic acid content (85–92%) have been developed with the intention of using the oil as a feedstock in chemicals manufacturing processes [122]. Recently, the sunflower seed industry has developed mid-oleic acid oil, containing ~65% oleic acid. Also, high-oleic confectionery-type sunflower seed has been developed, enabling extension of the shelf life of roasted seed.

Peanut

Peanut, groundnut (*Arachis hypogae*), also is a New World crop and was grown in the Upper Plata River Basin of Bolivia in 2000 BC. It was taken to Europe by early explorers and was returned to the southeastern United States from Africa by slave traders. Broad cultivation did not occur domestically until the early 1920s, when the southeastern United States was looking for a crop substitute for cotton, which was severely ravaged by the boll weevil [160]. Most domestically grown peanut is consumed as food, with over one half of the crop produced used in making peanut butter. Peanut is very susceptible to *Aspergillus flavus* mold invasion in the soil, which produces carcinogenic aflatoxins. World production of peanut for oil has slowed because of limitations on feed uses of the meal. Direct food uses of peanut have increased in developing countries.

Some Miscellaneous Oils Processed at the Food Protein R&D Center, Texas A&M University

Pistachio Oil

Pistachio oil is extracted from the fruit of *Pistacia vera* by expeller or screw press. Compared to other nut oils, pistachio oil has a strong flavor. It tastes like the nut from which it is extracted. Pistachio oil is high in Vitamin E, containing 20–25 mg/100 g. It contains 12.8% saturated fats, 54.8% monounsaturated fats, 33.7% linoleic acid, and 0.7% omega-3 fatty acid. Pistachio oil is used as a table oil to add flavor to foods such as steamed vegetables. The fruit is a drupe, containing an elongated seed, which is the edible portion. The fruit has a hard, whitish exterior shell. The seed has a light green flesh, with a distinctive flavor. When the fruit ripens, the shell changes from green to a yellow or red and abruptly splits. This is known as dehiscence. Each pistachio tree averages around 40–50 kg of seeds, or around 60,000, every 2 years. The shell of the pistachio is naturally a beige color, but it is sometimes dyed red or green in commercial pistachios. Most pistachios are now picked by machine and the shells remain unstained, making dyeing unnecessary except to meet consumer expectations. The kernels are often eaten whole, either fresh or roasted and salted. Pistachios are also used in ice cream, pistachio butter, and confections. Recently, the FDA approved the first qualified health claim specific to nuts lowering the risk of heart disease.

Emu Fat/Oil

Emu is the largest bird native to Australia. The soft-feathered, brown, flightless birds reache up to 2 m (6.6 ft) in height. They have long thin necks and legs. Emus are farmed primarily for their meat, leather, and oil.

The Food Protein R&D Center at Texas A&M University has special expertise in emu fat processing to produce premium quality emu oil. There are presently no established special methods to process emu oil. However, current oil processing technology can be used. The entire refining process involves rendering the fat, cooking, fat separation, and fat drying. The crude oil is treated by caustic (caustic neutralization) for free fatty acid removal followed by bleaching and a deodorization. The general physical properties of emu oil are similar to those of a vegetable oil with intermediate content of saturated fatty acid. Emu oil is similar in oleic acid content to oils such as peanut or canola. With regard to saturated fatty acid, emu oil resembles cottonseed oil. The ultimate quality of the refined emu oil depends strongly on the starting material. For example, if the crude oil has blood tissues producing a high content of free fatty acids, even though the refining process will eliminate the acidity of the oil, the final product will contain a high concentration of mono- and diglycerides which easily degrades during storage. Emu meat is a low-fat meat (less than 1.5% fat), and with cholesterol at 80–90 mg/100 g, it is comparable to other lean meats. There is some evidence that the oil may have anti-inflammatory properties.

Mink Oil

Mink oil is produced from the thick fatty layer which lies just under the skin. This fat is removed from the pelt when the mink is skinned and is then rendered into mink oil. The crude oil is dark yellow in color. Crude mink oil is directly bleached and deodorized without being caustic neutralization. The bleached and deodorized oil is light yellow in color. Mink oil and its fatty acids are unique among animal-derived fats and oils. The total unsaturated fatty acids in mink oil account for more than 75% of the fatty acid content, but the oil, nevertheless, has a greater oxidative stability than other animal or vegetable oils. Mink oil is a source of palmitoleic acid, which possesses physical properties similar to human sebum. It is used in several medical and cosmetic products, and also for treating, conditioning, and preserving nearly all kinds of leather.

Ostrich Oil

The Ostrich (*Struthio camelus*) is a large flightless bird native to Africa. It is farmed around the world, particularly for its feathers, which are decorative and are used as feather dusters. Its skin is used for leather products and its meat is marketed commercially. The refining process for ostrich oil is similar to that for EMU oil. It starts with the cleaning of fat to eliminate blood, skin, dirt followed by fat separation (separate the solids by decantation), fat drying (heat under vacuum), refining (remove free fatty acid), bleaching (to eliminate color, peroxides, metals), and deodorization (to remove odor, peroxides, free fatty acids). Ostrich oil is high in omega-3, omega-6, and omega-9 EFAs. It contains around 20% saturated fatty acid followed by 35% monounsaturated and 39% polyunsaturated fatty acids. It is used in cosmetic and in the treatment of lesions, burns, contact dermatitis, dry skin, and many other ailments. Ostrich oil and EMU oil have very similar properties.

Trans Fats Nutritional Labeling

Trans fats have been created, often intentionally, during hydrogenation of TAG to obtain fat solids with specific properties for use in shortenings, spreads, confections, and other semisolid foods. Lesser amounts result during hydrogenation to reduce linolenic acid content and slow the oxidation of fats and their breakdown during frying. Small quantities (seldom more than 2.0%) are created during high-heat treatments in deodorization or physical refining. *Trans* fat formation is part of the oxidative sequence at ambient temperatures and is accelerated in frying. *Trans* fats in ruminant meat and dairy products, produced by biohydrogenation, are considered “natural” as described earlier. The US FDA has estimated the average domestic intake of *trans* fat at about 5.8 g, or 2.6% of calories per day for individuals 20 years of age or over. Consumption of saturated fat is estimated at four to five times more. Estimated sources of *trans* fat are: cakes, cookies, crackers, pies, bread, and the like ~40%; animal products ~21%; margarine and spreads ~7%; fried potatoes ~8%; potato chips, corn chips, popcorn ~5%; household shortening ~4%; salad dressing ~3%; and breakfast cereals and candy ~1% each [161].

In response to a petition by a consumer advocate group, the FDA published a proposal in the *Federal Register* [64] (221, 62745–62825, Nov 17, 1999) to include *trans* fat information in Nutrition Facts labels of packaged foods. (Listing of saturated fat and dietary cholesterol has been required since 1993.) The proposal initiated one of the most intensive dialogues ever in domestic food regulation between consumer advocates, public sector researchers, commodity producers, food processors, and government officials. The net effect was that separate listing of *trans* fat content was required on packaged food Nutrition Facts panels as of January 1, 2006. Also, dietary supplements, which contain 0.5 g or more *trans* or saturated fats per serving, must list their amounts in the Supplement Facts panels. The Nutrition Facts per serving format is:

	Percent daily value[a]
Total fat, 12 g	18
Saturated fat, 3 g	15
Trans fat, 1.5 g	—
Cholesterol, 30 mg	10

[a]Percent daily values are based on a 2000-cal diet

Daily value (DV) has not been established for *trans* fats; thus, calculation of a DV is not possible. Amounts of *trans* fat less than 0.5 g per serving are recognized as "*Trans* fat 0 g." Processors, who wanted more favorable nutrition facts labels on their packaged foods, reformulated their products. Those, choosing to claim "0 *trans,*" *typically* provided fat solids from interesterified fats, fully hydrogenated fats (because these do not contain *trans* bonds), or higher melting fractions of palm and other oils. Exclusion of conjugated structures from the FDA *trans* fats definition enables claiming products where all the fat is from ruminant sources (butter, cheese, cream, dips, ice cream, beef, and tallow) as containing "0 trans." Saturated fat content also must be shown on the label. *Trans* fats are unsaturated. Currently, some nutritionists and regulators are seeking means to inform restaurant, institutional, and fast foods customers about *trans* fats contents of specific products sold at these outlets.

The FDA's regulatory chemical definition for *trans fatty acids* is: "all unsaturated fatty acids that contain one or more isolated (i.e., nonconjugated) double bonds in a *trans* configuration." Under this wording, CLA would be excluded from the definition of *trans* fat, but *trans* vaccenic acid would be included. This definition was reconfirmed as of June 24, 2004 [162].

Copies of announcements, communications, and transcripts of open hearings during regulation making are available at the US-FDA Internet Web site. Readers may gain insight into regulator–nutritionist thinking on the *trans* issue from the transcript of the Nutrition Subcommittee Meeting on Total Fat and *Trans* Fat on April 27–28, 2004 [163].

During the review process, each responding interest group typically rationalized its position using generally accepted research findings. However, the diverse findings have yet to be integrated into a coherent system. FDA's communications have taught that consumption of saturated and *trans* fats causes undesired increases of total cholesterol in blood serum. They also has taught that *trans* fats cause undesired increase of LDL (LDL that carry cholesterol), which is termed "bad" cholesterol and is the rationale for listing *trans* fat in Nutrition Facts labels. However, it is generally accepted in the scientific community that not all saturated fats are equal, and likewise for *trans* fats. More specifically, stearic acid has been considered "neutral" in raising total cholesterol or LDL based on research initiated nearly a half-century ago and frequently reconfirmed. In response to concerns about defamation of palm oil as its world sales increased, that industry initiated extensive nutrition research, even involving noted United States scientists, on palm oil's dietary effects in the mid-1980s. Palm oil contains approximately 44% palmitic acid, 5% stearic acid, 39% oleic acid, and 10% linolenic acid. Repeatedly, diets containing palm oil (palm olein) raised neither total cholesterol nor LDL, and in some experiments decreased total cholesterol and increased "good" HDL (high density lipoprotein). In comparative experiments, *trans* fats fared negatively compared to palm oil, as also did palm kernel and coconut oils and lard. (In palm oil, 75% of the fatty acids in TAG position 2 are unsaturated.) It has been suggested that myristic acid (C 14) leads in increasing LDL, followed by lauric acid (C12). But, myristic and lauric acids, in limited amounts, also play essential nutrition roles. Current nutrition thought is shifting to dietary ratios of saturated and unsaturated fatty acids as more important in the diet than individual fatty acids. However, the world's major saturated fatty acid is palmitic, and together with stearic acid (which occurs in beef tallow consisting of approximately 25% palmitic acid, 19% stearic acid, 36% oleic acid, and 3% linolenic acid) accounts for the far majority of domestic dietary saturated fat intake. This leads to the question: If the majority of saturated fatty acids are essentially cholesterol neutral, why are they grouped with the LDL-raising saturated fatty acids in nutritional labeling?

With exception of the dairy and ruminant animal industries, which are recommending exemption of *trans*-vaccenic acid, no one is promoting increased consumption of *trans* fats. However, the dichotomy of procholesterolemic and cholesterol-neutral saturated fats has divided the nutrition community into two groups, each with well-respected members and spokespersons. Many scientists strongly recommended against combining saturated fat and *trans* fat into one number on the Food Nutrition Facts label in invited letters to the FDA. Others repeated that a better system for classifying "good" and "bad" nutrients is needed. Some nutritionists have suggested that earlier FDA Nutrition Facts labeling and the USDA Food Pyramid have: (1) been interpreted by the public as "all fats are bad;" (2) participated in causing current national obesity problems; and (3) led to development of "low carb" weight loss diets which include increased fat intake [163]. More discussions and new diets can be expected in the future.

Historically, the presence of CLA in vegetable oils was considered minimal, well below 1%, and little was known publicly about their actual occurrence in hydrogenated fats. In 2002, a group of Korean researchers published a pioneering study on effects of catalyst types and concentration, hydrogen pressure, and operating conditions on CLA formation in different oil species. Using selective catalysts, they obtained 23.2, 23.7, and 23.3% total CLA production in corn, cottonseed, and soybean oils (originally containing ~58, 51, and 51% linoleic acid, respectively). As much as 2.4% 9 *cis*-11 *trans*-18:2 and 1.7% 10 *trans*-12 *cis* 18:2 were obtained in mixtures from individual trials [164]. Biological activity of the CLA was shown later. Another publication from the same laboratory, incorporating the latest separation techniques developed in the United States CLA research program, and mass spectroscopy for CLA identification, reported separation and identification of 20 different CLA isomers in hydrogenated soybean oil [165]. These studies report that CLA production can be a result of hydrogenation conditions, and concentrations as much as eight times greater can be produced chemically than in ruminant sources. Thus, CLA are no longer unique to ruminant fats, and potentially can be made available in quantity. However, because of the unknown biological activity of other CLA isomers produced simultaneously, they should be carefully reviewed before broad authorization in the food supply.

New regulations can markedly change equipment requirements and practices of an industry. While regulators, lobbyists, and lawyers negotiated the new law, the fats and oils industry had to start preparing for new market requirements, and food formulators had to develop new products. Likely, the first consideration was selection of ingredients that make Nutrition Facts panels look attractive compared to competitor's products and also have functional properties necessary for making the products. But formulators are well experienced in matters such as ingredients listings. Research is continuing on ways to better control *trans* fat production during hydrogenation [166], but the industry has to formulate with currently available technologies. Reserving the use of hydrogenation for making fully hydrogenated hardstocks (~5 IV), the two most promising routes for obtaining fat solids are blending selected fractionated fat solids, hard stocks, and oils and interesterification of the above. However, as mentioned earlier, skepticism exists about increasing the number of saturated fatty acids in the TAG 2 position by randomization. Enzymatic interesterification at the 1,3 positions seems more appealing in theory.

Trans-free margarines and spreads were well established in Canada and European markets, before promotion in the United States. Many major domestic snack food fryers have adopted *clean ingredient listings* in which only potatoes or corn, vegetable oil, salt, and flavorings are listed (but no hydrogenated products or preservatives). Formulators of more complex foods have options of using emulsifiers, antioxidants, and an ever-increasing variety of hydrocolloids and gums.

Not all fat products have been reformulated for "0 *trans*" labels. Margarines must contain 80% fat, just like butter. Both have been largely replaced by spreads, or "lite" spreads, which contain lower amount of fat; a far majority of spreads claim "0 *trans*." Many home makers prefer to cook or bake with the higher fat content margarines, or (100% fat) shortenings, which often show positive *trans* fat content on their labels.

Edible Uses of Fats and Oils

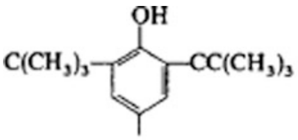
Antioxidants

Fats and fat-containing products may be stabilized against oxidation by the addition of antioxidants as adjuvants. These compounds are believed to act as hydrogen donors or as free radical acceptors that intercept and hold quantums of energy that otherwise might induce oxidation. The major food grade synthetic antioxidants used include: butylated hydroxyanisol (BHA), butylated hydroxytoluene (BHT), tertiary butylhydroquinine (TBHQ), propyl gallate (PG), and 2,4,5-trihydroxybutyrophenone (THBP). Generally, they are allowed in food products at 0.01% of total fat weight each, with a total of 0.02% if a combination is used.

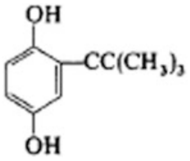
Synergists used in combination with antioxidants include citric acid, isopropyl acid, phosphoric acid, ascorbic acid, ascorbyl palmitate, iso-ascorbic acid, tartaric acid, and lecithin, most of which act as chelators of pro-oxidative metals.

Natural antioxidants include gum guaiac, tocopherols (including vitamin E), and oil of rosemary (containing rosmaridiphenol). antioxidant properties but are not counted in the maximum amounts of antioxidants permitted in the food. Types, mechanisms, and health effects of natural antioxidants are described in the references [168, 169].

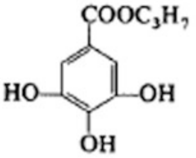
The cross-linking that occurs in drying oils also is an oxidation reaction. Polymerization of oils can occur rapidly in products such as fish meal that contain significant amounts of polyunsaturated triacylglycerols, even to the point of building up sufficient heat for a pile to spontaneously burst into flame. Thus, feed grade antioxidants, such as ethoxyquin, often are added before storage. Ethoxyquin also has been used for sparing the natural tocopherols for vitamin E activity in nutrition.



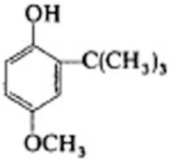
Butylated hydroxytoluene (BHT)



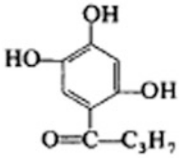
Tertiary butylhydroquinone (RBHQ)



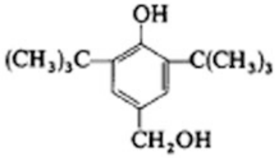
Propyl gallate (PG)



Butylated hydroxyanisol (BHA)



Trihydroxybutyrophenone (THBP)



Hydroxymethylditertiarybutylphenol

Because of cost, they usually are reserved for premium-priced foods. Many types of tocopherols and associated compounds exist, and understanding of the mode of action is continually improving. The forms that give the most antioxidant protection do not always show the greatest vitamin E nutrition activity [167].

Assemblers of grocery store or food service convenience foods have the same general limitations. However, they also have access to natural cereal ingredients and herbs, herb extracts such as oils of rosemary, sage, and other materials in formulation that possess.

Synthetic and natural antioxidants are known to be heat- and steam-distillable and preferably should be added to oil products after the maximum heat encountered in processing. It further is known that many natural phenolic compounds in cereal–oilseed products, as well as reaction products of Maillard (nonenzymatic amino acid-reducing sugar) browning and natural wood smoke, have antioxidant effects. Additional techniques for reducing oxidative activity include maintaining nitrogen blankets on oils stored in tanks and distributing oils in opaque containers or brown bottles to limit exposure to ultraviolet light, although this is not popular with many marketing departments who want to display clear, light-colored oils.

Nonionic Surfactants and Emulsifiers

Nonionic surfactants and emulsifiers, whose molecules have both aqueous (polar) and alkane (nonpolar) compatible sectors, also are common adjuvants. Their molecules have regions that are sufficiently similar to become part of either the water or oils phases, and other regions sufficiently dissimilar to repel that phase. For example, when added to a crystallizing fat, some nonionic surfactants may orient themselves to become part of the crystal, thus preventing further replication and limiting crystal size. Likewise, some will react with gelatinized starch in aqueous and bread systems and prevent its recrystallization (retrogradation). Furthermore, surfactants may orient around discrete droplets to stabilize water in oil (W/O) or oil in water (O/W) emulsions. When this occurs, the compatible end becomes associated with the discrete droplet, leaving the other end turned outward to associate with the compatible continuous phase [170].

The major groups of commercial emulsion stabilizers include: (1) glycerol esters, favoring W/O emulsions; (2) esters of mono-acylglycerols with hydroxycarboxylic acids (including lactic, succinic, malic, and tartaric); (3) sodium stearoyl-2-lactylate (SSL), favoring O/W emulsions; (4) fatty acid monoesters of ethylene glycol; (5) sorbitan fatty acid esters, known as SPANS favoring W/O, and TWEENS favoring O/W emulsions; (6) phosphorlipids; (7) water-soluble gums, including gum arabic, tragacanth, xanthin, agar, pectin carrageenan, and methyl- and carboxymethylcellulose; and (8) proteins [170, 171].

Glycerol has three exposed hydroxyl groups, resulting in a compound completely miscible in water at all concentrations. Fatty acids have both hydrophilic (water-attracting) and lipophilic (oil-attracting) ends. As they are esterified to glycerol, the molecular structure that originally was primarily water-soluble becomes increasingly oil-soluble. Finally, when all three positions are esterified to fatty acids, the molecule is nonpolar and soluble only in organic solvents.

Table 34.12 HLB (hydrophile–lipophile balance) numbers for some surfactants

Name	CAS number	HLB value
Oleic acid	112-80-1	1.0
Acetylated monoglycerides (film formers)	–	1.5
Sorbitan trioleate (SPAN 85[a])	26266-58-6	1.8
Glycerol dioleate	25637-84-7	1.8
Sorbitan tristearate (SPAN 65[a])	26658-19-5	2.1
Glycerol monooleate	25496-72-4	3.4
Glycerol monostearate	31566-31-1	3.8
Sorbitan monooleate (SPAN 80[a])	1333-68-2	4.3
Sorbitan monostearate (SPAN 60[a])	1338-41-6	4.7
Soy lecithin	8020–84–6	8.0
Sodium stearoyllactylate (anionic type)	18200-72-1	8.3
POE sorbitan monooleate (TWEEN 81[a])	9005-65-6	10.0
POE sorbitan monostearate (TWEEN 60[a])	9005-67-8	14.9
POE sorbitan monooleate (TWEEN 80[a])	9005-65-6	15.0
POE stearic acid (monoester) (TWEEN 20[a])	9004-99-3	16.9

[a] Atlas brand names

The hydrophilicity of nonionic surfactants can be characterized numerically as their hydrophile–lipophile balance (HLB). An HLB value of 3–6 indicates that the compound is a likely W/O emulsifier; 7–9, a wetting agent; 8–13, an O/W emulsifier; 13–15, a detergent; and 15–18, a solubilizer (of oil or other non-polar compounds) in water. The HLB values of some common compounds are presented in Table 34.12 [170]. An HLB value of 8.0 is shown in Table 34.12 for lecithin, but manufacturers are able to supply modified lecithins with values of 2–12.

Table Oils

The processing of RBWD table oils has been mentioned earlier. Depending on positioning in the marketplace, these products may also serve as light duty cooking/frying oils and may be brush hydrogenated and contain added natural or synthetic antioxidants and methyl silicones at 0.5–3.0 ppm to slow breakdown, initiation of foaming, and smoking during frying [172]. United States and Northern Europe markets consider light product color and bland flavor as indicators of oil quality. Olive oil aficionados often prefer the green color and its stronger taste as an indication of virgin or minimum processing. All oils revert to stronger flavors in time. Slight flavor reversion generally is acceptable in cottonseed oil, but not in soybean oil. Many years of research were required to reduce the problem in soybean oil, and the mechanism still may not be fully understood. Converting the industry to only stainless steel oil contact surfaces (specifically avoiding iron and copper), reducing residual phosphatides content to essentially "zero," early inactivation of phospholipases and associated enzymes when preparing seed for extraction, nitrogen blanketing and reduced storage temperatures of bulk oil, and the use of antioxidants are believed to have helped.

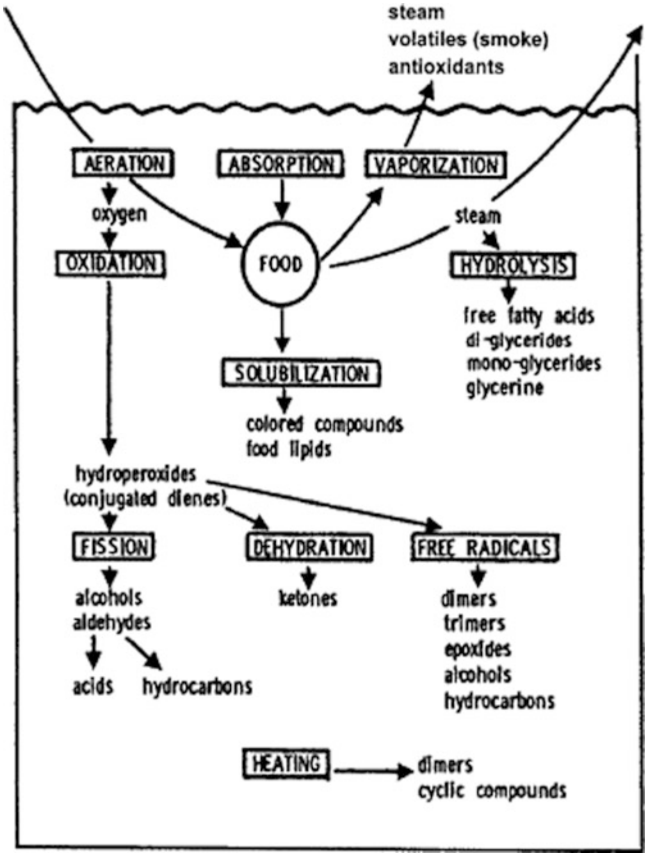


Fig. 34.32 Changes occurring during deep fat frying. (From Fritsch [173], With permission)

Frying Oils

Oil acts as a heat transfer medium during the frying process which rapidly heats, cooks, and sometimes puffs, dehydrates, or forms a crust around the product. The mechanics of deep fat frying are shown in Fig. 34.32 [173]. The entering material carries oxygen that temporarily aerates the hot oil, and water that forms steam, with both rapidly swept out because of low solubility of gases in the hot liquid. Besides enzymes, which are likely to be inactivated rapidly, raw materials also carry oxidation catalysts such as iron in the hemoglobin of fresh meats, chlorophyll, and color pigments. Alkali may accompany the material, especially in chemically leavened doughnut batter and alkali-treated corn snacks. In short, almost every concern mentioned thus far in seed preparation, extraction, and oil processing occurs during frying.

The high temperatures of frying, 177–204°C/350–400°F, are deleterious to oil quality. *Trans* fats are produced, and some of the degradation-prone anisidine value products remaining after refinery deodorization deteriorate. Oils start breaking down into simpler compounds, and also form cyclic compounds and polymers. This process continues while food shop fryers are left hot and inactive during afternoons between preparation of lunch and dinner meals. However, limited self-cleansing occurs in the system by steam distillation, as noted by reduction of carbonyl compounds in oils after inactive fryers are put into service again. An excellent review on frying has been edited by Perkins and Erickson [174].

Unlike in some countries, few United States households have dedicated frying pots containing oil in readiness for meal preparation. Most domestic deep frying is done by commercial operations such as: snack foods processors, producers of convenience fried foods such as Chinese egg rolls and frozen meals, and fast food vendors who prepare French fries, fried chicken, and other products.

Industrial fryers generally are of two types: (1) those who sell all the purchased oil with the product, and (2) those who must recondition and occasionally dispose of oil. Requirements for the selection and use of frying oils in these applications differ [175]. Industrial snack food fryers, designed to use all the oil, are equipped with clean-out systems and inline filters. In theory, they should be able to operate indefinitely, only adding make-up oil as needed. In *clean label,* operations, company policy dictates that no, or minimum, additives can be used. Typically, company personnel periodically inspect and monitor refinery operations of self-certified suppliers. Antioxidants or silicone defoamers are not permitted in these oils. Oils are received in bulk, typically at less than 0.05% FFA, < 1.0 PV, < 4 ppm phosphorous, and <0.75 ppm chlorophyll, and are kept in stainless steel tanks under nitrogen [175].

Stability against oxidation of soybean oil has consistently increased in frying trials as linolenic acid content decreases, whether by plant breeding, mixing with other oils, or by hydrogenation [176–178]. Less than 2% linolenic acid content has been a long-term industry goal, with the way by which linolenic acid is reduced (hydrogenation or breeding) appearing to have little effect on fried product stability [179]. Oils with PVs higher than 2.5 are not used in well-disciplined frying operations.

The freshness of fried snack foods is dependent on use of packaging impermeable to moisture and oxygen, including laminates of windowless aluminized films that block 99+ percent of the light to prevent photo-induced oxidation. Additionally, the pouches are nitrogen flushed before sealing, thus creating pillow packs that further protect the product against crushing. The snacks often are delivered by company delivery personnel, who place them on store shelves or racks. Products may have shelf lives of 6–8 weeks, but inventories are carefully managed to ensure rapid turnover and fresh products [175].

Normally, the warm surface oil serves as the binder (“tacking agent”) for adsorbing salt and dry flavorings to snack foods. The concept of reducing or entirely eliminating oil from snack foods appeared during the early 1990s. Rather than deep fat frying, snacks were dried at high temperatures in fluidized bed continuous dryers. A far smaller amount of oil was then sprayed onto the dried product for flavor. Where a “fat-free” snack food was desired, solutions of edible gums or specialty starches were sprayed onto the snack to serve as tacking agents, for water-soluble flavors; the product then required an additional drying step (to less than 1.5–2.0% moisture content to ensure crispiness) [175].

Industrial fryers, who cannot turn the oil over completely, have lengthened its life 3–10 times by using polydimethylsiloxane (methyl silicone), which is not allowed in some countries. Levels as low as 0.2–0.3 ppm have been found effective, with commercial usage of 0.5–5.0 ppm reported. Users are advised to minimize levels of usage to 1.0–3.0 ppm. Dispersion of polydimethylsiloxane in oil is difficult. The compound operates by suppressing foaming and polymerization and increasing smoke points of oils by up to 13.9°C/25°F. Antioxidants steam distill out of the oil during normal frying. Their initial inclusion essentially protects the oil only until the time of use, but some operators insist on periodically adding them to the fryer. For greater effectiveness, antioxidants in fresh oil are best sprayed onto the product after frying, or included in dry seasoning mixes [175]. Additional steps to prolong the use of frying oil include: inline filters, periodic cleaning of fryers to remove settled charred product, neutralization of fatty acids, and refreshing the oil by passing through adsorbent earth filters continuously, or at the end of the day. Numerous kits and advisory services are available.

Large commercial frying operations, and fast food franchises that prepare French fries and chicken, have provisions and personnel trained to care for frying oils. The greatest food safety concerns are about small restaurants which do occasional frying during midday and evening meals. Several countries have imposed standards on the quality of frying oil in use. For example, Germany requires that the smoke point be not lower than 170°C/338°F, and total polar compounds not exceed 24% [180]. Products fried in oils usually would be objectionable in taste to most Americans long before they reach the unusable specifications. Yet, with exceptions of reduced digestibility and depletion of vitamin E (which can be supplemented), laboratory animals directly fed thermally degraded fats have not done as poorly as anticipated from the history of the oil and known presence of mutagens. Obviously, gaps exist in our toxicology and nutrition knowledge. Many practical details are presented in a new book [181].

Specialty Oils

A variety of specialty oils exists. Pumpkin seed oil is popular in Central Europe. Sesame oil is used throughout the Middle East and the Orient; and roasted sesame oil is used in very small amounts to flavor Asiatic foods. Nut oils, grape seed oil, tomato seed oil, and herb-flavored oils are available.

The organic and "natural" foods markets demand nonchemically treated oils. These often are made by hard pressing the source and water degumming the oil, followed by bleaching with "natural" (nonchemically treated) earths and removing FFA by physical refining.

In 2004, ADM Kao LLC, a joint venture of Archer Daniels Midland Company (United States) and Kao Soap Company (Japan), launched a new type of cooking oil, Enova™ Brand in the United States market. Enova™ is a 1,3 diglyceride (diacylglycerol, DAG) intended for use in nutritional beverages/drinks, nutritional bars, salad dressings, and general cooking [182]. Enova™ oil was initially developed by the Kao Corporation and is the best-selling cooking oil in Japan under the "Econa® Healthy Cooking Oil" brand name. The manufacturer claims the US product is made from all natural soybean and canola oils, has the lowest saturated fat content of any cooking and salad oil, contains zero grams of *trans* fat, and is metabolized differently by the body: instead of being stored as fat, the majority of DAG is burned as energy.

In the early 1970s, Triple "F" Incorporated, a feed manufacturer in Des Moines, IA, developed a low-cost extruder for on-farm use, which inactivated trypsin inhibitor in whole soybean by heat produced by friction shearing. This made soybean directly usable by cattle, pigs, and poultry, without sending it to extraction plants and buying back soybean meal. However, the oil content of the product was too high for direct feeding, and purchase of defatted soybean meal often was necessary. In the mid-1980s, the INTSOY program at the University of Illinois developed a procedure for hard-pressing the sheared soybean, reducing fat content of the meal by about a half and producing saleable crude soybean oil. As domestic livestock operations increased in size, large farmers and cooperatives installed the InstaPro International Company ExPress® Extruder/Press System to process soybean, cottonseed, and occasionally other oilseeds for feed. Production of oil byproducts also increased. Research has shown many desirable properties in the crude hard-pressed ("expeller") soybean oil, including highly hydratable phosphatides content which are easily water degummed and a high resistance of the refined oil to oxidation. After physical refining, the oil shows unexpectedly good stability as frying oil, without hydrogenation or formation of *trans* fats while still retaining its linolenic acid (omega-3) content. After 35-h potato frying tests, soybean oil had a total polar compound level of 17.7% (degradation), whereas the "expeller" oil had 9.1%. Soybean oil with TBHQ antioxidant had 8.8% and hydrogenated soybean oil had 9.7%. Potatoes fried in the expeller oil had significantly higher flavor scores than those fried in the other oils [183]. Test results have been repeated in other laboratories [184]. The reason for the good performance is not yet known. Some believe the high heat (~130°C/266°F) generated in the process almost instantaneously inactivates most of the destructive enzymes (lipases, phospholipases, and especially lipoxygenases) at the time the seed is first sheared [185]. Others believe the benefits come from physical refining [186]. The industry is very excited about market potential for expeller-produced oil.

Margarines and Spreads

Margarine is one of the major temperature-profiled fat products. It was invented in France in 1869 intentionally as a butter substitute and was first produced in the United States in 1873. Originally, it was made from animal fat; coconut oil became the lead fat in margarine in 1917, partially hydrogenated cottonseed oil in 1934, and partially hydrogenated soybean oil in 1956. Various legal principles have been tested through margarine. Among the first was protectionism for butter and the dairy industry, with a requirement that a "Oleomargarine Sold Here" sign be posted on the door of every store offering it. (In retrospect, modern marketers might see this as the best free advertising possible.) In response to dairy industry claims of product inferiority, a law was passed in 1923 requiring that margarine be fortified to the same level of vitamin A as butter; vitamin D was added later. The principle that an intended direct substitute be at least as nutritious as the replacement had become part of the FDA's expectations for new products. Sales of colored margarine were approved on a state-bystate basis in the 1950s and 1960s, and United States per capita consumption of margarine surpassed butter in 1956.

The major enabling technical break-throughs were invention of the internal scraped surface heat exchanger (SSHE, "Votator") in 1937 for chilling the margarine oil mixture, and the lifting of emulsifier restrictions in 1992, which permitted development of a wide range of reduced-fat spreads. Margarine typically has had the same fat content as butter in most nations and is 80% in the United States. Spreads contain less than 80% fat and have become the most popular of the group. Some contain as little as 20% fat, the minimum amount being technical rather than legal.

SFI profiles of several margarine types, an all-purpose shortening, frying oil, and two shortenings are shown in Fig. 34.33 [128]. The SFI profile of butterfat varies with

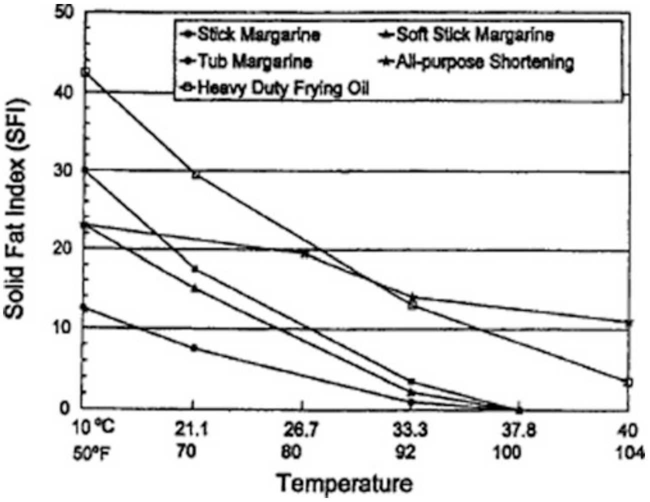


Fig. 34.33 Solid Fat Index (SFI) profiles for hard stick, soft stick and tub margarines, and for all-purpose shortening and heavy duty frying oil. (Data plotted from Erickson and [128])

Table 34.13 Crystal forming tendencies of hydrogenated oils, collected and updated from various sources

Beta prime (β′) type	*Beta* (β) type
Cottonseed	Soybean
Rapessed—HEAR[a]	Canola—LEAR[a]
Lard—modified[b]	Lard—nonmodified[b]
Palm	Palm kernel
Rice bran	Sunflower
Herring	Olive
Menhaden	Corn
Milk fat (butter fat)	Peanut
Tallow	Safflower Cocoa butter Coconut Sesame

[a] By elimination of erucic acid from traditional high erucic acid rapeseed, the C22 fatty acid family also was eliminated. The replacement canola (low-erucic-acid rapeseed) had one less variety of fatty acids and became a *β* crystallizer

[b] A reverse situation occurred in lard, which has few predominant TAG types. Rearrangement created additional types which led to *β*′ crystallization

the season and feed and typically is in the 40–50% solids range at 10°C/50°F. Soft stick margarine is softer than butter when taken from the refrigerator, and tub margarine is spreadable.

The basic sequence for making margarines/spreads includes:

- Formulation of an oil mixture from hydrogenated base stock or interesterified fats, hard stock, and oil that has the desired SFI or SFC profile.
- Ensuring that sufficient diversity occurs in the species and fractions to provide a variety of TAG that will form small *β*′ crystals.
- Compounding the margarine oil blend at the refinery; final deodorization; shipping the melted blend to the margarine plant; and storage under a nitrogen blanket.
- Preparing an oil-soluble additives mixture containing mono- or diglycerides and other emulsifiers, flavorings, oil-soluble vitamins, and yellow color at the margarine/spread plant.
- Preparing a water-soluble additives mixture containing water, salt, anti-microbial compounds, viscosity thickeners, and water-soluble flavorings. Formula amounts of thickeners and water-soluble flavorings increase as the fat content of a spread is reduced. The water should be microbiologically potable and deionized to remove calcium, magnesium, iron, and copper.
- Blending the heated oil-soluble additives mixture into the warm oil in a mixing tank. Slowly adding the warmed water-soluble additives mixture into the blend to produce a water-in-oil emulsion.
- Pasteurization; partial cooling.
- Chilling and working the emulsion in a series of (typically ammonia-chilled) internal scraped-surface heat exchangers (“A” units) and “picker bar” workers (“B” units), with resting tubes interspersed to ensure adequate time for establishment of the desired crystal structures.
- Filling.
- Holding under controlled temperature conditions for establishment of the desired crystals.
- Shipping.

Much of the physical chemistry for making temperature-profiled fats, including Figs. 34.1 and 34.2, has already been reviewed. Additional references are available [53, 187]. We cannot change the laws of thermodynamics, but can slow or hasten arrival at equilibrium. The margarine literature often shows a table similar to Table 34.13, but does not explain the practical implications. Basically, the beta prime (β′) tending oils contain a greater variety of fatty acids than the beta (β) tending. A greater diversity makes it more difficult for identical TAG to align and crystallize. It is easier to make whole cottonseed oil margarines/spreads than whole soybean oil counterparts because the former contains about 2.5 times more palmitic acid than soybean oil. For years it was common practice to include about 10% cottonseed hard stock in soybean margarines. A reduction in TAG diversity increased the problems of making margarine when erucic acid was eliminated from canola (rapeseed); but increasing diversity, as by rearrangement of lard and increased presence of *trans* fats, has assisted in making smooth-textured margarines. A similar lack of diversity is encountered in making palm oil margarines, and addition of C 18 hard-stocks and lengthening pin working and resting times have been employed. The newer high oleic acid oil varieties often present more difficulties in texture development than the traditional varieties in the same oil species. Crystals grow by aligning

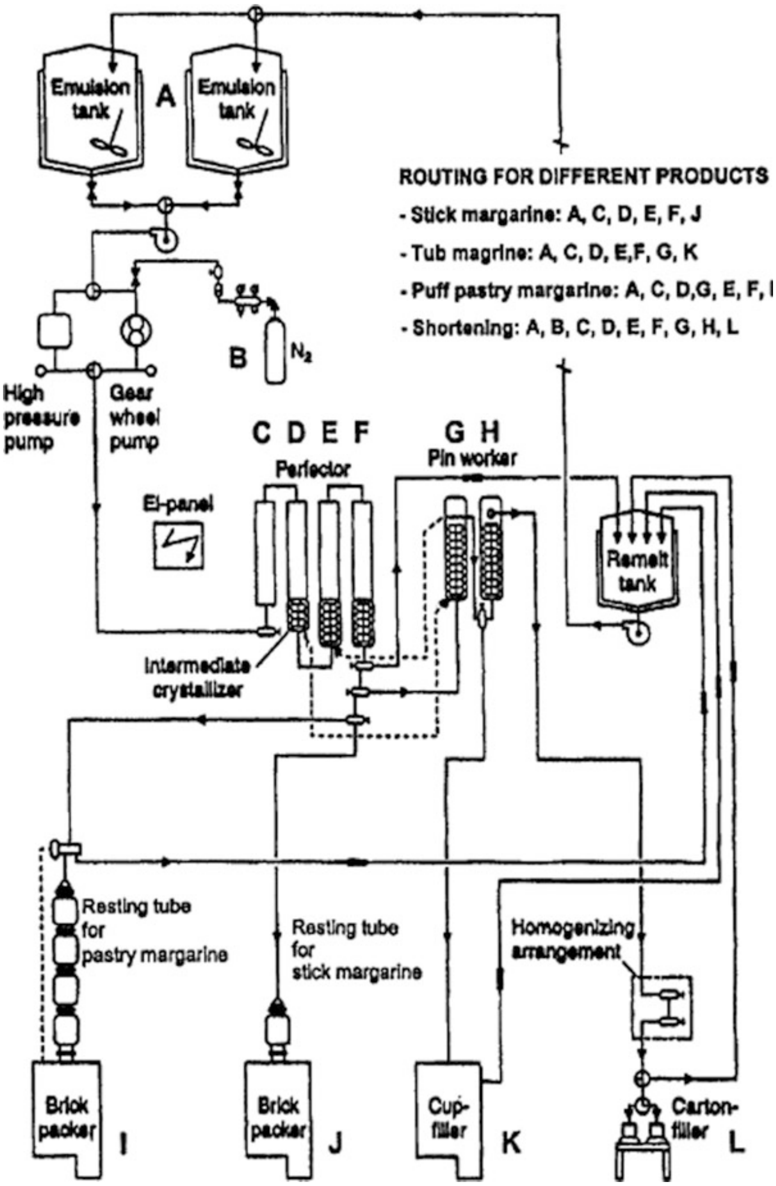


Fig. 34.34 Flexible Perfector Plant™ for making stick, tub, and pastry margarines, and shortening. (Courtesy of Gerstenberg Schröder, Brondby, Denmark)

identical molecules side by side. Emulsifiers, such as mono- or diglycerides, can align next to the crystal and be accepted as part of the lattice. But, not being identical, they foul the surface and stop crystal growth. Increasing the viscosity of lower oil content spreads makes it more difficult for similar TAG to align and crystallize, especially if temperature cycling occurs during storage.

Margarine, like butter, and spreads, is a water-in-oil emulsion. The water and all water-soluble ingredients are encapsulated in a continuous phase of oil. This has many practical advantages. Growth in individual droplets is limited by available nutrients, and bacteria cannot cross over to a second nearby food supply. If the product is salted at 1.5%, the water droplets will have a salt concentration of 7.5%, which is inhibitory to many bacteria. If the emulsion has been properly made and pasteurized, many droplets will be small and not contain bacteria cells. Lecithin is included in margarines/spreads at 0.1–0.5% as an antispattering agent used for frying; it is introduced with the oil-soluble additives initially added to the base oil.

The margarine/spread production flow sheet in Fig. 34.34 shows how these principles are applied. The figure represents a multipurpose plant, in which various products can be made. The legend in the upper right corner shows results of the sequences employed. The emulsions are prepared in tanks (A), although the provisions for pasteurization and cooling before processing are not shown. Consumer margarines must meet FDA Standards of Identity, which specify the permissible additives. Industrial buyers may arrange to have other FDA-approved ingredients added as a means to simplify their dispersion in manufactured products. The emulsion is pumped to an internal scraped-surface chiller and exits at

10–19°C/50–65°F as a soft glass (stiffening when held in the hand). Because this is not the type of final crystal desired, the chilled emulsion is then sent through a series of pin workers (tubes through which the margarine is pumped while a shaft with fingerlike pins mixes it), chillers, and resting stops to craft the desired crystal. The final temperatures for table margarines and margarine–butter blends are −18–−20°C/0–14°F, and −5°C to −1°C/12–23°F) for low-calorie spreads [188].

Figure 34.34 shows that tub margarine is pin-worked considerably more to keep it more liquid until filling. The solids content profile also is different from the stick margarine, and a smooth-textured product with minimum stiffening is desired.

Some bakery products ("puff pastries," phyllo) consist of flaky layers, made by rolling the dough thinly, covering it with shortening, folding, rerolling, and repeating the process many times. On baking, they puff up. The gluten layers must stay intact, and the shortening also must be flexible but not give off free oil during the machining. Puff pastries such as croissants, whose fat melt and clean-up in the mouth, can be made from butter, but must be prepared at low temperatures (10°C/50°F) in refrigerated rooms using chilled equipment. A very flat SFI curve (30.0–33.5 at 10°C; 28.0–30.0 at 21.1°C; 24.5–26.5 at 33.3°C; and 19 minimum at 40.0°C) is recommended for tallow–vegetable oil puff pastry margarines. They melt slightly above mouth temperature, but some people notice greasiness. Historically, puff pastry margarine was made using chilling rolls, but processes exist now for using scraped-surface heat exchangers, "B" units, and large resting tubes [188].

Shortenings

Typically, the function of fats in baking is to "shorten" or control development of wheat flour gluten and avoid or limit toughness in the product. In doing so, starch becomes the predominating matrix, and tenderizing complexes can be established between the fat and carbohydrates. A fat with a flatter solids temperature profile like the "All-purpose shortening" in Fig. 34.33 can accomplish its "shortening effect" during machining of the dough and baking even if it does not melt substantially during eating.

Numerous emulsifiers are available for dough (continuous gluten) and cake (sugar, corn sweetener, starch, and fat emulsion) systems. Emulsifiers act as conditioners in dough systems by: improving tolerances to variations in flour and other ingredients; increasing resistance to mixing and mechanical abuse; providing increased gas retention, shorter proof times, and increased product volume in yeast-leavened systems; improving uniformity of gas cell size, cell walls strength, and grain texture; improving slicing; and extending product freshness by delaying starch retrogradation and staling [171].

Cakes essentially are emulsified slurries before baking. Considerable use is made of sugar (with corn sweeteners increasingly used), starch in low-protein content flours, and fat. Emulsifiers have three functions in cake systems: to improve air incorporation; to disperse shortening into smaller particles to maximize the number of air cells; and to improve moisture retention. Complexes occur between fats, emulsifiers, and starches that result in smooth, tender, moist cakes. Emulsifiers include broad classes of lecithin and lecithin derivatives, mono- and difatty acid glycerol esters, hydroxycarbocylic acid and fatty acid esters, lactylated fatty acid esters, polyglycerol fatty acid esters, ethylene or propylene glycerol fatty acid esters, ethoxylated derivatives of monoglycerides, and Sorbitan™ fatty acid esters [171]. Many oil processors sell proprietary brands of shortenings containing emulsifiers, and large bakers can arrange to have emulsifiers added to their melted fat mixtures before shipping.

As shown in Fig. 34.34, nitrogen is injected into the first scraped-surface chiller, with the shortening oil mixture, to give an opaque white appearance and increase plasticity. Some bakers prefer to use semi-plastic shortenings in applications where oil may shorten the dough too rapidly, and for "creaming" sugar and shortening as in making cake frostings. Some small restaurants prefer semi-solid "shortenings" and have fryers able to accept standard 18.2-kg (40-lb) cubes. Care should be taken to avoid development of large air spaces when melting cubes over the (electric) heating rods to avoid burning the shortening.

Other Edible Applications

Cocoa butter (CB) has a challenging chemistry and has attracted many efforts to develop lower cost, acceptable alternatives. The following definitions provide a quick introduction to this field: (1) cocoa butter equivalents (CBEs) are compounded mostly from tropical oils other than palm. Because their melting and crystallization properties closely resemble CB, they are compatible as diluents at all levels of substitution; (2) cocoa butter replacers (CBRs) are made from nondairy oils (typically soybean, cottonseed, or palm) partially hydrogenated for maximum *trans*-C18 isomer formation to acquire a steep melting profile. They are best used for enrobing bakery products, but their melting profiles can be improved by chill fractionation; and (3) cocoa butter substitutes (CBSs) are made primarily from lauric-type fats (C8–C12 of palm kernel and coconut origin) and are hydrogenated under conditions that favor *trans* formation, but may contain a limited amount of hydrogenated C18 trigly-cerides. CBS limitations include: they must be used alone because of incompatibility with cocoa butter, they require the use of cocoa for "chocolate" flavor, and all ingredients with active lipase systems must be avoided to prevent the development of a soapy flavor [189].

The reader is referred to the volume by O'Brien [172] for other food applications of fats, including icings for sweet goods; spray-dried nondairy creamers; coffee whiteners; aerated whiteners with encapsulated air to produce a cappuccino effect when added to coffee; dried powders designed to be reconstituted, pasteurized, homogenized, and packaged for restaurant use as creamers; vending machine dry creamers; whipped topping shortenings, also used for making bakery cream pie fillings and cake toppings, aerosol toppings, powdered toppings, and frozen ready-to-use toppings; cheese analogue shortenings, frozen mellorine dessert shortenings, sour cream analogue, and dip base fats; and sweetened milk and sweetened condensed milk analogue fats.

Industrial Uses of Fats and Oils

Timeline

Industrial uses typically means nonfood-nonfeed applications, although feed uses are not consistently excluded. No one knows when man first used oils or fats for lighting, medical, and cosmetic applications, lubricants, or combined them with wood ashes to make soaps. The following sequence has been published for soybean oil [21].

- <AD 980—China: soybean oil likely used for illumination in lubricating fluids and coatings.
- AD 980—China: documented use of soybean oil boat caulking materials.
- 1908—Europe: soybean oils used in soaps; glycerin sought for making explosives for Panama Canal project and printer's inks; oil additionally used in rubber substitutes and linoleum flooring.
- 1910—United States: soybean oil classified as a drying oil; used as cheap replacement for linseed oil in paints.
- 1914–1918—United States: largest soybean oil industrial market is soaps, followed by paint, varnish, enamel, linoleum, oilcloth, asphalt, and other waterproofing materials.
- 1919—Blowing warm air through heated soybean oil found to increase viscosity by initiating oxidation and polymerization; blown oil improves properties of printing inks.
- 1926—Soybean oil used for plasticizing and increasing the elongation of rubber.
- 1930s—United States: Kienle and Hovey of General Electric developed soybean oil alkyd resins used in paints to improve drying, adherence, endurance, and color; Ford Motor Company used soybean oil and its derivatives in enamel paints for automobiles; DuPont's "four-hour enamel," based on soybean oil, is considered the most important factor in furthering soybean oil usage in paint. Japan also initiates programs.
  – Quality of soybean oil first reported suitable for food use and hydrogenation; replacement of cottonseed oil begins.
  – Strong chemurgic movement initiated; credited with nearly 200 industrial uses during its ten-year existence. Ford Motor Company uses significant amounts of soybean oil for enamel paint, glycerin for shock absorbers, in 1937 automobiles.
- 1945—Chemurgic programs decline after World War II.
- 1950s—D. Swern, USDA scientist, develops epoxy plasticizers from oils or monohydric fatty esters for use in plastics.
- Late 1970s—Converting oil triglycerides to methyl or ethyl fatty acid esters by transesterification ("alcoholysis") reduces injector fouling, carbon deposits, and degradation of lubricating oils in diesel engines, compared with direct use of vegetable oils or diesel fuel–vegetable oil mixtures.
- Early 1980s—American Newspaper Publishers Association develops first-generation soybean oil-based inks to replace uncertain mineral oil supplies.
- Mid-1980s—Degummed, alkali-refined cottonseed and soybean oils used in pesticide aerial spays.
- 1987—USDA Federal Grain Inspection Service allows the use of soybean and other edible oil sprays to reduce the risk of grain dust explosions in elevators. Use of one to two percent soybean oil in livestock feeds reduces dust in pig-rearing facilities, improves animal health, and gives five to ten percent increase in weight gains.

Chemurgy Revisited

Chemurgy is the use of replenishable farm crops as feedstocks for industrial processes. The concept became very popular with onset of the Great Depression in the early 1930s because it promised new uses for agricultural crops and development of a self-sustaining national economy. Famous advocates of the era included Henry Ford, industrialist, and George Washington Carver, crop scientist. The movement was heavily subsidized by the federal government, slowed as World War II approached, and ended in the prosperity of the postwar era. Currently, the United States again has agricultural surpluses, especially soybean oil. Naturally, the following question arises: Are business prospects for chemurgy products different now than in 1937, especially with the United States committed to an open global trade policy? New factors include: (1) Replenishable materials offer the promise of biodegradability, which is becoming increasingly popular with an environment-concerned public. (2) The global politics of petroleum carry many hidden costs, including threats of supply interruption, occasional price fixing, and potential involvement in wars to keep trade sources and routes open. (3) The petroleum companies are driven by economies of scale and have limited flexibility in responding to small, although reliable, markets.

(4) Biotechnology offers the promise of tailored crops, able to produce high levels of specific chemical feedstocks, previously unavailable. But (5) the portion of the public engaged in agriculture and agribusiness is the smallest ever, and federal subsidies may not be easily obtainable in the future.

Nevertheless, current chemurgy research and applications is the highest ever, and many products have been launched by small entrepreneurs to regain markets previously lost by vegetable oils to petroleum feed stocks (Table 34.14). Several factors seem obvious. (1) Supplies of fossil carbon sources (coal and petroleum) and minerals are limited and access to new sources is increasingly expensive. It seems almost certain that plastics and composites will become more important as structural and manufacturing materials, and people will have to increasingly replenish more of their carbon needs through agriculture. (2) The question of where replenishable resources will be grown and converted into industrial and consumer products is heavily political and yet to be answered as national policies compete with global companies. (3) the transition will not occur smoothly. Interests vested in the status quo are likely to first act to preserve current positions as long as possible, but will provide capital when change becomes inevitable.

Industrial Oils Utilization

Estimates of recent domestic edible and industrial oils and fats use are shown in Table 34.15, and industrial uses of soybean oil in Table 34.16.

Fatty acids chemistry and processes have been summarized by Johnson and Fritz [190] and oleochemicals manufacture and use by Gunstone and Hamilton [191]. Pathways for converting oils and fats into various oleochemicals are shown in Fig. 34.35 [21, 192].

Crude soybean oil has limited uses as sprays. Spray nozzles are in danger of clogging by phospholipids, which also leave repeatedly sprayed surfaces sticky. Generally, *once-refined oil* (degummed, alkali-neutralized, water-washed, and dried) is the minimum quality used for dust control and aerial spraying of pesticides. Anticorrosion and anti-polymerization agents are added in lubricants and also in hydraulic fluids. Current industrial applications include: inks, paints and coatings, biodiesel fuels and additives, lubricants, ion exchange resins, adhesives, foams, fatty alcohols, fatty amines, and associated processing [191, 193–196].

Table 34.14 Examples of soybean oil and lecithin industrial uses[a]

Soybean oil technical uses	Oleochemicals from oil	Soybean lecithin
Anti-corrosion agents	Methyl esters	*Wetting agents*
Anti-static agents	Soy diesel fuel	Dry powders
Candles	Solvents	Cosmetics
Caulking compounds		Plant pigments
Composite materials	Fatty acids	
Concrete release agents	Fatty alcohols	*Nutritional*
Core oils	Glycerin—industrial and explosive uses	Medical
Crayons		Vitamins—animal feeds
Dust control agents		
Electrical insulation		*Anti-foaming agents*
Epoxy resins		Alcohol
Fungicides		Yeast
Hydraulic fluids		
Printing inks		*Dispersing agents*
Linoleum backing		Inks
Lubricants		Pesticides
Metal casting/working oils		Magnetic tapes
Oiled fabrics		Paints
Paints		Papers
Pesticide carriers		Synthetic rubber
Plasticizers		
Protective coatings		*Other*
Putty		Viscosity modifications,
Soaps/shampoos/detergents		concrete, drilling muds
Solvents		Softening and curing leather
Vinyl plastics		
Wallboard		
Waterproof cement		

[a] Courtesy of American Soybean Association, St. Louis, MO

Fatty Acid Methyl Esters, Biodiesel

Fatty acid methyl esters (FAME) are the gateway to many products. Use for glycerol, a byproduct of alcoholysis interesterification. Current FAME uses include: cleaning graffiti stains and sticky deposits, light lubricants, degreasing baths, inclusion in penetrating oils, asphalt and concrete

Table 34.15 Reported fats and oils uses in the United States, 2002 (million pounds)[a,b]

		Percent of reported	
Fat or off	Amount (million pounds)	Crop oil	Total edible oils
Edible uses			
Coconut oil, Total edible	294	100.0	1.3
Corn oil, Total edible	950	100.0	4.3
Cottonseed oil:			
Baking or frying fats	200	37.1	0.9
Salad or cooking oil	317	58.8	1.4
Total edible	539	100.0	2.4
Lard:			
Margarine[c]	14	5.9	0.1
Total edible	238	100.0	1.1
Palm oil, total edible[d]	W	–	–
Peanut oil, total edible[d]	W	–	–
Canola oil (Edible rapeseed):			
Salad or cooking oil	732	81.0	3.3
Total edible	904	100.0	4.0
Soybean oil			
Baking or frying fats	8,572	48.1	38.3
Margarines	1,242	7.0	5.6
Salad or cooking oil	7,880	44.2	35.2
Other edible	125	0.7	0.6
Total edible	17,818	100.0	79.7
Sunflower oil, total edible	269	100.0	1.2
Tallow, total edible	252	100.0	1.1
Nonidentified edible	1,101	–	4.9
Total fats and oils:			
Baking or frying fats	9,704	–	43.4
Margarines	1,333	–	6.0
Salad or cooking oils	10,924	–	48.8
Other edible	403	–	1.8
Total edible uses	22,365	–	100.0
Selected Industrial Uses			
Fatty acids	2,178	32.8	–
Animal feds	2,670	40.2	–
Soaps	374	5.6	–
Paints and varnishes	111	1.7	–
Resins and plastics	138	2.1	–
Lubricants and similar oils	112	1.7	–
Other inedible products	1,054	15.9	–
Total Industrial Uses	6,637	100.0	–
Total US uses	29,002	–	–

[a] From Oil Crops Situation and Outlook Yearbook, Economic Research Service US Department of Agriculture, October 2003, OCS-2003
[b] US Census Bureau statistics
[c] Includes lard and edible tallows
[d] W = Withheld to avoid disclosing figures for individual companies

Table 34.16 Estimated uses of soybean oil in US industrial applications, 2003.[a]

Market	Soy oil (million lb)
Biodiesel	214.0
Solvents/specialty	1.8
Plastics/coatings	149.8
Paints/coatings	129.5
Other coatings/inks	99.5
Ttl. loss/metalwork lubes	10.7
Polyols (carpets, foams)	29.9
Soaps, fatty acids	80.2
Other uses	200.0
Total	915.4

[a] Courtesy of United Soybean Board, St. Louis, MO

mold release agents, and adjuvants in various applications. FAME are called *methyl soyate* if made from soybean oil. Large investments in manufacturing facilities are being made in expectation of rapid growth of this industry.

Vegetable oil fuels have been prepared by various methods, including micro-emulsification, transesterification, and pyrolysis. Tests during the mid-1970s and early 1980s showed that diesel engines can initially run on vegetable oils or animal fats, or their mixtures with diesel fuel; but, despite various additives to the fuel and engine oil, problems eventually were encountered with fuel injection valve clogging, cylinder head carbon deposits, and engine oil fouling by fuel blow-by.

The fuel properties of some vegetable oils and soybean FAME are presented in Table 34.17. The heat of combustion of various vegetable oils is nearly 90% that of No. 2 diesel fuel, but engine viscosity is 10–20 times greater. Conversion of soybean oil to methyl soyate reduces fuel viscosity to approximately twice that of diesel fuel. Cetane numbers (CN) indicate the comparative ignition delay time of fuels in the combustion chamber: the shorter the ignition delay, the higher the CN [195]. A flow sheet of a process currently installed for making methyl or ethyl fatty acid esters is shown in Fig. 34.36.

Problems encountered in handling vegetable oil-based fuels include: higher viscosity and higher cloud and pour points, which may require supplemental heating of fuel tanks in cooler weather. Various techniques have been tried to reduce cold temperature viscosity of methyl soyate. Winterization lowered the crystallization temperature by 7.1°C [197]. Increasing the molecular diversity by the addition of isopropyl and 2-butyl (branched alcohol) esters lowered the crystallization temperature of soybean methyl esters by 7–11°C and 12–14°C, respectively [198]. But, molecular diversity effect is reported significantly diluted in 20:80 mixtures of methyl soyate and No. 2 diesel fuel [195].

Domestically, methyl esters are made from soybean oil, spent frying oils, and inedible animal fats. Palm oil is used in Southeast Asia, and rapeseed in Europe. Although appreciable in quantity, supplies of such materials actually are small compared with the amount of TAG that would be required if significant quantities of methyl esters were used in fuels. Governmental support of biodiesel development has included funding of research and demonstration projects, and reducing or eliminating state or federal taxes collected for its fuel use. In turn, this funds the amount of methyl esters that can be blended to make biodiesel competitive with nonblended fuels.

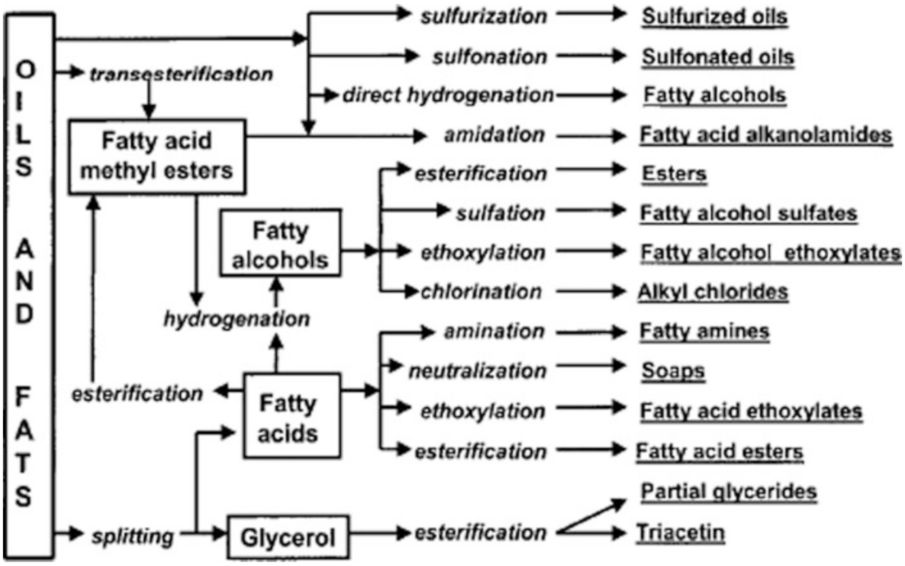


Fig. 34.35 Oleochemical derivatization pathways. (Modified from Zoebelein [192])

Table 34.17 Fuel properties of some vegetable oils and soybean esters[a]

Oil or ester	Viscosity (mm^2/s)	Cetane no.	Gross heat of combustion (kJ/kg)	Cloud point (°C)	Pour point (°C)
Oils[b]					
Castor	297.0	–	37,274	None	−31.7
Corn	34.9	37.6	39,500	−1.1	−40.0
Cottonseed	33.5	41.8	39,468	1.7	−15.0
Crambe	53.6	44.6	40,482	10.0	−12.2
Linseed	27.2	34.6	39,307	1.7	−15.0
Peanut	39.6	41.8	39,782	12.8	−6.7
Rapeseed	37.0	37.6	39,709	−3.9	−31.7
Safflower	31.3	41.3	39,519	18.3	−6.7
High oleic safflower	41.2	49.1	39,516	−12.2	−20.6
Sesame	35.5	40.2	39,349	−3.9	−9.4
Soybean	32.6	37.9	39,623	−3.9	−12.2
Sunflower	33.9	37.1	39,575	7.2	−15.0
Soybean esters[c]					
Methyl soyate	4.1	46.2	39,800	2	−1
Ethyl soyate	4.4	48.2	40,000	1	−4
Butyl soyate	5.2	51.7	40,700	−3	−7
No. 2 diesel fuel	2.7	47.0	45,343	−15.0	−33

[a] From: Foglia et al. [195], With permission
[b] Viscosity determined at 38°C
[c] Viscosity determined at 40°C

The attraction of biodiesel is reduced discharge of undesirable combustion emission compounds into the air. Undoubtedly, effectiveness is related to the degree of petrochemicals replacement. Biodiesel is sold with a numerical designation of the oil/fat-ester content. B20 is a frequent blend, containing 80% petroleum diesel fuel and 20% FAME. A major farm equipment manufacturer has announced it will ship its new equipment with B2 in the fuel tanks. A salt mine has begun using B 100 as a means to improve air quality for its workers.

Progress on biodiesel development can be followed on the National Biodiesel Board (NBB), Jefferson City, MO, Web site (www.biodiesel.org), and the United Soybean Board (USB) St. Louis, MO, Web site (www.unitedsoybean.org) for soybean oil-based fuels; also *Render Magazine*, Camino, CA (www.rendermagazine.com) has kept that industry's members informed on worldwide developments on biodiesel development [199–205].

While a number of bio feedstocks are currently being studied, algae (an autotrophic organism, ranging from unicellular to multicellular forms) have emerged as one of the most promising sources of biodiesel feedstock. Algae have some important advantages over other oil producing crops. It can be grown in almost any enclosed space and it requires

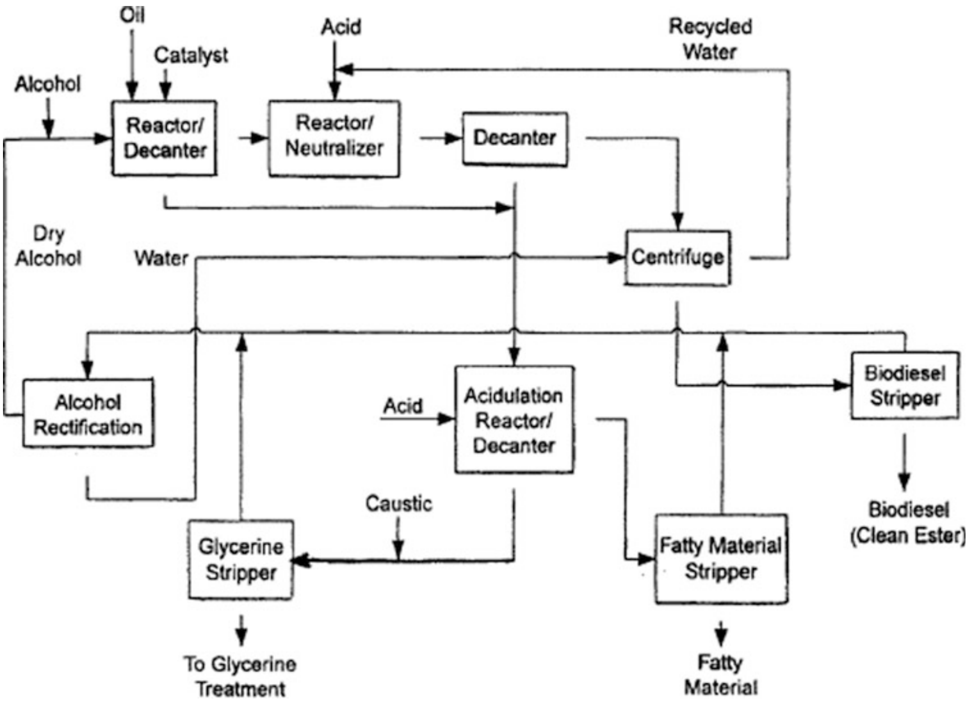


Fig. 34.36 Basic flow diagram of Crown biodiesel ester process. (Courtesy of Crown Iron Works, Minneapolis, MN)

Table 34.18 Global Biodiesel Production, 2008-2010 (metric tons)[a]

Country	2008	2009	2010	% Change (10/09)
EU-27	7,752,825	9,046,000	10,296,033	14
Brazil	1,027,078	1,415,089	2,156,007	52
Argentina	756,802	1,179,204	1,848,006	57
United States	2,302,016	1,815628	1,049399	-42
Thailand	394,241	536,802	580,802	8
Colombia	70,400	290,401	369,601	27
Indonesia	96,800	308,001	352,001	14
Korea, South	158,400	264,001	352,001	33
Malaysia	195,000	222,000	275,000	24
Canada	118,800	164,561	167,201	2
Philippines	63,360	132,000	109,120	-17
Australia	47,520	86,240	52,800	-39
Paraguay	8,800	7,040	10,560	50
Uruguay	2,640	4,400	7,000	59
Turkey	24,992	6,952	6,952	0
Honduras	968	880	880	0
Total	13,020,643	15,479,198	17,633,364	15

[a] *Source*: USDA/Foreign Agriculture Service GAIN reports, National Biodiesel Board, European Biodiesel Board. (Render April 2011)

very little support to grow such as sunlight, water, and carbon dioxide. However, maintaining a pure culture is a major challenge, including during the oil extraction process. Currently, efforts are being focused on these two areas by academicians and by the bio-industry. Algae contain anywhere between 2 and 60% of lipid oil by weight depending on type of strain. Microalgae have been investigated for the production of a number of different biofuels including biodiesel, bio-oil, bio-syngas, and bio-hydrogen.

The global market for biodiesel is changing rapidly, creating both uncertainty and opportunity. The first-generation biodiesel markets in the United States and Europe have reached impressive production levels, but remain constrained by feedstock availability (see Table 34.18). As a result, a surge in demand for alternative feedstock is driving new growth opportunities in this sector. Biodiesel from non-food feedstock is also gaining attention around the world. For example, jatropha plantations in India and China. In Brazil and Africa, there are significant programs underway dedicated to producing non-food crops such as jatropha and castor for biodiesel. Also, an increasing number of second-generation biodiesel projects are now emerging in anticipation of growing sustainability concerns by governments, and in response to market demands for improved process efficiencies and greater feedstock production yields. A great deal of research is under way on reducing the cost of feed-stocks such as algae, jatropha, castor, used vegetable oils, tallow. Given recent efforts in these areas, it is reasonable to expect significant advances at the industrial scale in the near future.

Other Industrial Applications

Essential oils, used in perfumes and cosmetics, are extracted in several ways by perfusion into fat, distillation, batch solvent extraction, and critical CO_2. The oldest method, used today only for extremely valuable essential oils, is *enf eurage–def eurage*. This consists of layering blossoms of plants to be extracted between lard-coated glass plates and allowing the essential oil to perfuse into the fat. The blossoms are renewed daily. At the end of the season the fat is scraped from the plates, melted, poured into containers, and sold as pomade or is batch-extracted with cold ethanol and sold as extraits. Another technique consists of macerating the material and extracting with hot fat. The most common extraction process for essential oils is steam distillation of a mash of the leaves or seeds to be extracted. Selective solvent batch extraction also is used.

New oilseed crops, currently studied as potential sources of specialty fatty acids, include *Crambe abyssinica* for erucic acid, *Limnanthes alba* for very long-chain fatty acids, *Dimorphotheca pluvialis* for dimor-phecolic acid, *Lesquerella fendleri* for lesquerolic acid, *Calendula officinalis* for calendic acid, and *Euphorbia lagascae* and various *Vernonia* species for vernolic acid.194 The lowest cost sources (inedible fats and oils and palm oil fractions) are most likely to be exhausted first as world trade in industrial applications grows. Cornstarch is becoming a major feedstock for plastics production. This may compete with potential oil uses, but also will increase production of corn oil.

Other current areas of soybean oil industrial applications research include plastics, coatings, lubricants, and hydraulic fluids. Potential applications are only limited by imagination, economics, and the business skills of the respective entrepreneur. The United Soybean Board maintains a Web site (www.unitedsoybean.org) of current soybean oil-based industrial products manufacturers, listed under the categories of adjuvants, alternative fuels and fuel additives, building and construction, cleaners, concrete, dust suppressants, engine oils, hydraulic fluids, ingredients, metal working fluids, printing, and miscellaneous. The categories list suppliers and as many as several hundred products each. In many cases, suppliers list their Web sites for interested persons to learn more about uses and specifications of their products.

Analytical Methods

Selected analytical methods, adopted by the American Oil Chemists' Society (AOCS) [22] for characterizing the composition, structure, physical properties, and stability of fats and oils, are summarized below. Prescribed equipment must be used and conditions followed. Some of these techniques are limited to specific oil species, but adaptations are available for other species. Procedures for sample drawing and preparation also are specified.

AOM for fat stability (Cd 12–57): determines the time (in hours) for a sample of fat or oil to attain a predetermined peroxide value (PV) under the conditions of the test. The method is used to estimate the comparative oxidative stability of fats and oils. The method has been placed in surplus, in favor of Cd 12b-92 (Oil Stability Index), but retains official status and is still used in domestic industry. *p*-anisidine value (AV) (*Cd 18–90*): determines the amount of aldehydes (principally 2-alkenals and 2,4-dienals) in animal and vegetable fats and oils. These are degradation products of peroxides, which are not removed by bleaching. Some fats and oils chemists propose increased use of this method in purchase specifications.

Bleaching Test for Soybean Oil (Cc 8e-s63): determines the color of a sample of soybean oil after treatment with a specified bleaching earth. Specific methods exist for other oil species.

Boemer Number, Foreign Fats in Lard (Cb 5–40): estimates the presence of tallows and similar fats, based on differences in melting points of foreign glycerides and fatty acids as compared with pure pork fat. *Cloud Point Test (Cc 6–25)*: determines the temperature at which a cloud first forms in cooling a sample of melted fat to the first stage of crystallization

Cold Test (Cc 11–53): measures the relative resistance of a sample to crystallize at an established temperature in terms of time. The 5.5-h test at 0°C is used as an index of stearin removal in the winterization of salad oils.

Color Measurement by Lovibond–Wesson (Cc 13b-45): determines the color of clear oil samples by comparison with glasses of known color characteristics.

Fatty Acid Composition by Gas Chromatography (Ce 1–62): *quantitatively* determines saturated and unsaturated fatty acids with 8–24 carbon atoms in animal fats, vegetable oils, marine oils, and fatty acids after conversion to their methyl ester forms.

Fatty Acid Composition by GLC (Ce Ic-89): measures the fatty acid composition and levels of *trans* unsaturation and *cis, cis* methylene-interrupted unsaturation of vegetable oils using capillary gas liquid chromatography.

Flash Point–Pensky–Martens Closed Cup for Fats and Oils (Cc 9b-55): determines the temperature at which an oil sample will flash when a test flame is applied. This technique sometimes is used to estimate levels of residual hexane and to ensure the safety of workers handling the oil. Some refineries use gas chromatography methods instead.

Free Fatty Acids in Crude and Refined Fats and Oils (*FFA*) *(Ca 5a-40)*: determines FFA, as oleic acid, by ethanolic sodium hydroxide titration.

Halphen Test for Detecting Cottonseed Oil (Cb 1–25): estimates the presence of cottonseed oil in vegetable or animal fats or oils as the result of a pink color formed between the reagent and cyclopropenoic fatty acids (sterculic and malvalic) normally present in cottonseed oil.

Hexane Residues in Fats and Oils (Ca 3b-87): determines, by gas chromatography, the "free" volatile hydrocarbons remaining in fats and oils after extraction with hydrocarbon solvents. The results are expressed in terms of hexane.

Insoluble Impurities in Fats and Oils (Ca 3a-46): determines dirt, meal, and foreign substances that are insoluble in kerosene and petroleum ether.

Iodine Value of Fats and Oils—Cyclohexane Method (IV) (Cd lb-87): measures the unsaturation of fats and oils in terms of centigrams of iodine absorbed per gram of sample. The method is applicable to all normal fats that do not contain conjugated double bonds. It often is used to estimate the degree of hydrogenation of oils.

Melting Point–Capillary Tube Method (Cc 1–25): determines the temperature at which a sample of fat in a closed capillary becomes completely clear and liquid; broadly applicable; popular for tropical fats.

Melting Point–Mettler Dropping Point (Cc 18–80): determines the temperature at which a sample becomes sufficiently fluid to flow in a specified apparatus; the major melting point determination method used by domestic industry.

Melting Point–Open Tube-Softening Point (Cc 3–25): determines the temperature at which a solidified fat, in an open capillary tube, softens sufficiently to slip and rise to the top of the heating bath. This method is applicable to fats such as coconut oil, stearin, hydrogenated fats, and hard tallows. The results sometimes are reported as the "melting slip point," but the method is different from the AOCS Slip Point (Cc 4–25).

Melting Point–Wiley Method (Cc-38): determines the temperature at which a sample disk of solidified fat assumes a spherical shape while suspended in a heating bath with an alcohol–water density gradient. A seldom-used method, primarily replaced by the Mettler Dropping Point.

Moisture-Distillation Method (Ca 2a-45): determines only moisture in triacylglycerols and emulsions by distillation with an immiscible solvent (toluene).

Moisture and Volatile Matter —Air Oven Method (M&V) (Ca 2c-25): determines the moisture and volatile matter by heating in a hot air oven. This method is applicable to animal and vegetable fats, but not to drying oils, coconut group fats, or oils with added monoacylglycerols.

Oil (Aa 4–38): determines oil content in a dried sample of oil-bearing material by extraction with petroleum ether. This method is specific for cottonseed, which first must be fumed with hydrochloric acid to prevent oil adsorption to the fiber. Additional methods exist for other oilseeds.

Oxygen Stability Index (OSI) (Cd 12b-92): measures the oxidation induction period of fat sample (essentially the time for a sample to exhaust its antioxidant properties) under conditions of the test.

Peroxide Value, Fats and Oils (PV) (Cd 8–53): determines all substances, in terms of milliequivalents of peroxide per 1000 g of sample that oxidize potassium iodide (KI). These substances generally are assumed to be peroxides or products of fat oxidation.

Phosphorus in Oils (Ca 13–55): estimates the phospholipid content of crude, degummed, and refined vegetable oils in terms of phosphorus. Refineries often use ICP spectrographs to analyze divalent cations rapidly in aspirated crude oil. The calcium and magnesium measured are mainly responsible for NHP and are determined directly. An AOCS method for analysis by ICP is being developed.

Refining Loss, Vegetable Oils, Crude (Ca 9a-52): determines the loss of free fatty acids and impurities when crude oils are refined under specified procedures.

Residual Lint (Aa 7–55): determines the lint content of cottonseed by fuming (digesting) with hydrochloric acid.

Saponification Value of Fats and Oils (Cd 3–25): determines the number of milligrams of (alcoholic) potassium hydroxide necessary to saponify a 1-g sample of a fat or oil.

Schaal Test (Schaal Oven Method): an accelerated test for determining the oxidative stability of a fat or a fat-containing food product. Results are reported as the time elapsed until a rancid odor is detected. This is not an AOCS method; see American Association of Cereal Chemists' Method Manual. Modifications, using OSI apparatus, have been reported.

Smoke, Flash, and Fire Point—Cleveland Open Cup Method (Cc 9a-48): determines the temperatures at which fats and oils smoke, flash, or burn. Smoke point determinations sometimes are used to follow degradation of frying oils with use.

Solid Fat Content of Fats and Oils by NMR (SFC) (Cd 16–81): estimates the percentage of solids in a semi-solid fat on the basis of the pulsed NMR signal of hydrogen in the liquid fraction. The method is used in the palm oil industry and widely throughout the world.

Solid Fat Index—Dilatometric Method (SFI) (Cd 10–57): estimates the percentage of solids in a semi-solid fat on the basis of changes in volume with temperature. This method utilizes glass dilatometers and is the primary method in the United States. *Totox Value:* an estimate of the degree of oxidation of a fat or oil, calculated as:

$$\text{Totox} = 2 \times (\text{PV}) + \text{AV}.$$

Triglycerides by GLC (Ce 5b-86): quantitatively determines triglycerides (triacylglycerols) in liquid vegetable oils in terms of molecular weight and degree of unsaturation as a function of their equivalent carbon number using high-pressure liquid chromatography.

Unsaponifiable Matter in Fats and Oils, Including Marine Oils (Ca bb-5, 3): determines substances dissolved in fats and oils that cannot be saponified (turned into sodium salts) by the usual caustic treatment, including higher aliphatic alcohols, sterols, pigments, and hydrocarbons. This method is not suitable for marine oils or feed grade fats.

References

1. Staff Time Line (1999) *Biotech Chronicles*, Access Excellence @ the National Health Museum, The National Health Museum, Washington, DC (www.accessexcellence.org/RC/AB/BC, downloaded, January 10, 2005)
2. Wiener N (1948) Cybernetics: or control and communication in the animal and the machine. Wiley, New York
3. Wiener N (1961) Cybernetics: or control and communication in the animal and the machine, 2nd edn. MIT, Cambridge
4. Wiener N, Schade JP (eds) (1965) Cybernetics of the nervous system. Elsevier, Amsterdam
5. Wiener N (1954) Human use of human beings: cybernetics and society. Houghton Mifflin, Boston
6. Wiener N (1967) Human use of human beings: cybernetics and society. Avon, New York
7. Liu Q, Singh SP, Green AG (2002) High-stearic and high-oleic cottonseed oils produced by hairpin RNA-mediated post transcriptional gene silencing. Plant Physiol 129:1732–1734
8. Scott WJ (1957) Water relations of food spoilage microorganisms. In: Mrak EM, Stewart GF (eds) Advances in food research, vol 7. Academic, New York, pp 83–127
9. USDA–ERS (2004) Genetically engineered crop varieties gain further acreage share in 2004. In: Amber waves. U.S. Department of Agriculture, Economic Research Service, Washington, DC, September, 2004
10. Clive J (2004) PREVIEW: global status of commercialized biotech/GM crops: 2004. In: ISAAA Briefs 32–2004, International Service for the Acquisition of Agri-Biotech Applications, c/o International Rice Research Institute, Metro, Manila
11. Mellon M, Rissler J (2004) Gone to seed: transgenic contaminants of the traditional seed supply. Union of Concerned Scientists, Cambridge
12. NRC (2004) Biological confinement of genetically engineered organisms. National Research Council, National Academies Press, Washington, DC
13. IM-NRC (2004) Safety of genetically engineered foods: approaches to assessing unintended health effects. Institute of Medicine and National Research Council, The National Academies Press, Washington, DC
14. NRC (1996) Ecologically based pest management: new solutions for a new century. National Research Council, National Academies Press, Washington, DC
15. Cooperative Extension Service (1987) Soybean production handbook, C-449 revised. Kansas State University, Manhattan
16. Stalcup L (2005) High-tech for success. Corn and soybean digest 65(2):12–13
17. Stalcup L (2005) Pinpoint accuracy. Corn and Soybean Digest 65 (2):19–21
18. Ege SN (1984) Organic chemistry. D. C. Heath, Lexington
19. National Research Council (1989) Recommended dietary allowances, 10th edn. National Academy, Washington, DC
20. Foreign Agricultural Service (2005) Oilseeds: world markets and trade, Circular Series FOP 1–05, U.S. Department of Agriculture, Washington, DC, January 2005
21. Johnson LA, Meyers DJ (1995) Industrial uses for soybeans. In: Erickson DR (ed) Practical handbook of soybean processing and utilization. AOCS, Champaign, pp 380–427
22. Methods AOCS (2009) Official methods and recommended practices of the AOCS. AOCS, Champaign
23. Schmidl MK, Labuza TP (eds) (2000) Essentials of functional foods. Aspen, Gaithersburg

24. AFCO (2004) Official Publication 2004. Association of American Feed Control Officials, Sharon Senesac, Assistant Secretary-Treasurer, Oxford
25. Hutt PB (2000) U.S Government regulation of food with claims for special physiological value. In: Schmidl MK, Labuza TP (eds) Essentials of functional foods. Aspen, Gaithersburg, pp 339–352
26. Council NR (1993) Nutrient requirements of fish. National Academy, Washington
27. Robertson JA, Morrison WH III, Wilson RL (1979) Effects of planting location and temperature on the oil content and fatty acid composition of sunflower seeds, USDA Agric. Results ARR-S-3. U.S. Dept. Agriculture, Washington, DC
28. Robertson JA (1975) Use of sunflower seed in food products. Crit Rev Food Sci Nutr 6:201–240
29. Pryde EH (1979) Fatty acids. AOCS, Champaign
30. Pryde EH, Princen LH, Mukherjee KD (1981) New sources of fats and oils. AOCS, Champaign
31. Cahn RS, Ingold CK, Prelog V (1956) The specification of asymmetric configuration in organic chemistry. Experientia 12:81–94
32. Hirschman H (1960) The nature of substrate asymmetry in stereoselective reactions. J Biol Chem 235:2762–2767
33. Nawar WW (1996) Lipids. In: Fennema OR (ed) Food chemistry, 3rd edn. Marcel Dekker, New York, pp 225–317
34. Beckman HJ (1983) Hydrogenation practice. J Am Oil Chem Soc 60:234A–242A
35. Sevanian A (1988) Lipid peroxidation in biological systems. AOCS, Champaign
36. Frankel EN (1988) Lipid oxidation. Oily, Dundee
37. Min DB, Smouse TH (eds) (1989) Flavor chemistry of lipid foods. AOCS, Champaign
38. Packer L, Traber MG, Xin W (1996) Proceedings of the international symposium on natural antioxidants molecular mechanisms and health effects. AOCS, Champaign
39. Min DB, Lee SH, Lee EC (1989) Singlet oxygen oxidation of vegetable oils. In: Min DB, Smouse TH (eds) Flavor chemistry of lipid foods. AOCS, Champaign, pp 57–97
40. Yurawecz MP, Mossoba MM, Kramer JKG, Pariza MW, Nelson GJ (eds) (1999) Advances in conjugated linoleic acid research, vol 1. AOCS, Champaign
41. Sébédio J-L, Christie WW, Adolf R (eds) (2003) Advances in conjugated linoleic acid research, vol. 2. AOCS, Champaign
42. Corl BA, Baumgard LH, Dwyer DA, Griinari JM, Phillips JM, Bauman DE (2001) The role of delta 9-desaturase in the production of Cis-9, Trans-11 CLA. J Nutr Biochem 12(11):622–630
43. Guyton AC, Hall JE (2000) Textbook of medical physiology, 10th edn. W. B. Saunders, New York
44. Lampert D (2000) Processes and products of interesterification. In: O'Brien RD, Farr WE, Wan PJ (eds) Introduction to fats and oils technology, 2nd edn. AOCS, Champaign, pp 208–234
45. Hagemann JW, Tallent WH, Barve JS, Ismail LA, Gunstone FD (1975) Polymorphism in single acid triglycerides of positional and geometric isomers of octadecanoic acid. J Am Oil Chem Soc 52:204–207
46. Rossell JB (1967) Phase diagrams of triglyceride systems. In: Paoletti R, Kritchevsky D (eds) Advances in lipid research, vol 5. Academic, New York, pp 353–408
47. Small DM (1986) Handbook of lipid research. The physical chemistry of lipids, vol 4. Plenum, New York
48. Bockish M (1998) Fats and oils handbook. AOCS, Champaign
49. Hagemann JW (1988) Thermal behavior and polymorphism of acylglycerides. In: Garti N, Sato K (eds) Crystallization and polymorphism of fats and fatty acids. Marcel Dekker, New York, pp 9–95
50. Hernqvst L (1988) Chocolate temper. In: Beckett ST (ed) Industrial chocolate manufacture and use. Avi–Van Nostrand Reinhold, New York, pp 153–171
51. Hernqvst L (1988) Crystal structures of fats and fatty acids. In: Garti N, Sato K (eds) Crystallization and polymorphism of fats and fatty acids. Marcel Dekker, New York, pp 97–137
52. Shipley CG (1986) X-ray crystallographic studies of aliphatic lipids. In: Small DM (ed) Handbook of lipid research, vol 4, The physical chemistry of lipids. Plenum, New York, pp 97–147
53. Sato K (1988) Crystallization of fats and fatty acids. In: Garti N, Sato K (eds) Crystallization and polymorphism of fats and fatty acids. Marcel Dekker, New York, pp 228–263
54. Widlak N, Hartel R, Narine S (eds) (2001) Crystallization and solidification properties of lipids. AOCS, Champaign
55. Poot C, Dijkshoorn W, Haigton AJ, Verburg CC (1975) Laboratory Separation of Crystals from Plastic Fats Using Detergent Solution. J Am Oil Chem Soc 52:69–72
56. Rivarola G, Anon MC, Calvelo A (1986) Crystallization of waxes during sunflowerseed oil refining. J Am Oil Chem Soc 62:1508–1513
57. Rivarola G, Anon MC, Calvelo A (1988) Crystallization of waxes in sunflowerseed oil. J Am Oil Chem Soc 65:1771–1773
58. Nichols PD, Nichols DS, Bakes MJ (1994) Marine oil products in Australia. Inform 5:254–261
59. Yankah VV, Jones PJJ (2001) Phytosterols and health implications: chemical nature and occurrence. Inform 12:808–811
60. Yankah VV, Jones PJJ (2001) Phytosterols and health implications: efficacy and nutritional aspects. Inform 12:899–903
61. Yankah VV, Jones PJJ (2001) Phytosterols and health implications: commercial products and their regulation. INFORM 12, 1011–1014, 1016
62. Gordon MH, Magos P (1983) The effect of sterols on the oxidation of edible oils. Food Chem 10:141–147
63. White PJ, Armstrong LS (1986) Effect of selected oat sterols on the deterioration of heated soybean oil. J Am Oil Chem Soc 63:525–529
64. Yan PS, White PJ (1990) Linalyl acetate and other compounds with related structures as antioxidants in heated soybean oil. J Sci Food Agric 38:1904–1908
65. Warner K, Su C, White PJ (2004) Role of antioxidants and polymerization inhibitors in protecting frying oils. In: Gupta MJ, Warner K, White PJ (eds) Frying technology and practices. AOCS, Champaign, pp 37–49
66. Cherry JP, Kramer WH (1989) Plant sources of lecithin. In: Szuhaj BF (ed) Lecithins: sources, manufacture & uses. AOCS, Champaign, pp 16–31
67. Foreign Agricultural Service (2005) Oilseeds: world markets and trade, FAS Circular Series FOP 4–05. U.S. Department of Agriculture, Washington, DC, April 2005
68. Rudbeck J (2001) Market report 2000. Render 30(2):10–15
69. Technical Committee–ISEO (1988) Food fats and oils. Institute of Shortening and Edible Oils, Washington, DC
70. USDA CNCPP (1998) Nutrition insight 5. Is total fat consumption really decreasing? USDA Center for Nutrition Policy and Promotion, U.S. Department of Agriculture, Washington, DC, April 1998
71. Unidentified (2000) State of the industry report. Snack Food & Wholesale Bakery, 89(6):SI-3–SI-74
72. A Brief History of Olestra (2002) Center for Science in the Public Interest, Washington, DC, www.cspinet.org/olestra/history.html). Downloaded, March 28, 2002
73. Martinez Suarez JM, Mendoza JA (1968) Olive oil processing and related aspects: extraction and refining: general aspect. In: Baldwin AR (ed) Proceedings world conference on emerging technologies in the fats and oils industry. AOCS, Champaign, pp 299–305
74. Boskou D (1996) Olive oil chemistry and technology. AOCS, Champaign

75. Firestone D, Fedeli E, Emmons EW (1996) Olive oil. In: Hui YH (ed) Bailey's industrial oil & fat products, vol 2, 5th edn, Edible oil & fat products: oils and oilseeds. Wiley, New York, pp 241–270
76. Canapi EC, Agustin YTV, Moro EA, Pedrosa E Jr, Bendaño MLJ (1996) Coconut oil. In: Hui YH (ed) Bailey's industrial oil & fat products, vol 2, 5th edn, Edible oil & fat products: oils and oilseeds. Wiley, New York, pp 225–254
77. Baisron Y (1996) Palm oil. In: Hui YH (ed) Bailey's industrial oil & fat products, vol 5, 5th edn, Industrial and consumer nonedible products from oils and fats. Wiley, New York, pp 271–375
78. Tuft LS (1996) Rendering. In: Hui YH (ed) Bailey's industrial oil & fat products, vol 5, 5th edn, Industrial and consumer nonedible products from oils and fats. Wiley, New York, pp 1–20
79. Wellons FH (1996) Industrial uses of rendered fats and oils. In: Hui YH (ed) Bailey's industrial oil & fat products, vol 5, 5th edn, Industrial and consumer nonedible products from oils and fats. Wiley, New York, pp 21–32
80. NRA (1993) Pocket information manual: buyer's guide to rendered products. National Renderers Association, Alexandria
81. Piette G, Hundt M, Jacques L, Lapoint M (2001) Influence of extraction temperature on amounts and quality of rendered chicken fat recovered from ground or homogenized skin. Poultry Sci 80:496–500
82. Søbstad GE (1990) Animal fats: the effect of process technology on fat quality. In: Erickson DR (ed) World conference proceedings: edible fats and oils processing: basic principles and modern practices. AOCS, Champaign, pp 31–36
83. Prokoph WH (1985) Rendering systems for processing animal by-product materials. J Am Oil Chem Soc 62:805–811
84. Love JA (1996) Animal fats. In: Hui YH (ed) Bailey's industrial oil & fat products, vol 1, 5th edn, Edible oil & fat products: general applications. Wiley, New York, pp 1–18
85. Schottelcotte W (1996) State of the art grease processing. Render 24(3):10, 11, 21
86. Lusas EW, Riaz MN (1996) Fats in feedstuffs and pet foods. In: Hui YH (ed) Bailey's industrial oil & fat products, vol 3, 5th edn, Edible oil & fat products: products and application technology. Wiley, New York, pp 255–329
87. Rudbeck J (2001) Market report 2000. Render 30(2):10–12, 13–15
88. Bimbo AP, Crowther JB (1992) Marine oils: fishing for industrial uses. Inform 3(988–996):998–1001
89. Pigott GM (1996) Marine oils. In: Hui YH (ed) Bailey's industrial oil & fat products, vol 3, 5th edn, Edible oil & fat products: products and application technology. Wiley, New York, pp 225–254
90. Burnham F (1989) Alaska's billion pounds of protein. Render 17 (5):8–12
91. Bimbo AP (1987) The emerging marine oil industry. J Am Oil Chem Soc 64(706):708–715
92. Moore HN (1983) Oilseed handling and preparation prior to solvent extraction. J Am Oil Chem Soc 60:141A–144A
93. Johnson LA, Lusas EW (1983) Comparison of alternative solvents for oils extraction. J Am Oil Chem Soc 60:181A–194A
94. Galvin JB (1997) Toxicity data for commercial hexane and hexane isomers. In: Wan PJ, Wakelyn PJ (eds) Technology and solvents for extracting oilseeds and nonpetroleum oils. AOCS, Champaign, pp 75–85
95. Wakelyn PJ (1997) Regulatory considerations for extraction solvents for oilseeds and other nonpetroleum oils. In: Wan PJ, Wakelyn PJ (eds) Technology and solvents for extracting oilseeds and nonpetroleum oils. AOCS, Champaign, pp 48–74
96. Technical Committee on Solvent Extraction Plants (2004) NFPA 36 standard for solvent extraction plants. NFPA International, Quincy
97. Sullivan DA, Campbell BD, Conway MF, Grimsby FN (1982) Isopropanol extraction of oilseeds. Oil Mill Gazetteer 87 (10):24–28
98. Baker EC, Sullivan DA (1983) Development of a pilot plant process for the extraction of soy flakes with aqueous isopropyl alcohol. J Am Oil Chem Soc 60:1271–1277
99. Lusas EW, Hernandez E (1997) Isopropyl alcohol. In: Wan PJ, Wakelyn PJ (eds) Technology and solvents for extracting oilseeds and nonpetroleum oils. AOCS, Champaign, pp 199–266
100. List GR, Freidrich JP (1989) Oxidative stability of seed oils extracted with supercritical carbon dioxide. J Am Oils Chem Soc 66:98–101
101. Liener IE (1994) Implications of antinutritional components in soybean foods. Crit Rev Food Sci Nutr 34(1):31–67
102. NOPA (2000) Yearbook & trading rules 2000–2001. National Oilseed Processors Association, Washington, DC
103. Farr WE (2000) Refining of fats and oils. In: O'Brien RD, Farr WE, Wan PJ (eds) Introduction to fats and oils technology, 2nd edn. AOCS, Champaign, pp 136–157
104. List GR (1988) Commercial manufacture of lecithin. In: Szuhaj BF (ed) Lecithins: sources, manufacture and uses. AOCS, Champaign, pp 145–161
105. Erickson DR (1995) Degumming and lecithin processing and utilization. In: Erickson DR (ed) Practical handbook of soybean processing and utilization. AOCS, Champaign, pp 174–183
106. van de Segers JC, Sande RLKM (1990) Degumming—theory and practice. In: Erickson DR (ed) World conference proceedings, edible fats and oils processing: basic principles and modern practices. AOCS, Champaign, pp 88–93
107. Dijkstra AJ (1993) Degumming, refining, washing and drying fats and oil. In: Applewhite TH (ed) Proceedings of the world conference on oilseed technology and utilization. AOCS, Champaign, pp 138–151
108. Dahlke K, Eichelsbacher M (1998) EnzyMax™ and Alcon™—Lurgi's route to physical refining. In: Koseoglu SS, Rhee KC, Wilson RF (eds) Emerging technologies, current practices, quality control, technology transfer, and environmental issues. AOCS, Champaign, pp 53–59
109. Erickson DR (1995) Neutralization. In: Erickson DR (ed) Practical handbook of soybean processing and utilization. AOCS, Champaign, pp 184–202
110. Hernandez, E (2001) Latest developments in refining and filtration of specialty oils with sodium silicates. Paper presented at American Oil Chemists' Society Annual Meeting, Minneapolis, May 13–17, 2001. Texas A&M University, College Station
111. Lin L, Rhee KC, Koseoglu SS (1998) Recent progress in membrane degumming of crude vegetable oils on a pilot-plant scale. In: Koseoglu SS, Rhee KC, Wilson RF (eds) Advances in oils and fats, antioxidants, and oilseed by-products. AOCS, Champaign, pp 76–82
112. Patterson HWB (1992) Bleaching and purifying fats and oils. AOCS, Champaign
113. Taylor DR (1993) Adsorptive purification. In: Applewhite TH (ed) Proceedings of the world conference on oilseed technology and utilization. AOCS, Champaign, pp 152–165
114. Zschau W (2000) Bleaching. In: O'Brien RD, Farr WE, Wan PJ (eds) Introduction to fats and oils technology, 2nd edn. AOCS, Champaign, pp 158–178
115. O'Brien RD (1995) Shortenings. In: Erickson DR (ed) Practical handbook of soybean processing and utilization. AOCS, Champaign, pp 363–379
116. List GR, Steidley KR, Palmquist D, Adlof RO (2001) Solid fat index vs. solid fat content: a comparison of dilatometry and pulsed nuclear magnetic resonance for solids in hydrogenated soybean oil. In: Widlak N, Hartel R, Narine S (eds) Crystallization and solidification properties of lipids. AOCS, Champaign, pp 146–152

117. O'Brien RD (2000) Liquid oils technology. In: O'Brien RD, Farr WE, Wan PJ (eds) Introduction to fats and oils technology, 2nd edn. AOCS, Champaign, pp 463–485
118. Tan BK, Flingoh CH (1981) PORIM Technol. Palm Oil Res. Inst. Malaysia No. 4
119. Tirtiaux A (1990) Dry fractionation: a technique and an art. In: Erickson DR (ed) World conference proceedings: edible fats and oils processing: basic principles and modern practices. AOCS, Champaign, pp 136–141
120. Kellens M, Hendrix M (2000) Fractionation. In: O'Brien RD, Farr WE, Wan PJ (eds) Introduction to fats and oils technology, 2nd edn. AOCS, Champaign, pp 194–209
121. Tirtiaux A, Gibon V (1998) Dry fractionation: the boost goes on. In: Koseoglu SS, Rhee KC, Wilson RF (eds) Emerging technologies, current practices, quality control, technology transfer, and environmental issues. AOCS, Champaign, pp 92–98
122. Davidson HF, Campbell EJ, Bell RJ, Pritchard R (1996) Sunflower oil. In: Hui YH (ed) Bailey's industrial oil & fat products, vol 2, 5th edn, Edible oil & fat products: oils and oilseeds. Wiley, New York, pp 603–689
123. Kaylegian KE, Lindsay RC (1995) Handbook of milkfat fractionation technology and applications. AOCS, Champaign
124. Hastert RC (1990) Cost/quality/health: the three pillars of hydrogenation, In: Erickson DR (ed) World conference proceedings: edible fats and oils processing: basic principle and modern practices. AOCS, Champaign, pp 142–152
125. Ariaansz RF (1993) Hydrogen theory. In: Applewhite TH (ed) Proceedings of the world conference on oilseed technology and utilization. AOCS, Champaign, pp 166–172
126. Kokken MJ (1993) Hydrogenation in practice. In: Applewhite TH (ed) Proceedings of the world conference on oilseed technology and utilization. AOCS, Champaign, pp 173–179
127. Patterson HWB (1994) Hydrogenation of fats and oils. AOCS, Champaign
128. Erickson DR, Erickson MD (1995) Hydrogenation and base stock formulation. In: Erickson DR (ed) Practical handbook of soybean processing and utilization. AOCS, Champaign, pp 218–238
129. Ariaansz RF, Okonek DV (1998) Isomer control during edible oil processing. In: Koseoglu SS, Rhee KC, Wilson RF (eds) Emerging technologies, current practices, quality control, technology transfer, and environmental issues. AOCS, Champaign, pp 77–91
130. Hastert RC (2000) Hydrogenation. In: O'Brien RD, Farr WE, Wan PJ (eds) Introduction to fats and oils technology, 2nd edn. AOCS, Champaign, pp 179–193
131. Okonek D (2001) Hydrogenated oil end products. In: Hernandez E, Rathbone SJ, Rhee KC (eds) Vegetable oils processing practical short course manual. Texas A&M University, College Station
132. O'Brien RD (2000) Shortening technology. In: O'Brien RD, Farr WE, Wan PJ (eds) Fats and oils technology, 2nd edn. AOCS, Champaign, pp 179–193
133. Going LH (1969) Interesterification products and processes. J Am Oil Chem Soc 44:414A, 416A, 418A, 420A, 422A, 454A–456A
134. Akoh CC (1998) Reduced-calorie fats and fat replacement. In: Koseoglu SS, Rhee KC, Wilson RF (eds) Emerging technologies, current practices, quality control, technology transfer, and environmental issues. AOCS, Champaign, pp 223–234
135. Fomuso LB, Akoh CC (1997) Enzymatic modification of triolein: incorporation, of caproic and butyric acids to produce reduced-calorie structured lipids. J Am Oil Chem Soc 3:269–272
136. Sonntag NV (1982) Fat splitting, esterification, and interesterification. In: Swern D (ed) Bailey's industrial oil and fat products, vol 2, 4th edn. Wiley, New York, pp 97–173
137. Hernandez E, Lusas EW (1997) Trends in transesterification of cottonseed oil. Food Technol 51(5):72–76
138. Sreenivasan B (1978) Interesterification of fats. J Am Oil Chem Soc 55:796–805
139. Rozendaal A (1990) Interesterification of oils and fats. In: Erickson DR (ed) World conference proceedings: edible fats and oils processing: basic principles and modern practices. AOCS, Champaign, pp 152–157
140. Kubow S (1998) Triglyceride structure and lipoprotein metabolism. In: Christophe AB (ed) Structural modified food fats: synthesis, biochemistry, and use. AOCS, Champaign, pp 149–159
141. Kritchevsky D (1998) Triglyceride structure and atherosclerosis. In: Christophe AB (ed) Structural modified food fats: synthesis, biochemistry, and use. AOCS, Champaign, pp 183–188
142. Würtz M, Pearce SW (2004) Enzymes as Processing Aids. In: Targut NT, Dunford HB (eds) Nutritionally enhanced edible oil and oilseed processing. AOCS, Champaign, pp 264–278
143. Kellens M, De Greyt W (2000) Deodorization. In: O'Brien RD, Farr WE, Wan PJ (eds) Introduction to fats and oils technology, 2nd edn. AOCS, Champaign, pp 235–268
144. Zehnder CT (1995) Deodorization. In: Erickson DR (ed) Practical handbook of soybean processing and utilization. AOCS, Champaign, pp 239–257
145. Stenberg O, Sjöberg P (1998) A novel deodorization plant design for seed oils. In: Koseoglu SS, Rhee KC, Wilson RF (eds) Emerging technologies, current practices, quality control, technology transfer, and environmental issues. AOCS, Champaign, pp 103–108
146. Hymowitz T (1970) On the domestication of the soybean. Economic Botany 24:408–421
147. Hymowitz T, Newell CA (1981) Taxonomy of the genus glycine, domestication and uses of soybeans. Economic Botany 35:272–288
148. Hymowitz T, Harlan JR (1983) Introduction of soybean to North America by Samuel Bowen in 1765. Economic Botany 37:371–379
149. Hymowitz T (1987) Introduction of the soybean to Illinois. Economic Botany 41:28–32
150. Britannica CD (1999) Encyclopedia Britannica CD 99 multimedia edition. Encyclopedia Britannica, Chicago, IL
151. Bowdidge E (1935) The soya bean: its history, cultivation (in England) and uses. Oxford University Press, London
152. Markley KS, Goss WH (1944) Soybean chemistry and technology. Chemical Publishing Company, Brooklyn
153. Staff (1977) Continuous solvent extraction: the early beginnings of a giant industry. J Am Oil Chem Soc 54:202A–207A
154. Govier GE (1928) European methods of crushing soya beans. Oil and Fat Industries 5:251–254
155. Burtis EL (1950) World soybean production and trade. In: Markley KL (ed) Soybeans and soybean products, vol 1. Interscience, New York, pp 61–117
156. Markley KS (ed) (1951) Soybeans and soybean products, vols. 1 and 2, Interscience, New York
157. Goss WH (1947) The German oilseed industry. Hobart, Washington, DC
158. Jones LA, King CC (1996) Cottonseed oil. In: Hui YH (ed) Bailey's industrial oil & fat products, vol 2, 5th edn, Edible oil & fat products: oils and oilseeds. Wiley, New York, pp 159–240
159. Michael Eskin NA, McDonald BE, Przybylski R, Malcolmson LJ, Scarth LJ, Mag T, Ward K, Adolph D (1996) Canola oil. In: Hui YH (ed) Bailey's industrial oil & fat products, vol Edible oil & fat products: oils and oilseeds, 5th edn, 2. Wiley, New York, pp 1–97
160. Young CT (1996) Peanut oil. In: Hui YH (ed) Bailey's industrial oil & fat products, vol 2, 5th edn, Edible oil & fat products: oils and oilseeds. Wiley, New York, pp 377–392
161. FDA (1981) Revealing trans fats. FDA Consumer Magazine, Pub. No. FDAAO4-1329C, U.S. Food and Drug Administration, Washington, DC, September–October, 2003
162. CFSAN-FDA (2004) Questions and answers about *Trans* fat nutritional labeling. CFSAN Office of Nutritional Products,

Labeling and Dietary Supplements, U.S. Food and Drug Administration, Washington, DC, Last updated June 24, 2004

163. CSAN-FDA (Transcript) Nutrition subcommittee meeting on total fat and trans fat, Loews L'Enfant Plaza Hotel, April 27, 2004, Washington, DC, Center for Food Safety and Applied Nutrition, US Food and Drug Administration, Washington, DC. Downloaded (www.fda.gov/ohrms/dockets/ac/04/transcripts/ 4035t1.doc), March 9, 2005
164. Jung OK, Ju JW, Choi DS, Yoon SH, Jung MY (2002) CLA formation in oils during hydrogenation process as affected by catalyst types, catalyst contents, hydrogen pressure, and oil species. J Am Oil Chem Soc 79:501–510
165. Jung MY, Jung MO (2002) Identification of conjugated linolenic acids in hydrogenated soybean oil by impregnated HPLC and gas chromatography-ion impacted mass spectrometry of their 4-D-dimethyloxazoline derivatives. J Agric Food Chem 50:6188–6193
166. Jang ES, Jung MY, Min DB (2005) Hydrogenation for low *Trans* and high conjugated fatty acids. Comprehensive Reviews in Food Science and Food Safety 1:22–30
167. Gregory JF III (1996) Vitamins. In: Fennema OR (ed) Food chemistry, 3rd edn. Marcel Dekker, New York, pp 531–617
168. Packer L, Traber MG, Xin W (eds) (1996) Proceedings of the international symposium on natural antioxidants: molecular mechanisms and health effects. AOCS, Champaign
169. Shahidi F (ed) (1997) Natural antioxidants: chemistry, health effects, and applications. AOCS, Champaign
170. Becher P, Schick M (1987) J. Microemulsions. In: Schick MJ (ed) Nonionic surfactants: physical chemistry. Marcel Dekker, New York, pp 435–491
171. Artz WE (1990) Emulsifiers. In: Branen AL, Davidson PM, Salminen S (eds) Food additives. Marcel Dekker, New York, pp 347–417
172. O'Brien RD (1998) Fats and oils: formulating and processing for applications. Technomic, Boca Raton
173. Fritsch CW (1981) Measurements of frying fat deterioration. J Am Oil Chem Soc 58:272–274
174. Perkins EG, Erickson MD (1996) Deep frying: chemistry, nutrition, and practical applications. AOCS, Champaign
175. Banks DE, Lusas EW (2001) Oils and industrial frying. In: Lusas EW, Rooney LW (eds) Snack foods processing. Technomic, Boca Raton, pp 137–204
176. Evans CD, Warner K, List GR, Cowan JC (1972) Room evaluation of oils and cooking fats. J Am Oil Chem Soc 49:578–582
177. Moulton KJ, Beal RE, Warner K, Bounty KB (1975) Flavor evaluation of copper-nickel hydrogenated soybean oil and blends with Unhydrogenated Oil. J Am Oil Chem Soc 52:469–472
178. Mounts TL, Warner K, List GR (1994) Performance evaluation of hexane-extracted oils from genetically-modified soybeans. J Am Oil Chem Soc 71:16–161
179. Tompkins C, Perkins EG (2000) Frying performance of low-linolenic soybean oils. J Am Oil Chem Soc 72:223–229
180. Firestone D (1996) Regulation of frying fat and oil. In: Perkins EG, Erickson MD (eds) Deep frying: chemistry, nutrition, and practical applications. AOCS, Champaign, pp 323–334
181. Gupta MK, Warner K, White PJ (eds) (2004) Frying technology and practices. AOCS, Champaign
182. Boice B, Egbert R, Sikorski DM, Stuchell YM, Widlak N (Inventors) Fods and drinks containing diglycerol. U.S. Patent Application 20,040,009,284, Jan 15, 2004
183. Warner K (2004).The effect of expeller pressed/physically refined soy oil on flavor of French fried potatoes and frying oil stability. In: Presented at AOCS Annual Meeting, May 9–12, 2004, Cincinnati, OH. USDA/ARS National Center for Agricultural Utilization Research. Peoria
184. Private communication. W. B. Wijeratne, Research Director, InstaPro International Corp., Des Moines, IA, January 3, 2005
185. Zu S, Riaz MN, Lusas EW (1996) Effect of different extrusion temperatures and moisture content on lipoxygenase inactivation and protein solubility in Soybeans. J Agric Fd Chem 44:3315–3318
186. Cheng V (2004) Frying leap pays off. Oil Mill Gazetteer 110 (3):16–17
187. Garti N, Sato K (eds) (1988) Crystallization and polymorphism of fats and fatty acids. Marcel Dekker, New York
188. Bumbalough J (2000) Margarine types and preparation technology. In: O'Brien RD, Farr WE, Wan PJ (eds) Fats and oils technology, 2nd edn. AOCS, Champaign, pp 452–462
189. Wainwright B (2000) Specialty fats and oils. In: O'Brien RD, Farr WE, Wan PJ (eds) Introduction to fats and oils technology, 2nd edn. AOCS, Champaign, pp 486–514
190. Johnson RW, Fritz E (eds) (1988) Fatty acids in industry. Marcel Dekker, New York
191. Gunstone FD, Hamilton RJ (eds) (2000) Oleochemical manufacture and applications. CRC, Boca Raton
192. Zoebelein H (1992) Renewable resources for the chemical industry. Inform 3:721–725
193. Johnsen PB (1998) Emerging industrial uses for soybeans. In: Koseoglu SS, Rhee KC, Wilson RF (eds) Emerging technologies, current practices, quality control, technology transfer, and environmental issues. AOCS, Champaign, pp 109–112
194. Derksen JTP, Cuperus FP (1998) Processing of oilseed constituents for nonfood applications. In: Koseoglu SS, Rhee KC, Wilson RF (eds) Emerging technologies, current practices, quality control technology transfer, and environmental issues. AOCS, Champaign, pp 113–120
195. Foglia TA, Nelson LA, Marmer WN, Knothe GH, Dunn RO, Bagby MO (1998) Improving the properties of vegetable oils and fats for use as biodiesel. In: Koseoglu SS, Rhee KC, Wilson RF (eds) Emerging technologies, current practices, quality control, technology transfer, and environmental issues. AOCS, Champaign, pp 121–125
196. Krawczyk T (1996) Biodiesel. INFORM 7, 800–801, 804–808 810–824
197. Lee I, Johnson LA, Hammond EG (1996) Reducing the crystallization temperature of biodiesel by winterizing methyl soyate. J Am Oil Chem Soc 73:631–636
198. Lee I, Johnson LA, Hammond EG (1995) Use of branched-chain esters to reduce the crystallization temperature of biodiesel. J Am Oil Chem Soc 72:1155–1160
199. Caparella T (2004) Biodiesel Bulletin: USDA awards biodiesel grants, proposes increased buying of biobased products. Render: The National Magazine of Rendering 33(1):12–15
200. Talley D (2004) Biodiesel: a compelling business for the rendering industry. Ibid. 33(1):16–21
201. Caparella T (2004) Biodiesel production mixed bag for renderers. Ibid. 33(1):22
202. Pearl GG (2004) Evidence shows fats safe to use in biofuels. Ibid. 33(1):23
203. Howell S (2005) Time to take the biodiesel plunge? Ibid. 34 (1):10–15
204. Caparella T (2005) Biodiesel bulletin: government grants help fuel industry. Ibid. 34(1):18–19, 21
205. Pearl GG (2005) Biodiesel is on the move! Ibid. 33(1):22–23, 25
206. Sundram J, Thiagarajan T, Gapor A, Basiron Y (2002) Palm tocotrienols: new antioxidants for the new millennium. Inform 13:634–640

Sugar and Other Sweeteners

35

Mary An Godshall

Introduction

Sugar and starch are among the most abundant plant products available, and large industries exist worldwide to extract and process them from agricultural sources. The world production of sugar (sucrose from cane and beet) in 2007/2008 was 170.4 million metric tons, raw value [1], with 20.7% being beet sugar and 79.3% cane sugar. The proportion of beet sugar to cane sugar has fallen steadily since about 1971, when it constituted 42.8% of total sugar production [2]. The decline in beet sugar proportion represents not so much a decline in beet production, which has remained in the range of 33–39 million metric tons, but rather a continued increase in cane sugar production from around 72 million metric tons in 1991 to 134 million metric tons in 2008. The total production of world sugar has risen dramatically since 1971/1972, when it was 71.7 million tons [3].

Although cane and beet are physically very different from one another, the factory processes for extracting and producing sugar are similar. Two major differences are the types of color produced during beet processing compared to cane processing, which allows direct processing of beets into white sugar; and the fact that cane factories are essentially self-sufficient in fuel because they can burn bagasse, the cane plant residue, whereas the beet factory does not produce a consumable by-product fuel.

Sugar is primarily used as a sweetener in the manufacture of food and in household use. When used for such purposes in industrialized nations, most of it is highly refined or purified, but considerable quantities are consumed in less industrialized areas of the world as a less refined sugar, called plantation white sugar or mill white sugar. In India, crude sugar products such as gur (also known as jaggery), khandsari, and cane juice are popular and represent as much as 29% of consumption [4]. Sugar is used to a limited extent in the production of other chemicals, such as sucrose esters, and in the form of by-product molasses, as a substrate for fermentation processes and to produce alcohol, both for consumption and power, usually mixed with gasoline.

Starch is widely used in the textile and paper industries, as well as for food. A major use of starch is its hydrolytic conversion to glucose and enzymatic conversion to fructose for use as a sweetener in the food industry. Technological developments and governmental price-support programs for the domestic (United States) cane and beet sugar industries begun in the mid-1970s greatly magnified the importance of starch as a source of nutritive sweeteners [5]. Starch-derived sweeteners compete directly and successfully with sugar in the United States, and increasingly in other countries.

The sugar industry has been characterized by steady, small incremental improvements in technology and production over the years, with subsequent improvements across the process, from field practices to new products [6, 7].

Sugar

History [8]

The ancestry of sugarcane and its use as a food have been traced to the island of New Guinea. An interesting legend by Hugill [9] associated sugarcane with the origin of the human race. Around 8000 BC, the plant started on its migration from New Guinea to many areas of southeastern Asia, Indonesia, the Philippines, Malay, Indochina, and eastern India, with humans probably acting as its main dispersal agent [10]. Sugarcane was first cultivated in Bengalese India as a field crop and the juice manufactured into various solid forms. A general knowledge of sugar was prevalent throughout India by 400 BC. By the tenth century AD, sugar cultivation and manufacturing had become important industries in Persia and Egypt. The early Islamic movement spread knowledge of the sugar industry throughout the Mediterranean area.

M. An Godshall (✉)
Sugar Processing Research Institute, Inc., New Orleans, LA, USA
e-mail: MA.Godshall@cox.net

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_35, © Springer Science+Business Media New York 2012

On the second voyage of Columbus to America, in 1493, sugarcane was introduced in Santo Domingo. It spread rapidly through the West Indies and Central America. Cortez brought cane to Mexico, and Pizarro introduced it in Peru. By 1600, the sugar industry was the largest industry in tropical America [11].

The modern sugar industry dates from the end of the eighteenth century, when steam replaced animal energy and made possible the development of larger and more efficient production units. The vacuum pan appeared in 1813, bag filters in 1824, multiple-effect evaporators in 1846, filter presses in 1850, centrifugals in 1867, dryers in 1878, and packaging machines in 1891 [12]. Cultivation of the sugar beet plant and the manufacture of sugar from the beet developed in the industrial nations of Europe during the eighteenth and nineteenth centuries. In 1747, the German chemist, Marggraf, established that sugar from beets was the same as sugar from cane. His pupil, Achard, in 1799, demonstrated that sugar can be commercially prepared from beets. During the Napoleonic wars, a short-lived beet sugar industry was established in France [13]. Today, beet is the major source of sugar in Europe.

Raw Sugar Production

Traditionally, raw sugar is produced at the cane mill and refined sugar is produced at a refinery, usually far removed from the source of the raw sugar, and close to areas of consumption. Figure 35.1 is a generalized flow diagram for the manufacture of raw cane sugar. In summary, the juice is extracted from the harvested cane stalks, the juice is purified (clarified), concentrated by vacuum evaporation, and crystallized. The several steps are described in the sections that follow.

Agriculture. The sugarcane is a large perennial tropical grass belonging to the genus *Saccharum*. Three basic species, *S. officinarum*, *S. robustum*, and *S. spontaneum*, make up the sugarcane complex, with a large number of varieties. Sugarcane is propagated commercially by cuttings, each cutting consisting of portions of the cane plant having two or more buds, or "eyes." The buds sprout into shoots from which several other shoots arise below the soil level to form a clump of stalks known as a "stool." From 12 to 20 months are required for the crop to mature from new plantings and about 12 months for ratoon crops (i.e., cane stalks arising from stools that have been previously harvested). Most of the field operations have been mechanized, but planting is still done by hand in some cane-producing areas. Fields are replanted after two to five or more ratoons (cuttings made from the original planting). As cane grows, its foliage is largely limited to the upper one-third of the plant, as the lower leaves die off from shading as the stalk grows taller. The cane stalk is round and jointed, and grows to more than 10 ft when mature (Fig. 35.2). It is covered with a hard rind that varies in color from light brown, green, yellowish green, purple, or striped, depending on the variety. There is usually a fine film of waxy material on the surface of the rind, heavier in some varieties than others. The stalk is divided into a series of internodes, from 3 to 10 in. in length, separated by nodes. The internodes contain the pithy

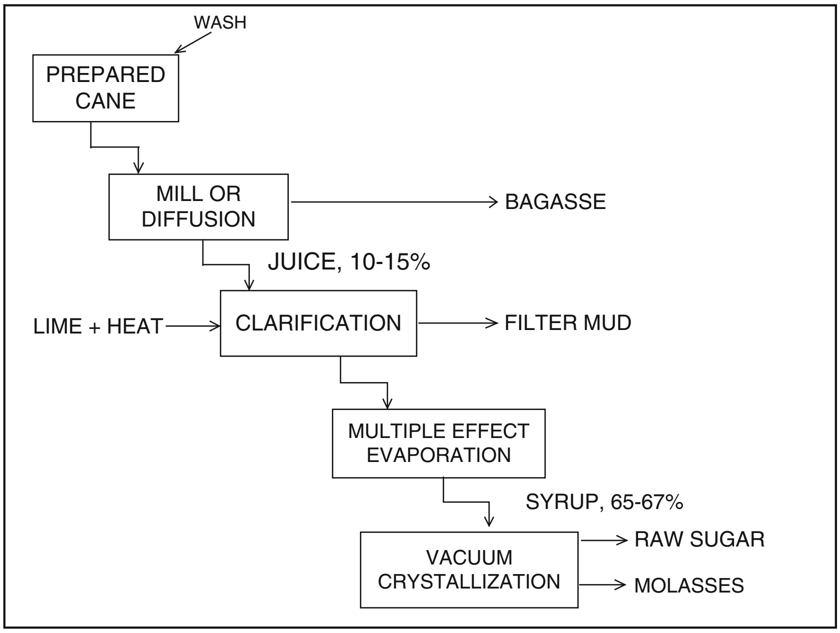


Fig. 35.1 Generalized flow diagram of steps in raw cane sugar manufacture (adapted from Clarke [104])

Fig. 35.2 Sugarcane in the field

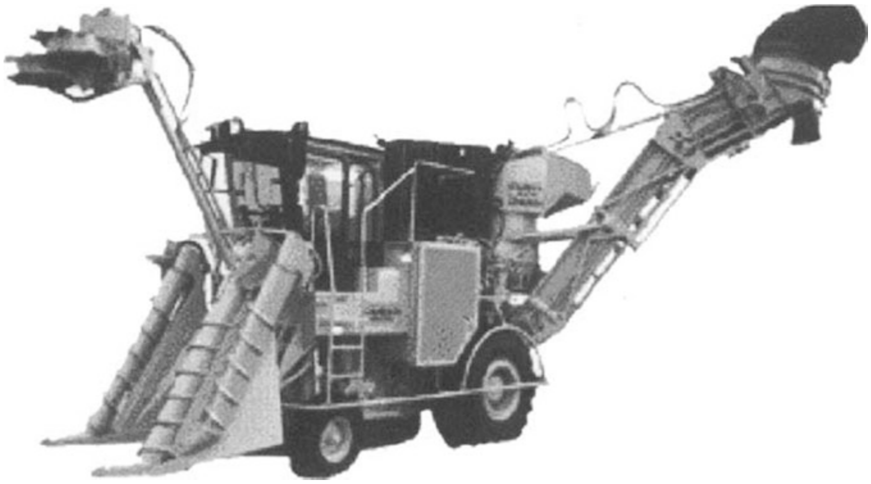


Fig. 35.3 Cameco single-row harvester

material from which most of the sucrose-containing cane juice is obtained. The nodes, somewhat woody in nature, contain one bud or "eye" on alternating sides, which sprout when planted.

Breeding programs throughout the world develop cane varieties that are suited to their particular climatic conditions, to give the highest yield, balanced with disease resistance and other desirable agronomic traits. In Louisiana, for example, cane varieties are bred for cold tolerance, among other traits, because of the short growing season and the possibility of freezing weather during the harvest. Varieties retain vigor for 10–15 years and then need to be replaced by new varieties. There is ongoing research in genetic manipulation of sugarcane for herbicide and disease resistance, but no commercial varieties exist to date.

Harvesting. Sugarcane is cut by hand with machete-type knives in a few producing areas. In hand-cutting, the tops and leaves are removed, and the cane is cut at ground level. In more industrialized areas, cane is harvested by machines. Two types of machine harvesting are in widespread use: whole-stalk harvesters, also known as soldier harvesters, and chopper harvesters, also known as combine harvesters, in which the cane is cut into smaller pieces, known as billets, about 8–12 in. in length. Combine harvesters exist as single-row (Fig. 35.3) and two-row

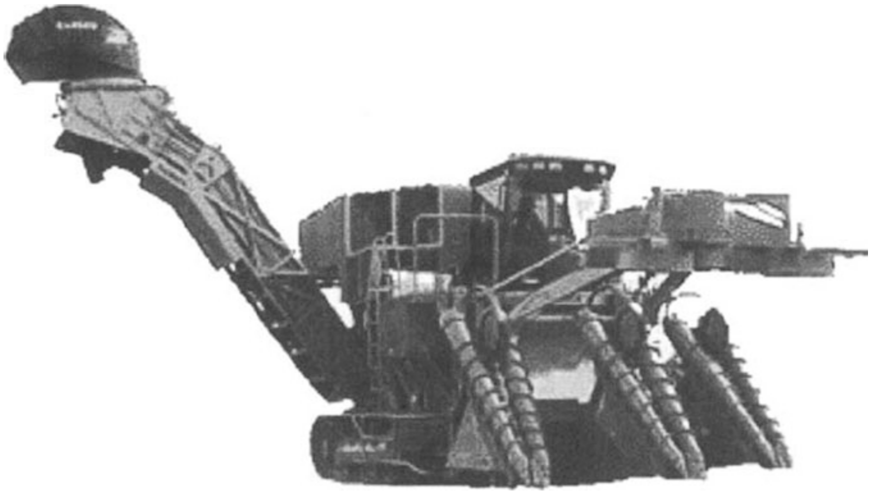


Fig. 35.4 Cameco two-row harvester

models (Fig. 35.4). The whole stalk cane may be transported immediately to the mill where it is dumped in large cane yards for future processing, or left in the field for future pick-up. Billeted cane should always be processed as soon as possible because of the faster deterioration rate of the juice quality caused by the shorter pieces. In many areas, cane is burned before harvesting to help remove leaves and trash. The practice of burning has become less environmentally and socially acceptable in recent years, and, in spite of many efforts by the industry to control adverse effects, the practice is in decline. The practice of harvesting unburned cane is called green cane harvesting, and it introduces a host of challenges to the processor because of the heavier load of impurities, such as color-producing molecules and polysaccharides, that are expressed into the juice [14, 15]. Cane may be transported to the mill by railcars, trucks, trailers, and carts.

A new development in harvest control for better quality of cane and beet has been the use of global positioning systems (GPS) based on satellite navigation facilities [16].

Preparation of cane for milling. Because of the presence of adhering soil, rocks, and field trash, mechanically harvested cane usually must be washed before milling. In some carefully controlled situations, when the weather has been dry, washing can be eliminated. Washing systems range from a simple spraying with warm water on the carrier or cane table to an elaborate system consisting of conveyors with water jets, stripping rolls, and baths for the removal of stones. In most mills, a biocide is added to prevent bacterial growth in the cane. Mill sanitation is important to control bacterial activity, which can result in direct sucrose loss as bacteria consume the sugar, and in the production of undesirable by-products, such as dextran, a polysaccharide that has deleterious effects in the production of sugar, or lactic acid. With the advent of billeted cane, more sucrose can be lost in the wash water due to the multitude of cut ends, resulting in some reduction in yield.

The cane is prepared for milling by mechanically knifing or shredding or a combination of the two to chop the cane into smaller pieces. It is important for the mill to pay attention to good cane preparation, in as much as this is one of the most important parameters affecting extraction of the juice [17]. The Unigator, developed in the late 1970s, combines the action of the chopper and the shredder into a single unit [18]. The advantages gained are a greater increase in juice extraction by the mills, an increase in the grinding rate, and, in some cases, a decrease in power consumption by the milling operation. The preparation index (PI) is a measure of how well the cane has been prepared for juice extraction, and represents the percentage of cane plant cells that are ruptured by the cane preparation equipment. It is measured as brix (percent of sucrose) in the ruptured cells as a proportion of the total brix in the cane.

A novel approach to the preparation of sugarcane for the extraction of sugar was demonstrated in a machine called a Tilby separator. Cane stalks, cut into 8–12-in. segments, are fed into the separator, split in half, and depithed by a depithing wheel, resulting in a clean, sucrose-containing pith and rejected rind segments (Fig. 35.5). This approach has not been adopted on any large scale, but is used in some tropical countries to produce edible cane juice and also has some uses for research purposes.

Juice extraction. The shredded cane passes through a series of three horizontal rollers (mills) arranged in a triangular pattern, with the top roll rotating counterclockwise and the bottom two rollers clockwise. A series of three-roller mills, numbering 3–7, is called a tandem. The pressure on the top roll is regulated by hydraulic rams and averages about 300 t. Below each mill is a juice pan into which expressed juice flows. The crusher and first mill extract 60–70% of the cane

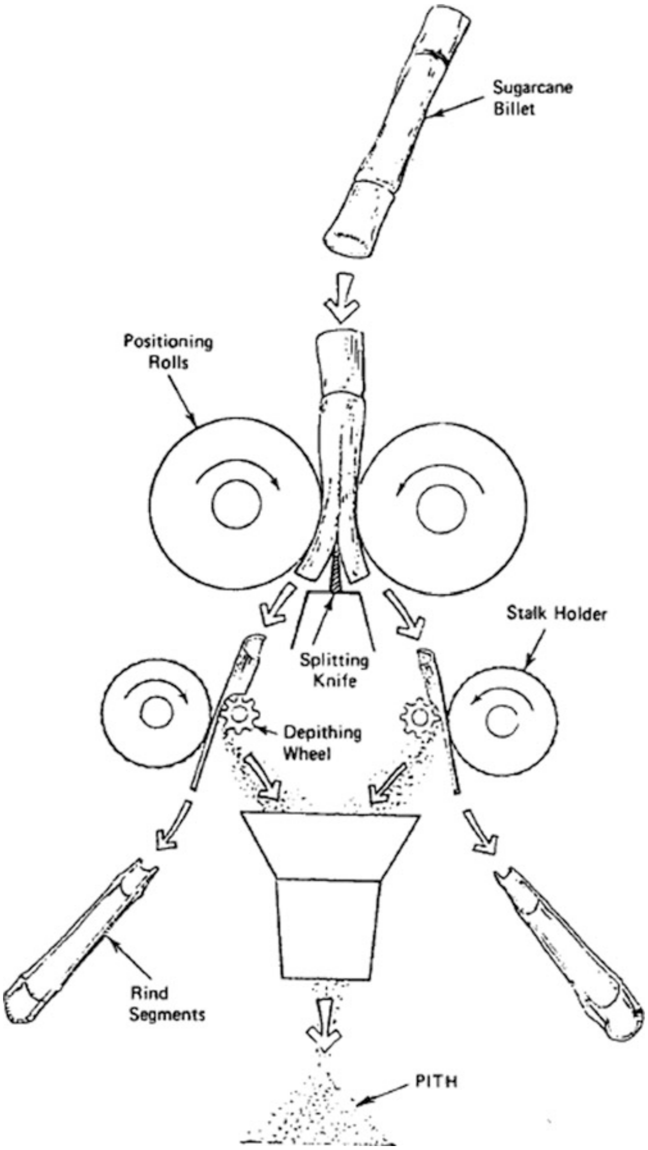


Fig. 35.5 Tilby separator process (Courtesy Ander-Cane, Inc., Naples, FL)

juice, and the remaining mills take out 22–25%. The fibrous cane residue remaining after juice is extracted is called bagasse. When the fiber content of the bagasse reaches about 50%, extraction by conventional milling approaches zero. Because the juice remaining with the fiber contains the same proportion of sucrose as the original cane juice, the yield of sugar would be substantially lowered if extraction were terminated at this point. Consequently, a process called "compound imbibition" is used to reduce the sucrose in the bagasse by repeated dilution and milling. In a five-mill tandem, water is added to the fourth mill, and the expressed juice from that mill is brought back to the second mill. The expressed juice from the third mill is recirculated to the first mill and the fifth mill to the third mill. In this way, the juice in the bagasse is always diluted ahead of each mill.

The amount of imbibition applied at each mill is approximately equal to the amount of water applied to the fourth mill or to the penultimate mill in a differently numbered tandem. The expressed juice is screened through perforated metal screens with 1-mm diameter openings. Additional screening may be provided by stationary or vibrating metal cloths.

The capacity of cane factories can vary from as little as 1,000 t of cane per day (tcd) up to 25,000 tcd. Factories operate only during the harvest, which may last for only 3 months or all year long, depending on the location.

Diffusion. Diffusion represents an alternative way to obtain juice from the cane. In diffusion, the cane is carefully shredded to disrupt the cells for maximum expression of juice. Diffusers have capacities ranging from 200 to 500 t of cane per hour (newer ones have capacities up to 875 t/h) and may be used alone or in conjunction with part of a milling tandem, the crusher and the first mill. Preparation of the cane for diffusion is essential for good extraction. In the diffuser, the shredded cane is countercurrently washed with imbibition water at an average temperature of 85°C. The high temperature helps to prevent microbial growth in the cane juice, but also increases the color of the juice. The last stage of a diffuser, the bagasse exit, receives water that gradually increases in sugar content as it proceeds to the final stage. Most of the sugar is extracted in the first four or five stages by simple displacement of sucrose from the ruptured cane cells. In later stages, diffusion appears to take place in unruptured cane cells. One or two mills are used to express water from the bagasse after discharge from the diffuser [19]. Chainless diffusers were introduced by Bosch Projects in 2006. The chainless diffuser eliminates the need for chains to move the cane, consisting instead of modules that move independently, actuated by hydraulic cylinders [20].

In comparing diffusers and milling, it is observed that sucrose extraction by a diffusion system averages 97–98%, whereas milling will produce 95–97% extraction. Diffusion provides a clearer juice (due to the filtering effect of the bagasse mat) with lower starch and gums but with about 25% higher color. Power requirements for diffusers are less than for mills and less filter cake is produced, as the mat of shredded cane in the diffuser acts as a primary filter, but more steam is needed to heat the juice [21, 22]. Cane diffusers have been installed in cane mills throughout the world and are manufactured by a number of companies. Many large cane mill installations have diffusers, the major exception being in the United States, where only one mill has a diffuser.

Biocides. Microbial inhibitors, or biocides, mentioned briefly above, are important for control in both beet and cane processing, as the juice contains sucrose and other nutrients in the right concentration for promoting microbial growth. The biocides in most common use around the world are dithiocarbamates, glutaraldehyde, and ammonium bisulfite. Combinations of these may act synergistically [23]. The use of natural products as antibacterials has received attention in recent years. Hop acids, rosin acids, and palm kernel fatty acids, especially myristic acid, have shown promise [24, 25]. They are recommended to be used in mixtures in rotation to prevent adaptation by microorganisms. These have the advantage that they are biodegradable and used in human food. Additionally, myristic acid is precipitated by lime during clarification [25].

Dextran. The major microbiological losses in sugar processing are caused by the common soil bacterium *Leuconostoc mesenerioides*, which uses sucrose as a food source, producing a long-chain polysaccharide, known as dextran, as a waste product. Dextran is composed of linear α-1,4-glucose with branches of α-1,6-glucose. In high enough concentration, dextran causes gummy solutions and difficulties in processing, along with sucrose loss. It also interferes with the polarimetric analysis of sucrose because, having a positive rotation, it is interpreted as sucrose, resulting in false high estimates of the sugar content in the cane or beet juice. Various methods exist for dextran measurement, all with some limitations.

Juice purification (*clarification*). The first stage of purifying the juice, designed to remove the maximum amount of impurities, is called clarification, or defection, and employs lime and heat. Clarification affects the juice filterability, evaporator scale composition, sucrose crystallization, and the quality of raw sugar produced [26]. The cane juice obtained from milling or diffusion is acidic (pH around 5.5–5.7), turbid, and dark brown in color. It contains 12–18% by weight sucrose (depending on the maturity of the cane), in addition to soluble and insoluble impurities, such as soil, protein, waxes, polysaccharides, starch, fine bagasse (bagacillo), organic and phenolic acids, soluble salts, and pigments. Cane juice also contains from 0.5 to 3% reducing sugars (glucose and fructose in approximately equal quantities). Lime serves the dual purpose of raising the pH and stabilizing the juice against hydrolysis of the sucrose in the acid juice and forming a precipitate to remove impurities. The heat helps to disinfect the juice of harmful bacteria and improves coagulation and precipitation of the impurities.

In simple lime defection, lime is added to the juice to raise the pH to 6.7–7.5 and to react with inorganic phosphate present in the cane juice to form a calcium phosphate floc, which entraps insoluble impurities from the cane juice, as it settles. The lime is added either in the form of milk of lime, lime-juice mixtures, or lime saccharate. Lime addition may be batchwise or continuous, and the addition of the lime may be to cold juice (35–40°C), to intermediate juice (72–76°C), or to hot juice (100°C) [26]. Polyacrylamide flocculants are added to aid in the coagulation of the precipitate.

In cold liming, milk of lime is added to the cold juice, and the limed juice is then pumped through heaters in which the juice is heated to 90–115°C. Many modifications to this process exist, especially with respect to the heating of the juice. Fractional liming with double heating involves liming the cold juice to pH 6.4, heating, liming to pH 7.6, and heating again before sending the juice to the clarifiers. Studies have shown the advantage of hot or intermediate liming over cold liming, resulting in less sucrose loss through inversion, less dextran formation, and better control of pH and turbidity with intermediate and hot liming [27, 28].

The combination of lime and heat forms a flocculent precipitate with various components in the juice, consisting mostly of insoluble lime salts, coagulated protein, and entrapped colloidal and suspended matter. The precipitate is removed by sedimentation or settling in continuous closed-tray clarifiers. The juice leaving the clarifier is a clear brown liquid. The flocculent precipitates that settle on the clarifier trays are called “muds.” They contain about 5% solid matter. Entrained sugar is recovered from the mud by means of rotary vacuum filters equipped with a perforated metallic screen cloth. The turbid filtrate is returned to the clarification system, and the press cake or filter mud is usually sent to the fields as fertilizer. Good clarification depends upon the formation of a stable flocculent precipitate that settles rapidly. Juices low in natural phosphates (less than 300 ppm P_2O_5 on juice solids) tend to be difficult to clarify and phosphate may need to be added. Very fine silty soil carried in from the field may also interfere with clarification. The main change in clarifiers over the last 20 years has been the development of the trayless SRI-type clarifier, which can handle much higher throughputs than the older-type clarifiers [29].

Syrup clarification. In some factories, a process known as syrup clarification is used to produce a higher-quality raw sugar and better yield. Evaporator syrup is treated with phosphoric acid, lime, and a polymer flocculent; insoluble impurities are floated to the surface of the syrup with aeration and removed by skimming. The clarified syrup goes to the vacuum pan for crystallization. Syrup clarification reduces the color and turbidity of the sugar produced, and can also be used in conjunction with sulfitation to produce an off-white, edible sugar [30].

Sulfitation. Sulfitation (alternate spelling, sulphitation) is the practice of adding sulfur dioxide (SO_2) to cane juice and syrup for the purpose of color control. In many parts of the world, sulfur dioxide is used to process cane juice into a semi-refined white sugar, known as plantation sugar or mill white sugar, used for domestic consumption (see the section "Direct Consumption Sugars"). Sulfitation is not generally used in the manufacture of raw sugar, which is produced for export to refineries, and is never used in the US cane sugar industry. Sulfur dioxide gas is produced by burning elemental sulfur. The gas is added to the cane juice during clarification, in the presence of lime, and many of the complex, little understood reactions that produce color in cane juice are inhibited by the presence of SO_2 gas. For example, enzymes responsible for the enzymatic browning reaction are inhibited, and reducing sugars (glucose and fructose) are blocked by sulfur addition at the aldehyde group and cannot react with amino acids to form browning polymers in the Maillard reaction.

Double sulfitation is practiced in many mills in India and China. In this process, a second dosage of sulfite and lime is added to the evaporator syrup at a solids content of 55–65% [31]. Double sulfitation provides a better quality white sugar.

There are many variations for the addition of SO_2 to juice, based on temperature, order of addition, and pH levels. These include liming first, sulfiting fist, simultaneous addition of lime and sulfite, and fractional procedures, as well as sulfiting cold or hot or stepwise heating of juice. If sulfitation is done first, because it lowers the pH of the juice, it is important to add the lime (about 0.20% on juice) in as short a time as possible to prevent sugar loss [31, 32]. The advantage of using SO_2 is a much lower color in the produced sugar. Disadvantages include increased scaling problems in the evaporators, high sulfite content in the final product, higher sugar loss due to lower pH, and secondary color formation (color return) on storage. Recently, the use of liquid SO_2 has been shown to have advantages over the gas, especially in a significant reduction in gas consumption (up to 69%) and a better quality of product [33]. There is a trend to decrease the use of SO_2 in the cane sugar industry in favor of alternative color control methods, such as polyaluminum coagulants and hydrogen peroxide.

Evaporation. The clarified juice (about 85% water) is pumped to evaporators where it is concentrated to a clear heavy syrup containing about 65% solids. Evaporation is carried out in multiple-effect evaporators in order to achieve maximum steam economy. Each effect is arranged in series and operated so that each succeeding one operates under a higher vacuum (lower pressure). This arrangement allows the juice to be drawn from one vessel to the next and permits it to boil at progressively lower temperatures. The concentrated juice (final evaporator syrup) is removed from the last effect by a pump. Triple, quadruple-, and quintuple-effect evaporators are used, with the quadruple-effect type plus a preevaporator being a common configuration. In a four-effect (quadruple) evaporator, 1 lb of steam evaporates 4 lb of water.

Several types of evaporators exist [34]. The older, more traditional, evaporators are the Roberts and the Kestner, both rising film, tubular evaporators. The Roberts, first introduced in the 1800s, is known as a short tube, natural circulation, vertical tube evaporator. The tubes, inside which the evaporation takes place, are in the range of 1.5–3 m in length. The Kestner evaporator consists of numerous long vertical tubes, 6–7.5 m long, inside a cylindrical shell. In both, the juice to be concentrated is fed to the bottom of the tubes and heated, causing the juice to boil and rise through the tubes. At the top, concentrated juice and vapor are separated in separators. The Kestner evaporator has a short residence time and a high heat transfer coefficient, but is more subject to scale deposits. The Roberts evaporator does not have a juice/vapor separator.

Newer types of evaporators are the falling film tubular evaporator and the plate evaporator, either rising film or falling film configuration. The falling film tubular evaporators have long tubes, like the Kestner, but the juice is fed at the top and flows down the tube wall as a thin film, resulting in higher heat transfer coefficients and reduced contact times. However, the tubes require feeding equal quantities of juice and must be kept wet at all times, as even a short time without adequate wetting can lead to severe fouling. Advances and innovations in evaporation are driven by the need for energy efficiency and to improve heat transfer coefficients. The goal in evaporation is to move the liquid through as quickly as possible at the highest temperature without causing degradation of the sucrose. The newest of the evaporators is the falling film plate evaporator (e.g., Balcke–Dür, Niro, Inc., GEA Ecoflex GmbH), the first one being installed in 1992. Recently, technology has become available that allows the retrofitting of existing Robert evaporators with falling film plate packs and some peripheral equipment, allowing the total steam demand of a cane sugar factory to be decreased by almost 30% on cane [35]. Plate evaporators are more common in the beet industry.

Evaporator scaling. A severe problem in cane sugar processing is the deposition of scale on the evaporator tubes or plates. Scale deposits consist of calcium and silicate salts, enmeshed in a matrix of organic material that can accumulate very rapidly, causing a severe drop-off in heat transfer and efficiency of evaporation [36, 37]. Cane mills are forced to clean their evaporators on the order of every 10–15 days, using mixtures of acid and alkali and mechanical cleaning. Various antiscalants are available, with some, though not a great deal, of efficacy. Evaporator scaling is a problem in cane sugar refining and beet sugar production as well, but not to the extent that it is in raw sugar production. Good control of clarification will help to minimize evaporator scale.

Crystallization. The final evaporator syrup is pumped to a vacuum pan, in which it is evaporated to supersaturation to cause sugar crystallization. The vacuum pan is a single-effect evaporator designed to handle viscous materials. It is a vertical cylinder with its bottom designed to allow easy removal of the crystallized mass. The heating elements used in vacuum pans either are short, large-diameter vertical tubes (calandria pans) or coil around the inner surface of the truncated cone of the pan (coil pans). A typical vacuum pan with a "catchall" or entrainment separator for separating syrup from vapors measures approximately 25 ft in height and 15 ft in diameter. The working capacity is about 1,000 ft^3 of massecuite (mixture of crystals and syrup or mother liquor). The shape of the pan and the positioning of the heating elements within the pan are important design factors in maintaining good circulation of the massecuite. For example, floating calandrias (calandria not attached to the shell of the pan), horizontal pans, and pans having mechanical circulators are used. The process of crystallization is called boiling or pan boiling in the industry.

The move toward continuous processing has resulted in the development of continuous vacuum pans, and low-grade continuous pan boiling is well established [38], and raw sugar boiling with continuous pans is increasing as well. The benefits of continuous pan boiling include control simplicity, improved product quality, low product losses, steam savings, steadier demand, and energy savings [39]. Most continuous pans are a variation of the original 1931 Werkspoor patent, of a horizontal cylindrical pan divided by 8–18 stationary partitions or compartments, each heated by steam. Continuous vacuum pans are used in both the cane and beet sugar industries.

It should be noted that batch pans are still required for the production of high-quality refined sugar. One version of the continuous vacuum pan is a horizontal cylinder with compartments in its lower part. The pan is provided with an additional evaporator called the concentrator, where the density of the syrup is raised to 78–80° Brix.[1]

Seed is added in the first compartment, and the resulting massecuite moves progressively through the compartments of the pan. Additional syrup is added to each compartment to control the fluidity of the massecuite. Approximately 20% of the syrup is introduced at the concentrator, and 80% of the syrup is fed to the various compartments through special feed headers.

The crystallization of sucrose in vacuum pans is called "sugar boiling" and each boiling is termed a "strike." Because a single crystallization does not recover all the sucrose from the syrup, mother liquor from a strike is recycled for recovery of additional sugar. There may be as few as two and as many as four strikes to recover the maximum amount of sugar. A three-boiling system is the most popular. The strikes are designated by letters. Thus, the first strike is the A boiling, and the products of the A massecuite after centrifugation are A sugar and A molasses. The boiling systems are based on "purity," the amount of sucrose remaining in solution, and usually are expressed as the ratio of the polarization value to the total solids as measured by a Brix refractometer. In a three-boiling system, the purity of the A massecuite is set between 80 and 85 by blending syrup with A molasses; the purity of the B massecuite is set between 70 and 75 by blending with syrup and A molasses; the purity of the C massecuite is set between 55 and 60 by blending with syrup and B molasses. Other variations are possible.

The A and B massecuites, after being discharged from a vacuum pan, are sent to centrifugals for separation. The A and B sugars are combined to become commercial raw sugar, the principal product of a cane sugar factory. The C massecuite is a low-purity, highly viscous material that is not immediately sent to the centrifugals because of the large amount of recoverable sucrose remaining in solution. Instead, it is placed in crystallizers, U-shaped horizontal containers equipped with coils attached to a hollow rotating shaft through which water circulates. The massecuite remains in the crystallizer from 1 to 4 days to allow additional crystallization to take place. Centrifugation of the C massecuite yields a final molasses (blackstrap molasses) and a C sugar that is used for seeding the A and B boilings.

Centrifugation. Massecuite from the vacuum pans is sent to centrifugal machines in which the crystals are separated from the mother liquor. A centrifugal consists of a cylindrical perforated basket lined with a screen of perforated sheet metal. The basket, enclosed in a metal casing, is mounted on a vertical shaft that rotates the basket, imparting centrifugal force to the massecuite. In a batch centrifugal, hot massecuite is fed into the basket through a short chute from the holding vessel. As the basket rotates, the massecuite forms a vertical layer on the screen lining. When the machine reaches operating speed (1,000–1,800 rpm), the syrup flows through the perforations lining the basket and is removed through an outlet at the bottom of the casing. The sugar on the lining is washed with a spray of water to decrease the amount of molasses adhering to the crystals. The basket continues to rotate until the sugar is fairly dry, at which time the machine is switched off and brakes applied. The sugar is discharged by an automatic plough.

In continuous centrifugals, the machines do not stop, but continue in motion while receiving fresh supplies of

---

[1] The Brix scale is a density scale for sugar (sucrose) solutions. The degrees Brix are numerically equal to the percentage of sucrose in solution (wt/wt). The term Brix solids refers to the solids in solution as determined by a refractometer.

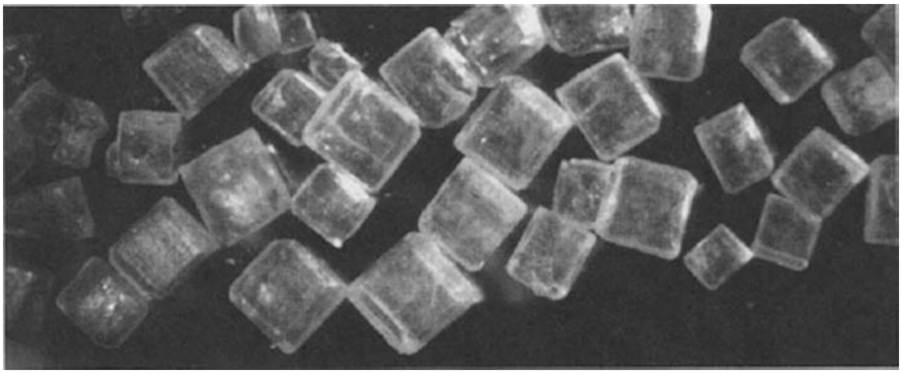


Fig. 35.6 Plantation sugar crystals, showing the ideal shape of sugar crystals

massecuites. The rate of feed must be carefully regulated to obtain the optimum separation of molasses from crystals, while at the same time not abrading or breaking too many crystals.

Raw sugar. The final product of the mill is raw sugar, about 98.0–98.5 purity, a bulk commodity shipped around the world to be refined by importing countries. All raw sugar produced in the United States is refined within the United States; none is shipped offshore. The quality of raw sugar is determined by its polarization (i.e., purity), color, moisture, ash, dextran, invert (i.e., glucose and fructose content), crystal uniformity, and proportion of color on the crystal surface. These parameters determine the commercial price of raw sugar.

Control. Factory control is based on temperature, pH, and polarization. Cane is tested by various means as it comes to the mill to determine its characteristics and to determine the amount to pay the farmer. Control of microbial infection, discussed above, is important. The starch content of juice from immature cane can be significant (e.g., in Louisiana and South Africa), and amylase enzyme may be added continuously to the clarified juice to reduce it. Starch will cause severe viscosity problems if not removed. If there is a problem with dextran, dextranase enzyme can be added, but it is expensive and not as effective as desired [40].

Molasses exhaustion. Molasses contains a significant amount of sucrose and its maximum exhaustibility (removal of sucrose) has always been a priority for the sugar industry. Well-exhausted cane molasses contains about 33–35% sucrose. Many factors contribute to poorly exhausted molasses, besides inefficiency of process, and these include the impurities present in the cane juice. Many of these impurities, such as polysaccharides, ash components, and organic acids are melassigenic (i.e., they increase the solubility of sucrose), making sucrose more difficult to crystallize, and therefore going preferentially into the molasses [41].

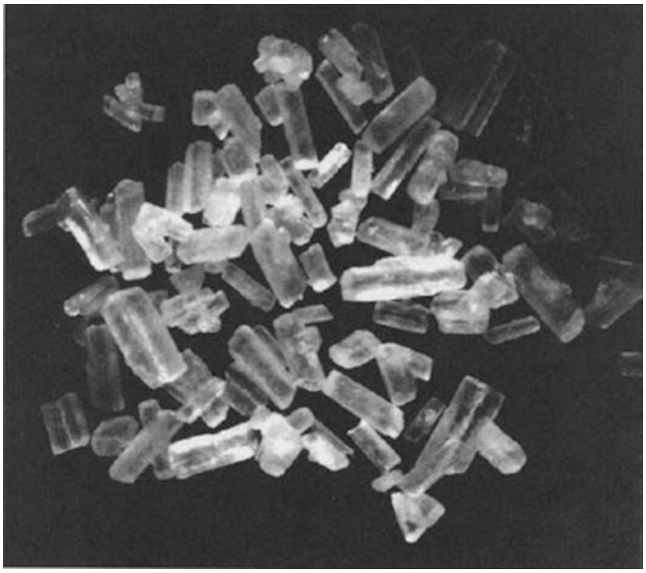


Fig. 35.7 Raw sugar crystals elongated in the presence of dextran and/or oligosaccharides

Crystal Quality

Sucrose forms anhydrous monoclinic crystals of the sphenoidic class [42]. In appearance, the ideal sugar crystal is a slightly asymmetrical square with 8–14 faces or planes (Fig. 35.6). Certain impurities in the mother liquor can affect the growth of the sugar crystal and change its morphology. Figure 35.7 shows the elongating effect that dextran and certain oligosaccharides have on the crystal shape. Sucrose also can form twinned crystals. Figure 35.8 shows refined sugar crystals.

Bagasse

Bagasse is the fibrous portion of the cane stalk that remains after extraction of the cane juice. When discharged from the milling train, bagasse contains about 50% by weight of water.

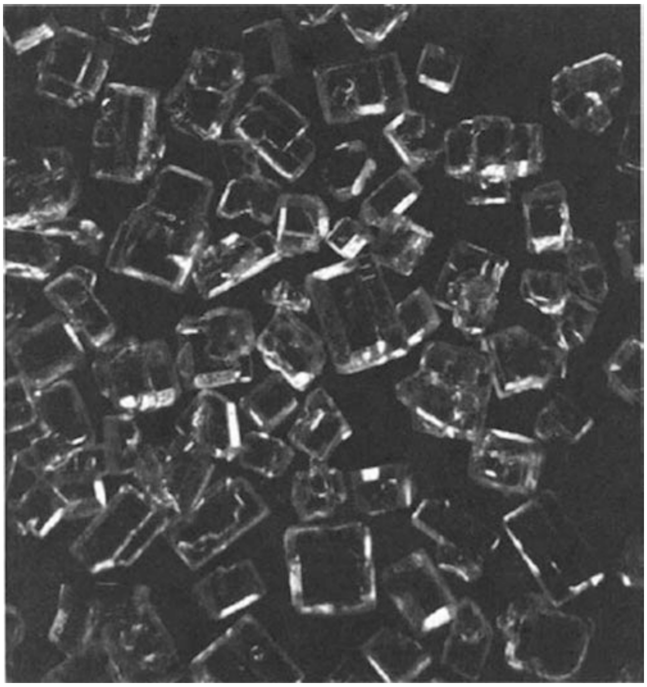


Fig. 35.8 Refined sugar crystals

The average fuel value of ash-free dry bagasse is 8,300 BTU/lb (gross calorific value 19,400 kJ/kg dry, ash free). It is a valuable by-product because it is used to fire the boilers in cane mills, making cane factories essentially energy-sufficient. Sometimes excess energy is produced which can be sold to the power companies. A sugar mill that processes one million tons of sugarcane per year produces in bagasse fuel the equivalent of about 420,000 barrels of imported crude oil, more than enough to operate the sugar mill and co-generate electricity [43]. Excess bagasse is used for paperboard and wallboard and as a chemical feedstock for conversion to furfural [44]. It is made into a number of paper products in China [45].

Direct Consumption Sugar

Many cane sugar mills in tropical countries produce white sugar by using combinations of sulfitation, syrup clarification, or carbonation. These edible sugars are known as mill white sugar or plantation white sugar. When compared to refined sugar, mill sugars are higher in color, ash, moisture, and reducing sugars. Because sulfur is used in production, the finished product may have sulfite in it.

Talodura. The Talodura process, used in syrup clarification, consists of using a combination of sulfite clarification followed by subsequent clarification of the evaporated liquor with lime and phosphoric acid to which a flocculant is added. It has led to improvements in the color of the mill white sugar, boiling house recovery, and chemical costs per ton of sugar [46].

Blanco Directo. Blanco Directo is a trademark of Tate & Lyle Process Technology. The process and similar processes (as described above) are used at sugarcane factories to produce a high-grade white sugar suitable for use in food and beverages. This product is economically competitive with imported refined sugars. Although the term “Blanco Directo” is the trade name for one particular process, it appears to be used somewhat generically for similar processes.

Carbonation. In this process the mixed juices from the mills are heated, clarified with lime, and evaporated to about 35° Brix. The syrup is relimed and treated with carbon dioxide, filtered, recarbonated, reheated, and refiltered. After carbonation, the syrup is given a double sulfur treatment and filtered. The resulting syrup is subjected to a three- or four-boiling system with the A and B sugars used as the white sugar product.

Edible Products from the Mill

Raw sugar is not considered an edible food in the United States, but in many other countries it is. There has been a growing trend to produce food-grade sugar at the cane mill (see discussion about plantation white sugar above). The growing societal interest in “unrefined” or “natural” products has led to a huge growth industry for products from sugar mills [47]. These include such products as demerara, Barbados, muscovado, and turbinado sugar, which are various grades of brown/raw sugar. Turbinado sugar is raw sugar that has been washed (centrifuged in the presence of steam) to remove some of the colored molasses coating on the surface of the crystal, resulting in a light golden sugar with large crystals and a mild cane taste.

Organic sugar. The largest area of growth for edible mill products is organic sugar. This sector is growing at 25–30% per year [48]. There are many regulations and restrictions in place for a mill to be certified to produce organic sugar; these deal with land preparation and the use of herbicides, fertilizers, and chemicals. The yields from organic production are decreased 30–60% from conventional sugar production, resulting in a product which is about four times more expensive than refined sugar. Brazil is the largest producer of organic sugar.

New Technology

In recent years, a number of new technologies have been explored for the production of raw sugar, refined sugar, and

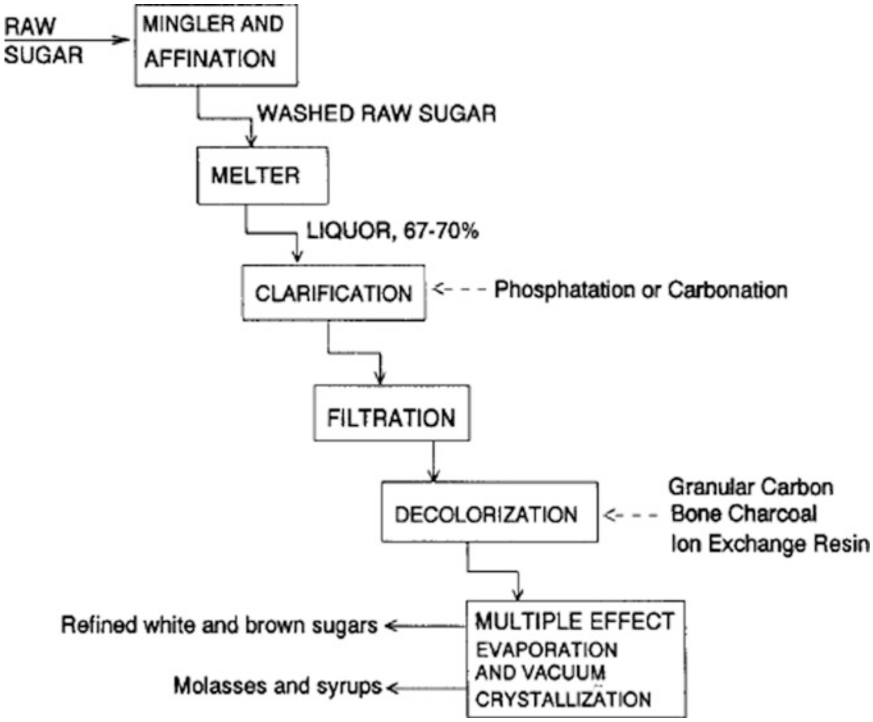


Fig. 35.9 General flow diagram of sugar refining (adapted from Clarke [104])

beet sugar, including ultra- and membrane filtration, ion-exchange, juice softening (the use of weakly cationic resin to remove calcium [49]), and chromatographic methods, which are, in the words of one author, "beginning to blur the demarcation of steps like clarification" [50]. These techniques are not limited to raw sugar production but are also being explored in the beet industry, although less so in cane refining. The goals in the use of these technologies are to reduce or eliminate the use of lime and/or sulfur, to produce a better quality sugar, or to produce refined sugar directly from cane juice at the mill.

Membrane filtration technology has not yet proven economically feasible in the industry, but continued research may eventually lead to viable commercial products [51, 52]. Carousel chromatographic systems, such as the continuous ion-exchange separation (ISEP) and continuous chromatographic separation (CSEP) systems, offer some promise in the sugar industry. They have been used successfully in other industries, and may be tailored to obtain a number of different fractions or products [53, 54].

Cane Sugar Refining

The raw sugar produced at the factory is light brown in color and contains about 98% sucrose. As described above, in most instances this product is not considered an edible food, but rather is a bulk commodity. The process of purifying raw sugar to make a white sugar is called refining, and the sugar produced is called refined sugar. Refining traditionally involves several unit stages, as shown in Fig. 35.9. Because of the increased quality of the raw sugar produced in some areas, not all the stages are used, and various combinations of processes are used at different refineries.

Affination and Melting

The first step in refining raw sugar is called affination (or washing), and consists of removing the film of molasses, which contains a large portion of the impurities in the sugar, from the surface of the raw sugar crystal. The raw sugar is mixed with saturated sugar syrup (72–75° Brix) at about 50°C in a U-shaped trough called a mingler. Saturated syrup is used to avoid dissolving the raw sugar. The mingler has a rotating agitator to maintain maximum contact between the sugar and the syrup. The mixture is centrifuged to separate the crystals from the syrup, and the crystals are then washed with a small spray of hot water or steam. The "washed" or "affined" sugar is "melted" or dissolved in water to a density of 55–60° Brix. Potable water, steam condensate, and "sweet waters" are used for dissolution. The liquor from the melter is screened to remove insoluble matter, such as sand, stones, wood, cane fibers, and lint. Screening is done by vibrating screens, cyclonic separators, or centrifugal screens.

Today, most refiners have eliminated the affination stage by buying very high-purity, low-color raw sugar, in effect, requiring the raw sugar mill to do the affination at the mill.

Purification

The melted liquor is then purified by any of several combinations that include clarification and decolorization. In clarification, the liquor is treated chemically to remove suspended matter, colloids, and other impurities that contribute to turbidity. Clarified liquor is optically clear, but is still highly colored. Decolorization procedures include treating the clarified liquor with various filtration media, such as bone char, granular activated carbon (GAC), or ion exchange resin, either alone or in some combination. It is necessary to clarify the liquor before the filtration medium because the particulate impurities will quickly "blind" or foul the filtration material and render it useless.

Clarification. All clarification treatments include the addition of lime. Lime increases the pH, stabilizing the sucrose against hydrolysis at high temperatures, and the calcium forms insoluble precipitates with many of the impurities in the raw liquor. Simple liming is rarely used in refining. The important clarification processes are phosphatation and carbonatation, which include combinations of lime with either phosphoric acid or carbon dioxide.

Phosphatation. In phosphatation, the screened liquor is heated to 60–70°C and mixed with phosphoric acid (0.005–0.025% P_2O_5 on solids). The mixture is immediately limed to pH 7.0–8.0, aerated with compressed air, and sent to a clarifier, a tank equipped with heating coils. The liquor enters the clarifier at one end and is heated to 88°C while flowing to the outlet at the opposite end. A flocculent precipitate of calcium phosphate forms, reacting with or entrapping colloids, organic acids, some coloring matter, and suspended particles. Air flotation is used to separate the precipitate from the liquor, causing it to rise to the top of the clarifier as a blanket of scum, which is skimmed off by moving paddles. (This process is sometimes known as phosflotation.) About 25–30% of the color is removed in this process.

Talofloc process (*color precipitation*). The addition of flocculants and surfactants has greatly improved the simple phosphatation process. The Talofloc process (developed by Tate & Lyle) consists of adding a cationic surfactant (Talofloc) to the raw melt liquor in concentrations ranging from 0.02 to 0.07% on solids. This is followed by the addition of lime and phosphoric acid and the incorporation of air. Just before the liquor enters the clarifying chamber, a polyacrylamide flocculant (Taloflote) is added that induces aggregation of the calcium phosphate precipitate, causing a faster and more complete separation of the impurities. As much as 60% of the color can be removed by this process. Other similar additives are commercially available.

Carbonatation. In carbonatation (sometimes also known as carbonation), the melter liquor, heated to about 60–80°C, is limed to about pH 10 (0.4–0.8% CaO on solids), gassed with carbon dioxide, heated to about 85°C, and regassed until the pH drops to between 7.7 and 8.5. Because of the two-stage gassing, carbonatation requires two clarifiers, a primary and a secondary. Washed flue gases are the source of carbon dioxide. The carbonated liquor resides in tanks called saturators for about 3 h to allow growth of the calcium carbonate. The calcium carbonate precipitate that forms entraps some colored matter, colloids, organic acids, and inorganic compounds. The precipitate, known as mud, is removed by filtration, such as Sweetland presses. Carbonatation alone achieves about 50% decolorization and, with the addition of polycationic additives, can achieve as much as 70% [55].

There are continuing discussions about the merits of one type of clarification over the other, phosphatation vs. carbonatation. In summary, the consensus appears to be that carbonatation reliably removes more color and colloidal material and is more flexible in handling different types of raw sugar, but is more expensive to operate, generates more waste, and may have slightly higher levels of sucrose loss to molasses than phosphatation. Environmental restrictions on disposal of phosphate-containing waste are a consideration for phosphatation refineries.

Filtration. The liquor from phosphatation and carbonatation contains small amounts of finely dispersed particulate matter that require filtration for removal. The filtration process is similar for both types of clarified liquor. A precoat of filter aid (diatomaceous earth or perlite) is first deposited on the filter surfaces of the press. Additional filter aid is added to phosphate-clarified liquors to improve press runs. Liquor at 70–85°C is fed to the press at a pressure of about 60 psig until the flow rate drops below a predetermined level, ending the filtration cycle. The calcium carbonate particles in carbonated liquor act as a fairly good filter aid medium, as long as the particles are developed to optimum size during carbonatation. Large quantities of sugar polysaccharides (starch, dextran, and others) in the melt liquor will restrict the growth of calcium carbonate, requiring the use of diatomaceous earth. A filter cycle may last from 2 to 12 h depending upon the quality of the feed liquor.

New developments. At least one refinery has, in recent years, been able to omit all chemical clarification by using only very high-purity, low-color raw sugar. Clarification is accomplished by a two-stage coarse filtration, followed by a finer filter-aid filtration. Again, this development is possible only because of the high quality of the incoming raw sugar.

Scums and muds. Both phosphatation and carbonatation produce scums or muds that have a moderate level of entrained sucrose, which must be recovered. Various schemes exist for removing the sugar, such as secondary clarification of phosphatation scums and sluicing of carbonation muds with dewatering on rotary vacuum filters.

Decolorization. Filtered clarified liquor is a clear brown liquid having a solids content between 55 and 65%, pH of 6.7–7.2, and temperature between 65 and 85°C. The next stage of refining involves removing this color to produce a clear, nearly colorless solution from which white sugar can be crystallized. During decolorization, the liquor is passed through one or a combination of decolorizing adsorbents, such as bone char, GAC, or ion-exchange resins. Bone char, consisting of the sintered long bones of cattle, is composed of calcium hydroxyapatite (80–85%) in a carbon matrix. It removes colorants, colloidal material, and a considerable amount of the ash components. GAC by itself only removes color; it is usually treated with magnesium oxide in order to buffer the sugar liquor and may remove a small amount of ash. Ion exchange resins absorb color and change the composition of the ash.

Bone char and granular carbon generally are used in cylindrical columns 20–25 ft high and about 10 ft in diameter. Liquor flow through bone char is about 1,500 gal/h over 30–60 h. Liquor flow through GAC is 3,000 gal/h for 20–30 days. After the decolorizing cycle is completed, the adsorbent is sweetened off (sugar removed) by displacing the liquor with water. After washing, the spent carbonaceous adsorbent (bone char or GAC) is transferred to regenerating equipment consisting of dryers, kilns, and coolers. Bone char is regenerated at 540°C in a controlled amount of air. GAC is revivified at 950°C in a limited oxygen and steam atmosphere. After regeneration, the adsorbent is returned to the system for a new decolorizing cycle.

Ion-exchange resins are used in columns 8–10 ft high and 6–10 ft in diameter, holding between 100 and 300 ft^3 of resin. Flow rates are rapid (3,000–4,500 gal/h), and the cycle is short (8–16 h). Regeneration is accomplished in the column with a 10–15% salt solution at 45–60°C. The chloride form of a strong anionic resin decolorizes the liquor. Resins can be either acrylic or styrenic in composition. Acrylic resins have a higher capacity for color removal, especially for high molecular weight color bodies, and good resistance to organic fouling. Styrenic resins, although having less capacity, can remove the remaining small molecular weight colorants. Because styrenic resins will foul more easily, acrylic resins are usually placed ahead of them to protect from fouling, and they are used mainly for "polishing" the fine liquor to remove the final traces of color. Both types of resins have ion exchange as the primary mechanism for color removal; styrene resins also exhibit an adsorption mechanism for color removal [56]. The sugar in the sweet water from washing the various adsorbent columns is recovered by returning it to the melter. In recent years ion exchange systems have made use of fractal technology to improve flow and prevent excess pressure drop.

Crystallization

Decolorized liquor is a pale yellow liquid (International Commission for Uniform Methods of Sugar Analysis [ICUMSA] color 200–300) with solids content of 55–65% (see section "Process Control and Instrumentation" for further information on ICUMSA). This liquor goes to multiple-effect evaporators to be concentrated to a solids content of about 72° Brix, and is then pumped to the "white pans" for crystallization. There are four stages in the crystallization process: (1) seeding or graining; (2) establishing the seed; (3) growth of crystals; and (4) concentration. Crystallization is governed by the degree of supersaturation, which is the state in which more solids are in a solution than can be dissolved by water at a given temperature. (The supersaturation coefficient is defined as the mass fraction of sucrose in water divided by the mass fraction of sucrose in water in a saturated solution at a given temperature.) There are three phases of supersaturation: [57]

1. The metastable phase—existing crystals grow in size, but no new ones form.
2. The intermediate phase—existing crystals continue to grow and new ones form.
3. The labile phase—crystals form spontaneously without the presence of others.

Sugar-boiling for proper crystallization is done in the metastable and intermediate phase; below the metastable zone, crystals dissolve and inefficiencies result; above the intermediate phase, uncontrolled "false grain" or extra, small, and agglomerated crystals form, resulting in poor-quality crystals. When the syrup is in the metastable phase, it is "seeded"; that is, a predetermined amount of very small crystals is added, which act as nuclei for crystal growth. A sufficient quantity of evaporated liquor is drawn into the pan to cover the heating elements, and water is evaporated from the syrup until its supersaturation coefficient approaches 1.25. At this point, the steam pressure is lowered, and seed crystals are added. The seed consists of finely pulverized sugar dispersed in isopropyl alcohol or sugar liquor. This method is called shock seeding because addition of the seed induces an immediate formation of crystal nuclei throughout the supersaturated syrup. The nuclei are grown to a predetermined size or grain. Once the grain is established, the crystals are grown to size by maintaining supersaturation between 1.25 and 1.40 through control of the steam pressure, vacuum, and the feed rate of the evaporated liquor. Adequate circulation during crystal growth is important.

When the volume of the massecuite reaches the maximum working capacity of the pan, the syrup feed is shut off, and evaporation is allowed to proceed until a thick massecuite is formed. When the massecuite concentration is considered “tight,” the steam and the vacuum are shut off, and the massecuite is dropped into a holding tank equipped with agitators, where it is kept in motion until discharged into the centrifuges.

The boiling system in a refinery is straightforward. The first strike is boiled from evaporator liquor; the second strike is boiled from the runoff syrup of the first strike, continuing on to three or four strikes. These strikes are blended to form the final white sugar product. The runoff of the last strike, which is quite high in color and ash, can be used in a variety of ways: as syrup for affination, as syrup for remelt, or as syrup for producing brown sugar, also called “soft” sugar in the industry.

Centrifuging. Refined sugar crystals are recovered from the mother liquor by centrifuging the massecuite in equipment similar to that used for affining raw sugar. However, at this point, the crystals are washed with a greater quantity of hot water. The washed crystals are discharged into a holding bin supplying a dryer.

Drying and conditioning. Sugar from the centrifugals contains 1–2% moisture and is too wet to be placed in storage or packaged. The wet sugar is fed to drying equipment called granulators, which are large rotating drums, 15–35 ft long and 6–7 ft in diameter, inclined slightly for gravity discharge. Screw conveyers, scrolls, vibratory “grasshoppers,” or linear belts are used to transport wet sugar from the centrifuge to the drier, where heated air is blown through the dryer, concurrent with the flow of the sugar. Sugar exiting the granulators has a moisture content of about 0.03%.

Remelt Sugar

The sugar contained in affination syrup must be recovered for economy, which is accomplished by crystallizing the sugar in a vacuum pan. The area of the refinery where various “low-grade boilings” take place is called the remelt house. The resulting sugar is “raw” in composition and is returned to the refinery at the affination station, hence it is “remelted.” The residual syrup from the remelt station is known as refiners’ blackstrap, from which no further sugar can be recovered. Various grades or streams of remelt are produced, depending on the refinery.

Packaging and Storing Refined Granulated Sugar

Sugar leaves the dryer at around 52–55°C and is allowed to cool to about 45°C. Sugar leaving the dryer still contains enough moisture to cause problems with caking on storage, and the proper bulk storage and conditioning of sugars in silos is a widely discussed topic [58, 59]. One of the most important considerations for refined sugar is its ability to store well, without caking or darkening in color.

Granulated sugar remains free flowing for a longer period of time if it is “conditioned,” as conditioning further reduces the moisture level, which is necessary to prevent caking. Conditioning takes place in large vertical silos, with up to three million pounds capacity. Conditioned dehumidified air is percolated into the sugar from the bottom of each silo. Residence times in silos vary, from about 24 to 72 h. Proper control of air flow, humidity of the air, and temperature are essential for good conditioning. Sugar leaves the silos with a moisture content of 0.025% or less. Following conditioning, the sugar is screened to meet the particle size requirements of customers. Screened sugar is then sent for packaging or for transport in bulk.

Sugar that is not immediately used is placed in storage silos for periods that can be as long as 2 months. Sugar storage is another area where problems with caking and solidification can occur due to moisture migration, and newer systems are beginning to appear where sugar in storage is treated with a slow stream of conditioned air.

Granulated sugar is packaged in 1, 5, 10, 25, 50, and 100 lb bags, 2,000 lb sacks, as well as cartons and heat seal packets. Other packaging configurations are possible.

Specialty Sugars

Refineries produce a number of specialty sugars and syrups designed to meet the needs of customers. These include sugars with various crystal sizes, powdered sugars, brown sugars, and liquid sugars [60].

Liquid sugars. Liquid sugars (sugar dissolved in water) are economically important because food manufacturers often prefer to use sugar in the form of syrup because of the ease of handling a liquid product. Liquid sugars are prepared by one of two methods: dissolving refined sugar in water; or further purification of in-process liquors by ion-exchange treatment to remove minerals, and further decolorization using bone char and pulverized or granular carbon. Liquid sugar can consist only of sucrose, as a mixture of sucrose and invert sugar, or as a blend with various glucose syrups from starch hydrolysis. Sucrose liquid sugars are usually distributed at a 67% sugar concentration, whereas invert sugar and mixtures of invert and sucrose are distributed at concentrations of 72–77% sugar. Partially inverted sugar syrups are commercially preferable because of the higher solids content and greater microbiological stability.

Partially inverted sugar syrups are made by one of three hydrolytic methods: acid hydrolysis with mineral acids (sulfuric or hydrochloric), hydrolysis by cation ion-exchange resin, or inversion by enzymes. Acid hydrolysis by mineral acids has the disadvantage that the resulting syrups have a high ash content because the solution is neutralized with sodium or potassium hydroxide. Enzymatic conversion has the disadvantage of being expensive and not very efficient at the high temperature and densities of the solutions being hydrolyzed. Cationic resin treatment offers the best alternative, resulting in nearly ash-free invert syrup. The most popular invert syrup is a 50% mixture of sucrose and invert [61].

Liquid brown syrups are also available, known as refiners syrups, and can be adjusted in color, flavor, ash level, and sucrose content according to the specifications of the customer.

Brown or soft sugars. Brown sugars range in color from light to dark brown. They have very small crystals and high moisture content, ranging from 0.5 to 3%. The combination of very small crystals and high moisture gives the sugars the "soft" appearance and texture. Brown/soft sugars also contain higher ash, invert sugar, and an array of compounds that contribute to the characteristic flavors [62]. They are made in two ways. The traditional method is to boil them from a low-purity process liquor to obtain the desired color, flavor, and composition. The newer and more common method is to "paint" or coat a lesser quality (third or fourth strike) refined sugar with refiner's syrup or molasses to produce a product with similar appearance and characteristics. A typical brown sugar composition is 85–90% sucrose, 2–5% invert, 2–4% moisture, and 1–3% ash.

Screened sugars. These are sugars that have been separated on various sieves to give a narrow defined range of particle size. Baker's special, known as caster sugar in England, has a fine grain size, about 75–150 μm, and is desirable because it mixes, blends, and dissolves evenly. Fruit sugar is another screened sugar with a slightly larger grain.

Pulverized sugars. A range of products are manufactured by milling granulated sugar to a desired size. Examples include powdered or confectioner's sugar, with particle sizes 45–75 μm, and characterized by the proportion of the sugar passing through certain sieve sizes. In the United States, these sugars usually are mixed with 3% corn starch to prevent caking. Elsewhere, tricalcium phosphate may be added for the same purpose.

Fondant sugar. Fondant sugar is a special form of pulverized sugar, which exists as a blend of sugar (in the 90–95% range) and another ingredient, such as maltodextrin, invert, glucose, or starch. It may be dry and free-flowing or a paste. It is also called icing sugar or glazing sugar. Dry fondants usually are agglomerated forms.

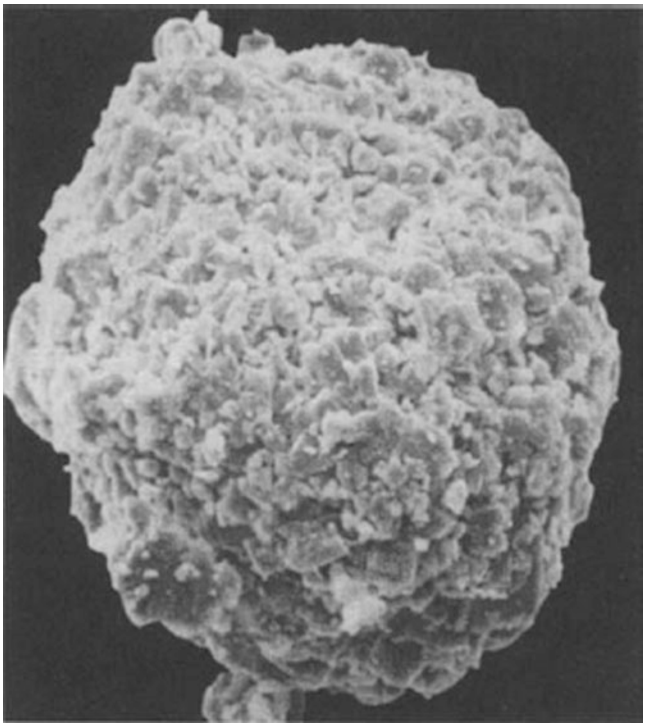


Fig. 35.10 Example of a co-crystallized sugar product. The ball represents many microsized sugar crystals agglomerated into a porous granule, which allows inclusion of a second ingredient throughout the structure

Agglomerated sugars. The agglomeration of sugar crystals helps to prevent caking, so products, such as soft sugars that are subject to caking, can be converted into a free-flowing dry product. Several companies manufacture "dry fondant" by agglomerating mixtures of invert sugar and very fine (less than 40 μm) sucrose particles. Particles below 40 μm cannot be distinguished by the tongue, so the product is perceived as a smooth paste.

Compound products. Compound products are made by manipulating the crystal size and shape and incorporating another ingredient in a process called co-crystallization. A very hot (120°C), supersaturated sucrose solution is allowed to cool with agitation in the presence of a second ingredient, resulting in a dry, free-flowing, agglomerated, porous granule, shown in Fig. 35.10 [63]. Typical products include granulated brown sugar, powdered brown sugar, molasses granules, honey granules, and some fondants. Flavors, such as apple or peanut butter, may be co-crystallized with sugar. The process is said to protect the flavors from loss of volatiles and from oxidation.

Beet Sugar

Like cane sugar processing, the main stages in the extraction and production of sugar from sugar beets have remained basically unchanged for the last 100 years, with cumulative improvements in sucrose yield and processing taking place over the years. Unlike cane sugar manufacture, white sugar is made directly from the sugarbeet. The process is a simple one, with many variations: beets are harvested, stored until needed, washed, sliced, and the sugar-laden juice extracted by diffusion. The juice is purified using lime and carbon dioxide, sometimes also sulfur dioxide, filtered, concentrated, and crystallized. The goal is to produce a good quality white sugar in the highest possible yield. Europe and the United States are the major beet sugar producing areas of the world. Sugar beets are grown in ten states, with Minnesota, North Dakota, Idaho, and Michigan the major producers [64].

Agriculture

The sugar beet, *Beta vulgaris*, is a temperate-zone root crop grown largely in the Northern hemisphere (Fig. 35.11). It is adapted to a wide range of climatic conditions, growing, for example, in the short cold summers of Finland and the hot dry climates of the Imperial Valley of California and southern Spain. Sugar beet agriculture began in earnest in Europe around 1830 and in the United States around 1880. Better-yielding beet varieties, mechanization of the harvest, and increasing efficiency in processing have helped the beet sugar industry grow to the extent that it competes in parts of the world with cane sugar.

The original sugar beet seed is of the multigerm type, which results in several plants germinating in one spot, necessitating laborious hand thinning. In 1950–1952, two Russian scientists, V.F. and Helen Savitsky, working in the United States, developed two lines of monogerm seeds. These seeds develop only one plant per seed, which reduced or eliminated the need to thin the young beet plants. Today all sugar beet cultivation uses monogerm hybrids.

Harvesting and Beet Handling

After the plants have reached maturity in late fall, they are harvested by machines that remove the top growth of leaves, lift the roots from the ground, and deliver them to a holding bin or a truck.

Storage. Because the harvesting season is shorter than the processing season, beets are stored in piles or "clamps" at the factory or at outlying points near transportation. Proper storage of sugar beets is very important to prevent deterioration of the beets. Frequently, storage piles are ventilated to lower the temperature of the beets, thus reducing sugar loss due to respiration during storage. Sucrose losses inevitably occur if the temperature fluctuates widely and if there are freeze–thaw cycles.

Fig. 35.11 A sugar beet (Courtesy of American Sugarbeet Growers Association)

In the Red River Valley area of Minnesota and North Dakota, storage sheds for sugar beets were constructed in the mid-1980s that measure 210 ft wide by 520 ft long and hold 50,000 t of beets. The temperature in these storage facilities is controlled at not much above freezing to minimize degradation of the sucrose content.

Washing. Beets are transferred from storage to the factory in water flumes. These flumes lead directly into a rock-catcher, which allows rocks to settle out, and then on into a trash-catcher and a rotary washer. A large amount of soil usually is entrained with the beet roots, and processes for recovering and returning the soil to the field are in use.

Slicing. The washed beets are sliced into thin V-shaped cossettes by means of specially shaped knives set in frames mounted around the periphery of a rotating drum. Good removal of rocks and trash is essential for a reasonable life for the knives.

Extraction of the Juice

Diffusion. The process of extracting the juice from the beet is called diffusion. The cossettes are weighed and transferred continuously into a diffuser where water passes countercurrent to the movement of the cossettes and where, by diffusion, sugar and some of the nonsugars of the beet are extracted. For optimum extraction of juice, the sucrose-containing plant cells need to be fully disrupted to allow the escape of the juice. This is done mechanically (5–15% of cells are disrupted by the slicers) and with heat. A balance between temperature and the time of heating is necessary: it should be high enough to maximally extract the juice, but not so high as to cause extraction of undesirable nonsugars, such as pectins, or to cause color formation and sugar degradation. The temperature of the juices during diffusion is around 70°C. Retention time is around 1 h. Antifoam agents are used to control foam in continuous diffusers, and bactericides are added for microbiological control. Continuous diffusers have largely replaced batch diffusers. Continuous diffusers come in a variety of forms and shapes, but all employ the same principle, namely the movement of juice countercurrent to the movement of the cossettes.

The RT-type diffuser is a large revolving drum with an internal helix that separates the drum into moving compartments. As the drum revolves, the cossettes travel by the action of the moving helix from one end of the drum to the other end, while the juice moves in the opposite direction. The BMA diffuser is a cylindrical tower with a conveyor mechanism attached to a central rotating shaft. Guide plates on the shaft direct the cossettes upward, while the juice exits through screens at the bottom.

Beet pulp. Beet pulp is the solid residual portion remaining after diffusion, and is a valuable by-product. Wet pulp, containing 85–90% moisture, is dewatered in horizontal double-screw presses, reducing the moisture content to about 75%. Pulp is further dried to about 8–10% moisture in pulp dryers, direct-fired horizontal rotating drums. Newer pulp-drying technology, utilizing a pressurized fluid bed, results in 40% more capacity and better energy efficiency [65]. Beet pulp is mainly used as a pelletized animal feed, either alone or mixed with molasses.

In the 1980s, several beet companies developed edible fiber from beet pulp to take advantage of interest in the health benefits of dietary fiber [66]. The fiber has a high water-holding capacity and other interesting properties, but the market has remained small and specialized, with uses as a fiber additive in both human and pet food.

Juice Purification

The diffusion juice (also called raw juice) is a turbid liquid ranging in color from lavender to nearly black, containing 15–18% sucrose and 1–3% nonsugars, including proteins, nitrogenous bases, amino acids, amides, inorganic material, and pectinaceous matter. It also contains 0.25% sodium and potassium. These impurities are removed by a series of purification processes using lime and carbon dioxide (carbonation), and sulfur dioxide (sulfitation). In recent years, there has been some use of ion exchange resins to produce high-quality refined granulated sugar.

Knowledge of the chemistry of the raw beet juice is important in controlling the phases of juice purification. The addition of increasing amounts of lime and carbon dioxide should precipitate pectinaceous materials and insoluble anions such as phosphate, sulfate, oxalate, and others, as calcium salts, cause alkaline degradation of invert sugar into lactic and other acids, and flocculation of colorants. Lime is also essential for promoting good filtration.

Classic juice purification consists of preliming, main liming, first carbonation, first sludge separation (clarification), second carbonation, second sludge separation, sulfitation, and filtration. Control of pH is very important throughout the process and depends on the quality of the beet juice, which is determined in the lab.

Preliming. A small amount of lime (0.12–0.3%) CaO is added to the juice, increasing the juice pH from about 6.2 to 10.8–11.4. The lime reacts with nonsugar impurities in the juice to make insoluble precipitates and soluble products. Calcium carbonate in the form of recycled first carbonation sludge is added to provide colloid adsorption and stabilization. Temperature may be cool (50°C) or hot (80°C) depending on the factory design. Insoluble precipitates are formed quickly. Retention time is 15–30 min.

Main liming. The addition of more lime (main liming) to an alkalinity of 0.8–1.5% CaO on juice results in a second set of reactions that produce soluble products. During these reactions, the small amount of invert sugar present in the raw juice is destroyed and amides are saponified. These reactions help to stabilize the juice. A great deal of color is formed during this stage, but this type of alkaline-derived colorant is fairly easy to remove in crystallization.

First and second carbonation. In this process, heated, limed juice is sent to a carbonator for gassing with carbon dioxide. The carbon dioxide reacts with the lime to form calcium carbonate, which helps to absorb the insoluble impurities, and some of the soluble impurities as well. The resulting mixture, containing insoluble lime salts, is pumped to subsiders (thickeners, clarifiers) to remove the insolubles by settling (sludge separation), and clarified juice is decanted. From the subsider, the partially clarified filtrate is recarbonated (second carbonation) to lower the pH further, and the residual lime is precipitated. The juice is then treated with about 150 ppm sulfur dioxide gas (sulfitation), which helps to inhibit further color-forming reactions in the juice, removes any remaining traces of calcium, and provides the final pH adjustment, to a pH of around 8–8.5. This treatment is followed by press filtration. The sludge from the subsiders is filtered on rotary-drum filters, and the sugar-containing filtrate is returned to the first carbonation step. The carbonation process can be either continuous or batch. In some factories the clarified juice is also sent to softeners (cation resins) to remove any remaining calcium.

Beet processing uses a large amount of lime (2–2.5% on beet) and efforts are being made to reduce the usage by effective control of temperature, pH, and other parameters [67], which can bring down usage to about 1.0–1.2% and inclusion of innovative processes, such as membrane filtration [68], which can further reduce lime usage to around 0.8% on beets. The recovery of lime by reburning is also being studied. The clarified juice is called “thin juice” and contains about 15% solids.

Evaporation and Standard Liquor

The thin juice discharged from the filter presses after clarification is evaporated to “thick juice” or evaporator syrup (50–65% solids) in multiple effect evaporators, containing five, six, or even seven effects. If decolorizing absorbents are used, they are usually added to the thick juice. Granular and powdered carbons have been used for this purpose. Thick juice that has been treated with powdered carbon usually must be double filtered to remove all the carbon.

Low raw sugars (sugars with purities below that of refined sugar) and white centrifugal wash water are added to the thick juice in the melters to make “standard liquor” from which white sugar is crystallized. Standard liquor is usually filtered with diatomaceous earth before going to the vacuum pan.

Extending the Processing Cycle: Thick Juice Storage

Depending on location, sugarbeet harvesting begins as early as August or as late as October. At least one American company has extended the season by allowing beets to freeze naturally so that they can be processed over the winter. Many other companies in the United States and Europe store excess thick juice over the winter, which can extend production time an additional 40–80 days. Proper storage conditions are critical to maintaining the integrity of the thick juice to prevent sugar loss or microbial contamination, including pH of 9.2, a high solids content of 67–69%, cooling to below 30°C, and addition of soda lye or hop β acid in the upper layer [69].

Crystallization, Centrifuging, and Drying

Crystallization practices in the US beet sugar factories are similar to those in a cane sugar refinery and result in a white granulated sugar comparable in quality to refined cane sugar. White sugar centrifuge stations as well as drying operations are comparable with those operations described earlier for cane sugar factories.

A major difference between cane and beet sugar refining is that beet refining is done in one stage, with refined sugar produced directly from the sugar beets, whereas refined cane sugar is produced first as a raw sugar and then refined. It is possible for beet sugar to be produced directly because the nature of the coloring material in beet sugar is fundamentally different from that of cane. A beet evaporator syrup with color over 3,000 ICUMSA color units produces a white sugar with an ICUMSA color of 25–40. However, for a cane white sugar of the same color to be produced, the color of the evaporator syrup must be only 200–300 color units. The nature of these differences appears to be associated with lower molecular weight colorants in beet, lower polysaccharide content in beet, and colorant mainly composed of alkaline degradation products of invert sugar [70].

Molasses Desugarization

Sugar Recovery from Beet Molasses

Molasses is the viscous, dark-colored material that remains when no more sugar can be economically extracted by crystallization. Beet molasses is generally high in sucrose, 60% dry basis, and represents considerable value. Sugar is routinely recovered from beet molasses by various ion exclusion chromatographic systems.

Historically, recovery of sucrose from molasses was accomplished by the Steffen process, in which sucrose is precipitated from dilute molasses at low temperature (9–14°C) as the tricalcium saccharate by adding finely ground burnt lime. The process uses an excess of lime (100–130% on molasses sugar) and does not work if there is more than 0.4% invert sugar, and other impurities, such as raffinose, which inhibit crystallization [71]. Molasses desugarization today is accomplished by various forms of ion exclusion. Chromatography is no longer done in batch form, which is inefficient, and commercial operations utilize any one of several procedures involving moving beds, simulated moving beds (SMBs), or coupled loops. These are discussed in detail in the literature [72, 73]. Up to 90% of the sucrose in the molasses can be recovered and other valuable products, such as betaine, can also be recovered. Most beet factories today desugarize molasses.

Desugarization is accomplished using cross-linked strong cationic resin in the potassium form. Sucrose is preferentially absorbed by the ion exchange resin and separates by lagging behind the majority of the nonsucrose components. Continuous processes are based on countercurrent technology. The SMB process avoids mechanical movement of resin, keeping the whole resin bed stationary while changing product inlets and outlets with an internal circular flow. Typically, molasses SMB systems have 8–20 resin compartments (or cells) in series. The countercurrent resin movement is simulated by changing the inlet and outlet ports forward in the direction of the liquid flow. In coupled loop technology, developed by Amalgamated Sugar in the 1990s, a cascade of several loops is provided instead of connecting several individual cells. This allows recovery of additional nonsucrose components, such as betaine, and provides a purer sucrose fraction. For proper operation of a molasses desugarization plant, the molasses must be low in calcium and magnesium ions (this is usually achieved by thin juice softening), free of suspended solids, and free of oxygen, which irreversibly damages the resin [74]. The sucrose fraction, called “extract,” goes to pan crystallization, is mixed with thick juice, and crystallized, or is further purified with activated carbon or ion exchange resin and sold as liquid sugar. The waste fraction, called the “raffinate,” is concentrated and sold as fodder molasses.

Desugarization of Cane Molasses

Desugarization of cane molasses is more difficult and less economical for several reasons: it has a lower content of sucrose (about 35%) and a higher content of invert sugar (about 10%); it is more viscous because of higher concentrations of polysaccharides; it has a lot more suspended material; and it has a higher salt and calcium concentration which has to be “softened” before it can be put on a resin. Therefore, cane molasses must be clarified to remove suspended matter and gums, and prior softening is necessary to extend the life of the resin. The resulting product fraction from cane molasses contains a mixture of sucrose and reducing sugars. There has been little adoption of this process in the cane sugar industry [75].

Process Control and Instrumentation

The sugar industry has always been in the forefront of chemical control. The most important areas of control include measurement of sugar content (done by polarization), pH control (because sugar is unstable at low or high pH, and more so when high temperatures prevail), color measurement, and control of vacuum pans (for good crystallization).

At the end of the nineteenth century, the ICUMSA was formed. This organization, which continues today, is dedicated to all aspects of sugar analysis, validation of methods, and promulgating official methods that are scientifically valid. A good early history of ICUMSA, covering the first session in 1897 up to 1936, can be found in Bates [76].

Raw sugar quality characteristics that are measured include polarization, color, grain size, moisture, invert, dextran, starch, and surface color (color that could be removed in affination). New techniques, such as gas chromatography, enzymes, high-pressure liquid chromatography (HPLC), and near infrared (NIR) analysis are in use for determining various parameters, particularly sucrose, glucose, and fructose. NIR is particularly promising because it is nondestructive and can theoretically measure a wide range of parameters, such as moisture in bagasse and polarization of juice. These methods are coming to the forefront because the older method of determining sugar content, polarization, has several drawbacks. Originally, lead acetate was used to clarify solutions for polarization measurement; with lead use restricted in most countries, aluminum and zinc salts have replaced it, but these are not as effective clarifiers. Polarization is also not a direct measurement of sucrose, so it can be inaccurate in low-purity materials, such as molasses.

Instrumentation is used extensively in sugar boiling to control the process of crystallization. Some of the principles used to provide control of vacuum pans for boiling a strike of sugar are listed below [57].

1. Boiling point rise (BPR): thermometers are used to measure the temperature of the massecuite and its vapors (the difference between these two temperatures is the observed BPR).
2. Electrical conductivity of the massecuite, based on the principle that conductivity is inversely proportional to the viscosity of the solution, which in turn has a similar relationship to the water content and thus the degree of supersaturation.
3. Fluidity of the massecuite: An ammeter is used to measure the current used by the motor of a mechanical circulator; changes in current indicate changes in the fluidity of the massecuite.
4. Soluble solids measured by a refractometer.
5. Radiofrequency.
6. Microwave density probe [77].

Cane and Beet Sugar Production

Sugar (cane and beet) is produced in 119 countries. Cane sugar is produced in tropical and subtropical areas, and beet sugar is produced in temperate zones. The United States is fourth among the world's sugar-producing countries and is the only country with significant sugarcane and sugar beet acreage [78]. Cane sugar production is approximately 79% of the world sugar production and beet sugar is 21%. The major producers of cane and beet sugar are shown in Table 35.1. The biggest change to occur in world sugar production is the ascension of Brazil as the major producer [79]. Brazil, India, and China alone account for 41.0% of world sugar production, but only Brazil is a major exporter.

Sugar Consumption and Usage

Per capita consumption of sugar varies widely from country to country and is often a measure of the affluence and level of industrial development of a country, except in Asia, where dietary customs do not include high sugar intake.

Table 35.2 shows the per capita consumption of selected areas [80]. Sugar usage patterns, particularly in developed countries, have not changed markedly over the years except in the United States, where high fructose corn syrup (HFCS) has replaced sugar in the manufacture of beverages (see Table 35.3). Nonfood uses of sucrose constitute a small portion of total use [81].

Derivatives of Sucrose: Sucrochemistry

Because of its high purity and wide availability, sucrose has been viewed as an attractive feedstock for organic chemical synthesis. Sucrochemistry has been pursued for many years, and many products have been produced [82–84]. However, the economics of production have not been good and the reactivity of sucrose is difficult to control. Among the derivatives of sucrose are ethers, fatty and other esters, acetals, and reduction products. Successful products from

Table 35.1 Major sugar-producing countries, 2008/2009 (in units of one million metric tons) [78, 105]

Total sugar production		Observations
Brazil	33.900	Cane
India	16.100	Cane
China	13.510	Cane
United States	7.847	57% beet; 43% cane
Thailand	7.472	Cane
Mexico	5.244	Cane
Australia	4.834	Cane
France	3.624	Beet
Germany	3.515	Beet
Pakistan	3.473	Cane

Table 35.2 Per capita consumption of sugar (cane and beet) in selected countries, 2003/2004 [80]

Country	(lb/year)
Australia	133
Brazil	121
Canada	97
China	19
Cuba	130
Germany	87
India	41
Indonesia	35
Japan	42
Mexico	110
Russia	103
United States[a]	66

[a]Does not include US consumption of corn sweetener, which is about 63 lb/year

Table 35.3 US sugar deliveries to industrial and nonindustrial users, selected years (1,000 short tons, raw value) [81]

Use	1990	1993	1995	2000	2005	2009
Bakery and cereal products	1,608	1,785	1,905	2,264	2,297	2,286
Confectionery products	1,279	1,292	1,372	1,328	1,131	1,082
Ice cream and dairy products	462	424	452	499	587	577
Beverages	228	158	169	168	237	351
Canned, bottled, and frozen foods	332	336	279	330	336	427
All other food uses	642	725	864	817	606	572
Nonfood use	102	85	64	85	92	83
Total industrial use	4,660	4,805	5,103	5,491	5,286	5,377
Hotels, restaurants, and institutions	108	108	103	71	115	127
Wholesale grocers, jobbers, sugar dealers	2,130	2,075	2,173	2,241	2,401	2,352
Retail grocers, chain stores	1,077	1,235	1,236	1,242	1,262	1,241
All other deliveries	76	171	189	339	248	179
Total nonindustrial use	3,391	3,589	3,701	3,893	4,020	3,899
Total United States	8,051	8,394	8,804	9,383	9,312	9,276

sucrose include surfactants, surface coatings, food and feed additives, polymers, textile chemicals, pharmaceuticals, and pesticides. A promising development is the conversion of sugar-based feedstocks into conventional liquid fuels and chemicals using a patented catalytic aqueous phase reforming process [85]. Other sucrose-based products in successful commercial distribution include olestra, a liquid sucrose polyester fat substitute, sucrose acetate isobutyrate, and sucrose-based detergents [86]. Sucralose, a sucrose derivative containing three chlorine molecules, was developed in 1976 by Tate & Lyle, and received approval for food use in the United States in 1998. It is 600 times sweeter than sucrose, is noncaloric, and stable to high temperature and low pH.

Sweeteners Derived from Starch

Starch is hydrolyzed to produce a range of sweeteners, including syrups and crystalline products. The bulk of starch-derived sweeteners comes from corn (maize) starch. Other starchy products are also used to produce starch-derived sweeteners, and small specialty markets exist for syrups from rice, barley, wheat, tapioca, and cassava.

The United States is the largest producer of corn sweeteners and the largest market for high-fructose corn syrup. The major use of HFCS is in beverages and soft drinks, which account for 75–80% of usage. Important milestones in the development and growth of HFCS are shown below [87].

- Late 1950s—Discovery that glucose isomerase enzyme converts glucose into fructose
- 1967—First commercial US production
- 1972—Immobilization of glucose isomerase for continuous production of HFCS-42
- 1977—Commercial production of HFCS-55
- 1980—HFCS approved as 50% of the sweetener in Coca-Cola
- 1983—HFCS approved as 50% of the sweetener in Pepsi-Cola
- 1984—100% level of HFCS approved in Coca-Cola and Pepsi-Cola

The 1984 approval of 100% HFCS as a sweetener in soft drinks resulted in phenomenal growth in the HFCS industry. In 1985, HFCS production was 5.3 million short tons (dry weight) and in 2000, it was 9.3 million short tons, with 9.6 million forecast for 2002 [87, 88]. In 2001, consumption of sugar and HFCS was almost equal in the United States.

In Europe, where other starch sources, such as wheat and potato, are used, these are referred to as high fructose syrups (HFS) or high fructose starch-based syrups (HFSS).

The corn industry makes widespread use of enzymes for carbohydrate conversion. The advent of enzyme technology in the corn industry in the 1960s dramatically changed the starch industry and allowed the development of new products. Today, enzyme hydrolysis of starch has largely replaced acid hydrolysis, which is used as an adjunct in starch conversion. Enzymes used to make corn syrups and HFCS include the following.

- Alpha amylase—Converts starch into dextrin; used in the manufacture of glucose syrup.
- Glucoamylase—Converts dextrin into glucose; specific for amylose.
- Pullulanase—Converts the amylopectin fraction of starch into linear segments.
- Beta amylase—Releases maltose units from the nonreducing end of a polysaccharide.
- Glucose isomerase—Converts glucose into fructose; used to make HFCS.

Starch Conversion

Starch conversion refers to the process of converting starch into other products. It involves gelatinization, liquefaction, and saccharification. Liquefaction refers to the acid- or enzyme-catalyzed conversion of starch into maltodextrin. Starch, usually from wet milling of corn, is pumped in a slurry to the conversion plant, where it undergoes one or more hydrolytic processes to yield mixtures of various carbohydrates in the form of syrups. The kind and amount of the various carbohydrates obtained depend upon the type of hydrolysis system used (acid, acid–enzyme, or enzyme–enzyme), the extent to which the hydrolytic reaction is allowed to proceed, and the type of enzyme(s) used. The fact that most starches consist of two different kinds of polymers (amylose and amylopectin) also influences the nature of the products obtained.

The extent to which starch is converted into simpler carbohydrates is indicated by the dextrose equivalent (DE), a measure of the reducing sugar content calculated as dextrose and expressed as a percentage of the dry substances. Hydrolyzates having dextrose equivalents ranging from 5 to 100 are produced. Those having a low dextrose equivalent frequently are referred to as dextrins and are not sweet. They are produced by minimal acid hydrolysis or roasting. Starchhydrolyzate syrups commonly are produced as “low,” “intermediate,” “high,” or “very-high” conversion products, as more or less standard products, as shown in Table 35.4. HFCSs are classified according to the fructose content, as a percent of solids (i.e., 42%, 55%, 90%), rather than by DE. Table 35.5 shows the composition of various corn syrups. In addition to carbohydrates, the syrups contain some sodium chloride and traces of nitrogenous substances. Maltodextrin is a dried corn syrup having a DE below 20, whereas the dry dextrins with a DE above 20 are called corn syrup solids.

Acid Hydrolysis of Starch

As mentioned above, this process now serves mainly as an adjunct to enzymatic conversion of starch and is rarely used alone. A starch slurry containing 35–45% solids is acidified with hydrochloric acid to about pH 1.8–1.9. The suspension is pumped into an autoclave (converter) where live steam is gradually admitted to a pressure of 30–45 psi. The conversion time largely determines the DE of the hydrolyzate; for example, 8 min may produce 42 DE syrup, and 10 min may produce 55 DE [89]. Converted liquors are neutralized with sodium carbonate to a pH of 5–7, with coagulation of insoluble protein, fats, and colloidal matter. The scum is removed by centrifugation.

The dark-colored clarified liquor is pressure-filtered and concentrated to 60% solids in multiple-effect evaporators. The concentrated liquor is decolorized with granular carbon in columns 12 ft in diameter and 30 ft high in a countercurrent manner; that is, liquor flows upward in the columns, while a portion of the carbon is removed from the bottom periodically. Carbon is used at a rate of 2.5% of dry solid processed, and approximately 5% of carbon is lost during revivification [90]. In some cases, granular carbon has been replaced by synthetic polymeric adsorbents to decolorize the syrup. “Low ash” syrups usually are deionized with ion-exchange resins. The processed liquor is evaporated to a final solids content of 75–85% in a single-effect evaporator.

Acid–Enzyme Hydrolysis

Starch is first liquefied and hydrolyzed to specific dextrose equivalents with hydrochloric acid. After evaporation to 60% solids, a saccharifying enzyme (fungal α-amylase) is added to continue hydrolysis to the desired level. By choosing two or more types of enzymes (such as α-amylase, β-amylase, glucoamylase, pullulanase) and adjusting the initial acid hydrolysis, syrups with different ratios of dextrose, maltose, and higher saccharides can be obtained [91].

Table 35.4 Conversion groups for starch hydrolyzates

Conversion	DE
Low (Type I)	20–38
Intermediate (Type II)	38–58
High (Type III)	58–73
Very high (Type IV)	73 and above

Enzyme–Enzyme Hydrolysis

Enzyme–enzyme conversion employs heat and an enzyme for starch liquefaction in place of acid. This is the most common form of corn processing today. Subsequent hydrolysis is by enzymes, as above. The choice of hydrolytic system depends upon economics and the kind of endproduct desired. Enzymes are usually inactivated by heating the syrup to 75–80°C, with the exception of the heat-stable α-amylases that have come on the market in the last 10–15 years.

A starch slurry of 30–40% dry matter is gelatinized and liquefied in a single dose jet-cooking process. Heat-stable α-amylase enzyme is metered into the starch slurry after pH adjustment to 6.0–6.5, and the slurry pumped through a jet cooker. Steam is injected to raise the temperature to 105°C and held for about 5 min, which is sufficient to gelatinize the starch (break up the starch granules). The partially liquefied starch is then cooled to 95–100°C and the enzyme is reacted at this temperature for 1–2 h, until the required DE is obtained [91]. Following liquefaction, the starch undergoes saccharification (conversion to sugars) using glucoamylase and pullulanase enzyme, resulting in maltose syrups, glucose syrups, and mixed syrups, as described above. Two enzymes are required at this point because starch is made up of two types of glucose polymers—amylose, which is a linear polymer, and amylopectin, a branched polymer—and each requires a different enzyme to break it up. Glucoamylase hydrolyzes the linear chains of amylose, and pullulanase specifically attacks the amylopectin at the branching points. The proper combination and time of hydrolysis will result in the desired range of products, all the way to syrups with as much as 97% glucose content. The glucose syrup may also undergo isomerization with immobilized glucose isomerase to make fructose syrups.

Crystalline Dextrose

Highly purified, high-dextrose syrup is crystallized to produce crystalline dextrose. The high-purity liquor is pumped to insulated crystallizers fitted with slowly moving

Table 35.5 Carbohydrate composition of glucose syrups (saccharides as a percent of total carbohydrates) [106]

DE	Glucose	Maltose	Tri-[a]	Tetra-[a]	Penta-[a]	Hexa-[a]	Hepta-[a]	Higher
15	3.7	4.4	4.4	4.5	4.3	3.3	3.0	72.4
35	13.4	11.3	10.0	9.1	7.8	6.5	5.5	36.4
45	21.0	14.9	12.2	10.1	8.4	6.5	5.6	21.3
55	30.8	18.1	13.2	9.5	7.2	5.1	4.2	11.9
65	42.5	20.9	12.7	7.5	5.1	3.6	2.2	5.5

[a] Refers to the number of glucose units in the oligosaccharide: tri maltotriose; tetra maltotetraose; penta maltopentaose, etc.

agitators for crystallization of dextrose monohydrate. A heavy seed base (about 25%) from a previous batch is mingled with the syrup and cooled to about 38°C. The seeded liquor is held at this temperature for several days until about 60% has crystallized. The mixture then is centrifuged to separate the crystals from the mother liquor. The wet sugar is dried in rotary dryers or recrystallized into anhydrous dextrose. The monohydrate also may be converted to anhydrous dextrose by drying in hot air. A second crop of crystals is taken from the mother liquor, and the runoff syrup from this step is final molasses or “hydrol.”

Crystalline Fructose

The newest sweetener from the corn industry is crystalline fructose, which came on the market around 1987, and is now available in retail outlets as a sweetener as well as used in commercial products.

High Fructose Corn Syrup

HFCS is produced from very pure glucose syrup using glucose isomerase. Commercial glucose isomerase is produced from a variety of microorganisms: *Bacillus coagulans*, *Actinplanes missourensis*, and several *Streptomyces* species. Glucose isomerase is always used in immobilized form because it is an intracellular enzyme. Immobilization is generally done by cross-linking with glutaraldehyde after the cells have been disrupted by passing through a homogenizer. The isomerization reaction is carried out at 60°C at pH 7–8. Reaction time is limited to prevent degradation of the fructose. All isomerization reactions today are done in continuous mode, rather than batch mode. The operating lifetime of a column of immobilized enzyme can be as long as 200–360 days, depending on the robustness of the enzyme used [91].

The basic feature of the isomerization process is a series of reactors containing the immobilized enzyme fixed in a packed bed. The feed material is of the highest practical purity possible because enzymatic activity is closely related to the purity of the glucose syrup. Glucose syrup of about 94 DE is filtered, treated with activated carbon to remove residual color, and deionized with ion-exchange resin to lower the ash content, particularly the calcium ion, which interferes with the magnesium enzyme activator. The dissolved oxygen is reduced by flash evaporation, which also concentrates the feedstock to 40–45% glucose and raises the temperature to 60–65°C. At this point, prior to entering the reactors, magnesium ions are added as enzyme activators. The pH is adjusted to about 8. Isomerized liquor is removed from the process when the equivalent of 42% fructose on a dryo basis is reached. After isomerization, the pH of the syrup is lowered to 4–5, and it is purified by ion-exchange resin and activated carbon. This results in a 42% HFCS. For use in soft drinks, it is converted to 55% HFCS by passing through large chromatographic columns of zeolites or the calcium salts of cation-exchange resins, which absorb and separate the fructose from the other components. This produces a stream of about 90% fructose, which is then blended with 42% fructose syrup to produce the desired 55% fructose syrup product.

HFCS technology and products have spread to Canada, Japan, and several western European nations.

Molasses

Molasses, defined as the residual mother liquor from which little or no additional sugar can be recovered economically, is a by-product of the cane, beet, and dextrose industries. Each type of manufacturer has designated this liquid by-product with its own name. The molasses from both raw cane sugar production and cane sugar refining is commonly called blackstrap molasses. In the cane mill, it is also called final molasses, and in the refinery it is sometimes called refinery molasses. Molasses from beets is simply called beet molasses, and that from starch hydrolysis is known in the United States as hydrol. The composition of molasses varies depending on location, varieties harvested, and the efficiency of the operation. The amount of sugar remaining in molasses is a measure of its “exhaustion.” The amount of refinery molasses is usually small, because as much as possible is recycled and a lot of it goes into the manufacture of brown sugar. Some cane refineries have no molasses output at all. The general composition of each type of molasses is shown in Table 35.6.

Blackstrap molasses from the cane mill contains significant quantities of sucrose and reducing sugars, some of it produced during the processing of the cane juice, caused by hydrolysis of sucrose, and some of it from the native cane juice. Refinery blackstrap has fewer impurities because the raw sugar starting material has fewer impurities than the cane juice but more sucrose. Beet molasses contains primarily sucrose and little or no reducing sugar because of the lower content of reducing sugars in beet juice and because the highly alkaline processing conditions destroy most of the reducing sugars. Hydrol contains no sucrose. Higher saccharides result from incomplete hydrolysis of starch. Cane and beet molasses contain polysaccharides that come from the plant (starch and cell wall polysaccharides) as well as some oligosaccharides that either occur naturally or are formed during processing.

Table 35.6 Composition of molasses from various sources[a]

Constituent	Cane final molasses	Refinery blackstrap	Beet molasses	Hydrol
Sucrose	31–40	50–60	60–63	0
Reducing sugars	20–30	5–10	0.5–1.5	55.0
Higher saccharides[b]	1.5–4.0	0.5–1.0	1.0–2.0	1
Ash	10–14	5–10	8–10	7.2
Nitrogen	0.4	0.1–0.3	2.0	0.07

[a] Values are on a dry basis
[b] Includes polysaccharides

The main uses of molasses are for cattle feed, production of some fermentation products, such as citric acid, and ethanol production. The sucrose in beet molasses is recovered by chromatographic processes described earlier. Cane molasses can also serve as a fermentation substrate for rum and "power" or fuel ethanol and for the production of baker's yeast.

Brazil has been in the forefront of producing fuel ethanol from cane sugar. Brazil unveiled the world's first alcohol-powered car after the 1970s oil crisis, using its abundant supplies of sugarcane. In spite of some difficulties in the 1990s, this innovative program was very successful. Brazil is the major producer of alcohol from sugarcane, with a total production in 2009 of 5.67 billion gallons. The United States is the major producer of alcohol in the world, with a total production in 2009 of 10.6 billion gallons [92]. US bioethanol is made from corn. Profitability is enhanced by the fact that the sugarcane mills produce both sugar and ethanol and can decide at any time on the product mix that gives the best return.

Other Sweeteners

Several carbohydrate-based sweeteners exist on the market today. Although sucrose is the "gold standard" for sweeteners, other sweeteners gain commercial viability because they are less expensive (such as corn-based sweeteners), have fewer calories, or are better suited for diabetic diets. Unlike the synthetic sweeteners, most of the carbohydrate-based sweeteners are less sweet than sucrose, with the notable exception of sucralose, which is 600 times sweeter than sucrose and is nonnutritive. Fructose and xylitol are reported to be sweeter than sucrose. Table 35.7 shows the relative sweetness of the carbohydrate-based sweeteners [93, 94].

An important group of carbohydrate-based sweeteners are the polyols, also known as sugar alcohols. These are not considered sugars for labeling purposes, so foods containing them as sweeteners can be called "sugar free." The polyols are usually 15–30% sweeter than their corresponding sugar. Some of the polyols give a cooling sensation on the tongue. Several of the new carbohydrate sweeteners are described below.

Table 35.7 Relative sweetness (RS) of carbohydrate-based sweeteners

Sweetener	RS
Crystalline fructose	180
Fructose in solution (5–15 %)	115–125
High fructose corn syrup	100–130
Invert syrup	105
Sucrose	100
Tagatose	92
Xylitol (10 % solution)	85
Xylitol (10 % solution)	100
Xylitol (10 % in solution)	120
Glucose (2 % solution)	50
Glucose (8–10 % solution)	60–70
Glucose (50 % solution)	90–100
Erythritol	70
Sorbitol (hydrogenated glucose)	70
Maltitol (hydrogenated maltose)	68
Maltose	50
Trehalose	45–50
Mannitol	40
Lactitol (hydrogenated lactose)	30–40
Lactose (milk sugar)	15–30

Erythritol, 70% as sweet as sucrose, is a polyol (sugar alcohol), currently used as a bulk sweetener in reduced-calorie foods. It has only 0.2 cal/g (sucrose has 4 cal/g). It was given Generally Recognized As Safe (GRAS) status by the Food and Drug Administration (FDA) in 1997, which allows it to be used in foods in the United States [95]. Erythritol is produced by fermentation with the fungus *Moniliella pollinis* [96].

Tagatose, a monosaccharide similar to fructose, received GRAS status from the FDA in April 2001. This sugar, 92% as sweet as sucrose, with 1.5 cal/g, is produced from galactose obtained from whey, using immobilized enzymes [97]. It has similar bulking properties to sucrose.

Trehalose, a disaccharide sweetener, 45–50% as sweet as sucrose, was given GRAS status in 2000. It is naturally found in mushrooms, honey, lobster, shrimp, and foods made with yeast. It has been used in Japan for decades, and is commercially produced from starch by bacterial enzymes [98]. Besides its mild sweetening power, it maintains cell structure during freezing and dehydration of foods. It is a nonreducing sugar, so it does not participate in the Maillard reaction (will not brown) and helps to protect the color of processed foods [99].

Other naturally occurring sweeteners include palm sugars, maple sugar, honey, and agave syrups (composed of about 90% fructose).

Regulation and Trade in Sugar

Raw cane sugar, refined sugar, sugar syrups, specialty sugars, and sugar-containing products enter the United States under a variety of tariff-rate quotas (TRQs). Low within-quota tariff rates facilitate access to the US market. The minimum sizes of the TRQs and the corresponding duties reflect obligations made by the United States. When Congress ratified the NAFTA and WTO (Uruguay Round) treaties, these TRQs became part of the US Harmonized Tariff Schedule.

Most of the sugar produced in the world (about 70%) is used for local or domestic consumption. The remaining 30% is in excess of local demand and makes up the global trade in sugar. In recent years, the trend has been toward the export of more white sugar instead of raw sugar. The European Union (EU-25) exports around 5.5 million tons of white sugar, and Britain, France, and Finland continue to import raw cane sugar for refining. Brazil is currently the largest exporter of sugar, at 27.4 million tons in 2008/2009 [78, 79].

Trade in sugar throughout much of the world is controlled by agreements and special trading arrangements. The United States regulated sugar production, distribution, and importation by means of the Sugar Act of 1948 as amended (US Congress, Public Law 339), but the act was allowed to expire in 1974. From 1981 to the present, the United States has had a price support program, connected to farm legislation. In 1982, an import quota system went into effect to supplement the price support program. In 1988, Australia filed a complaint against the United States, charging that the import restriction on sugar violated the General Agreement on Tariffs and Trade (GATT) rules. When the United States was found to be in violation on some of these rules, it implemented a change from an absolute quota to a tariff rate quota. The United States also has a program called the General System of Preferences (GSP) that exempts countries (mostly in the Caribbean area) from any duty. The EU countries import raw sugar within a special agreement with ACP (Africa, Caribbean, and Pacific) countries.

Raw cane sugar prices are based on the price of sugar delivered to New York and is quoted on the New York Coffee, Sugar, and Cocoa Exchange as Sugar Number 14 Contract (domestic). There is no futures market for US refined sugar. The world refined sugar price is based on the Number 5 Contract on the London International Financial Futures and Options Exchange (LIFFE), which is based on the London daily spot market price for refined sugar in free-on-board ships in European ports [100]. Copies of the Number 11 (world sugar) and Number 14 contracts (domestic) for raw sugar can be found in issues of *F.O. Lichts World Sugar and Sweetener Yearbook* prior to 2002.

The USDA's Sugar-Containing Products Re-Export Program is designed to put US manufacturers of sugar-containing products on a level playing field in the world market. The Refined Sugar Re-Export Program is designed to facilitate the use of domestic refining capacity to export refined sugar into the world market. The program establishes a license against which a refiner can export domestically produced refined sugar and later import world raw sugar, import world raw sugar for refining and distribution into the domestic market, and later export refined sugar, or import raw sugar, refine it and export it into the world market. The program was implemented to mitigate the imposition of restrictive quotas, which reduced the quantity of raw sugar allowed to enter the US domestic market.

Environmental Concerns and Sustainability

The sugar industry, in common with all other sectors of the food and chemical processing industries, is charged with protecting the environment. This means that close attention must be paid to all gaseous and liquid effluents, with the aim of minimizing their production and ensuring that they meet all minimum standards before being discharged into the environment. The goal of the sugar industry is to be ecologically sustainable and economically viable and efficient [101]. It also must take into account community concerns, such as the question of whether to continue the preharvest burning of sugarcane in various areas. Another factor of community concern is the odor produced by beet factories, and odor abatement programs are in place. It is certain that present standards will not be relaxed and that they can be expected to become more restrictive as time goes on [102]. Lead usage in polarization analysis was discussed earlier; most sugar producers now use alternatives (aluminum salts) or newer methods of analysis such as high wavelength polarimetry, NIR, or HPLC.

Issues of Genetic Engineering

In the United States, herbicide-tolerant genetically engineered sugar beet was approved in 1998, and has been available to the market since 1999. However, farmers have been wary about planting it because soft drink, food, and other industrial users are concerned about the growing consumer doubts and the inability to export such foods to the European Union. As of this writing, no genetically engineered commercial cane exists, but research in various countries has developed modified varieties. Brazil developed the first transgenic canes in 1994 and continues an intense program, with at least three varieties developed for herbicide tolerance, cold tolerance, and flowering control. The introduction of GMO cane is not foreseen for at least 5 years in Brazil due to social factors and strong regulations [103]. Successful introduction of modified cane or beet sugar varieties will depend on public opinion rather than scientific factors. Modified corn has been grown in the United States for many years, and several of the enzymes used are from modified organisms. Because these have been in place for some years, there is more acceptance of the products.

References

1. USDA Economic Research Service, Sugar & Sweetener Yearbook Tables, 2009. http://www.ers.usda.gov/publications/so/view.asp?f=specialty/sss-bb/
2. F.O Lichts world sugar statistics. F.O. Licht GmbH, 2004, World beet and cane sugar production, pp 14–15
3. F.O. Lichts world sugar statistics. F.O. Licht GmbH, 2001, World production, consumption and stocks of centrifugal sugar 1952–2000, pp 12–13
4. Sawhney DH (2002) The Indian sugar industry. Zuckerind 127:186–194
5. Tegge G (1978) Development and prospects of sweeteners obtained by glucose isomerization. In: Ahlfeld H (ed) International sweetener report no. 1. Licht, GmbH, Ratezeburg, pp 3–8
6. (2000) Recent developments in the sugar industry, by the Editor. Int Sugar J 102:43–60
7. Blackwell J (2003) Recent developments in sugar processing. F.O. Lichts World Sugar and Sweetener Yearbook, E5–E25
8. Moroz RD, Broeg CD (1992) Sugar and Other Sweeteners. In: Kent JA (ed) Riegel's handbook of industrial chemistry, 9th edn. Kluwer Academic, New York, Chapter 9, pp 315–316
9. Hugill JAC (1979) History of the sugar industry. In: Birch GG, Parker KJ (eds) Sugar: science and technology. Applied Science, London, pp 15–37
10. Meade GP (1959–1960) The sugar molecule, vol 11(1). Sugar Research Foundation, New York, pp 11–19
11. Deere N (1949) The history of sugar, vol 1. Chapman & Hall, London
12. Hugot E, Jenkins GH (1960) Handbook of cane sugar engineering. American Elsevier, New York
13. McGinnis RA (1951) Beet sugar technology. Van Nostrand Reinhold, New York
14. Eggleston G, Legendre BJ, Richard C (2002) Effect of harvest method and storage time on sugarcane deterioration. Proceedings of the sugar processing research conference, Porto, Portugal, pp 205–221
15. Godshall MA, Legendre BL, Richard C, Triche R (2002) Effect of harvest system on cane juice quality. Proceedings of the sugar processing research conference, Porto, Portugal, pp 222–236
16. Kördel P (1998/1999) New equipment and processes in the sugar industry. F.O. Lichts Guide to Equipment, Products and Services for the Sugar and Allied Industries, Directory of Suppliers, pp E3–E25
17. Rein P (2002) Cane preparation. Sugar Bull 8:9–11
18. Ducasse V (1978) Heavy duty chopper fiberiser. Proc Int Soc Sugar Cane Technol 16:1568–1601
19. Anon (1970) BMA cane diffusion. Sugar y Azucar, p 22, July 1970
20. Schroder KE, Voigt I, Moor B (2007) The conversion of a BMA diffuser to a Bosch Projects chainless diffuser—the first season's experience. Proceedings of the international society of sugar cane technologies, vol 24, CD, Durban, South Africa
21. Rein PW (1995) A comparison of cane diffusion and milling. Proc South Afr Sugar Technol Assoc 69:196–200
22. ISSCT Engineering and Energy Workshop on Factory Design, Pakistan, 1997. http://www.sugaronline.com/issct/energyengineering/engw.htm
23. Fowers M, Lloyd LE (1999) Biocide synergy: reducing costs and losses through physical and chemical interactions. In: Presented at the 30th general meeting of the American Society of Sugar Beet Technologists, Orlando, FL, Feb 1999
24. Pollach G, Hein W, Beddie D (2002) Application of hop beta acids and rosin acids in the sugar industry. Proceedings of the sugar processing research conference, Porto, Portugal, pp 51–76
25. Pollach G, Hein W, Beddie D (2004) The concept of different natural antibacterials for the sugar industry. Proceedings of the sugar processing research conference, Porto, Portugal, pp 237–259
26. Doherty WOS, Greenwood J, Pilaski D, Wright PG (2002) The effect of liming conditions in juice clarification. Proc Aust Soc Sugar Cane Technol 24:443–451
27. Eggleston G, Monge A, Pepperman A (2002) Preheating and incubation of cane juice prior to liming: a comparison of intermediate and cold lime clarification. J Agric Food Chem 50:484–490
28. Eggleston G, Monge A, Ogier B (2002) Further insights on the factory performance of cold, intermediate and hot lime clarification systems. Proceedings of the sugar processing research conference, vol 348–365, New Orleans, Louisiana, USA
29. Steindl RJ (1999) The clarification of cane juice. Int Sugar J 101:213–215
30. (1993) Raw syrup clarification processes. In: Chen JCP, Chou CC (eds) Cane sugar handbook, 12th edn (Section 5.77). Wiley, New York, pp 156–162
31. Chen JCP, Chou CC (1993) Cane sugar handbook (Chapter 5), 12th edn. Wiley, New York, pp 146–150
32. Sugar technology, beet and cane sugar manufacture. In: van der Poel PW, Schiweck H, Schwartz T (eds) Verlag Dr. Albert Bartens KG, Berlin, Chapter 9.2.4, pp 513–514, 1998
33. Moodley M, Bekker DJ, Pienaar PJ, Pillay R (1997) Replacement of sulphur burner with liquid sulphur dioxide system at the malelane refinery. Proc South Afr Sugar Technol Assoc 71:142–145
34. Rein PW, Kampen H (2002) Evaporators in cane sugar mills. Sugar J 65(7):18–21
35. Grant G, Anders M, Morgenroth B, Niepoth K, Ehrenreich S (2001) Application of the new falling-film plate technology in the cane sugar industry. Int Sugar J 103:70–73
36. Walthew DC (1996) Aspects of evaporator scale formation and control in the South African sugar industry. Proceedings of the sugar processing research conference, vol 22–43, Walthew
37. Godshall MA, Whartell L (2002) Composition of evaporator scale in Louisiana mills. Proceedings of the sugar processing research conference, New Orleans, LA, 379–385
38. Rein PN (1986) A review of experience with continuous pans in Tongatt–Hulett sugar. Proceedings of the South African sugar technologists association, Durban, South Africa
39. Moor B (2001) Continuous pans: part 1—history, theory and experience applied in a new design. Int Sugar J 103:542–546
40. Eggleston G, Monge A (2004) Optimization of sugarcane factory application of commercial dextranases in the U.S. Proceedings of the sugar processing research conference, Atlanta, GA, pp 371–394

41. (1998) Influence of individual nonsugars on molasses formation. In: van der Poel PW et al (eds) Sugar technology, beet and cane sugar manufacture. Verlag Dr. Albert Bartens, Berlin, pp 685–686
42. Vavrinecz G (1965) Atlas of sugar crystals. Verlag Dr. Albert Bartens, Berlin
43. Coerver JF (2002) Is bioenergy opportunity knocking? Sugar Bull 80(4):13–14
44. Manohar Rao PJ (1997) Industrial utilization of sugar cane and its co-products. ISPCK Publishers, Delhi, India
45. Yang W, Lan C, Lian H (2000) High quality paper products from Bagasse. Proceedings of the sugar processing research conference, Porto, Portugal, pp 274–285
46. Bennett MC, Elvin JR, Heineman HWB, Pottage PR, Rundell JT (1978) Syrup clarification for improving sugar quality and yield. Proc Int Soc Sugar Cane Technol 16:2797–2809
47. Bosworth M (2000) Billington's—the unrefined sugar specialists. Int Sugar J 102:654–655
48. Buzzanell PJ (2000) Organic sugar: short term fad or long term growth opportunity?. Paper presented at international sugar organization, 9th international seminar, Hot issues for sugar, London, Nov 2000. http://www.organic-research.co.uk/research/Papers/sugar.htm
49. Kochergin V, Tzchartzsch O (2005) Ion exchange thin juice softening using fractal technology. Int Sugar J 207:342–353
50. Blackwell J (2002) Recent developments in sugar processing. Int Sugar J 104:28–42
51. Kochergin V (2001) Practical applications of membrane filtration in the sugar industry. Int Sugar J 103:529–532
52. Steindl RJ (2001) Membrane filtration technology in the cane sugar industry. Proc Int Soc Sugar Cane Technol 24:3–10
53. Rossiter GJ (1999) Chromatography in sugar processing. Symposium on advanced technology for raw sugar and cane and beet refined sugar production, vol 181–193, New Orleans, LA
54. Rossiter G, Jensen C, Fechter W (2002) White sugar from cane at the factory: the impact of WSM. Proceedings of the sugar processing research conference, vol 162–177, Atlanta, GA
55. Cane sugar handbook, 12th edn (Section 12.15). In: Chen JCP, Chou CC (eds) Modification of phosphatation and/or carbonation. Wiley, New York, p 458, 1993.
56. MacDonald SEM (1996) Continuous ion exchange adsorption for the decolourization of sugar. In: Proceedings of workshop on separation processes in the sugar industry, Sugar Processing Research Institute, New Orleans, pp 29–54
57. Cane sugar handbook, 12th edn (Section 7). In: Chen JCP, Chou CC (eds) Pan control and use of instruments. Wiley, New York, pp 236–248, 1993.
58. Mikuz O, Budizeck L (1986) Sugar storage in silos. Sugar Technol Rev 13:53–129
59. Burke RT (1996) Conditioning of white sugar in silos. Proc Sugar Ind Technol 55:181–195
60. Godshall MA (2002) The specialty sucrose market—an important niche. Proc Sugar Ind Technol 61:251–265
61. van der Poel PW, Schiweck H, Schwartz T (eds) (1998) Liquid sugars: manufacture and properties. In: Sugar technology, beet and cane sugar manufacture. Verlag Dr. Albert Bartens KG, Berlin, pp 951–956
62. Godshall MA, Vinnet CH, Chew V (1986) Sensory analysis of brown sugars and its correlation with chemical measurements. Proceedings of the 1984 conference sugar process research, New Orleans, Louisiana, USA, pp 22–52
63. Awad A, Chen AC (1993) A new generation of sucrose products made by cocrystallization. Food Technol 47:146–148
64. Sugar Producer Magazine, Member Associations, Oct 2010, pp 20–22
65. Jensen AS (2003) Steam drying of beet pulp and bagasse. Int Sugar J 105:83–88
66. Tjebbes J (1988) Utilization of fiber and other nonsugar products from sugarbeet. In: Clarke MA, Godshall MA (eds) Chemistry and processing of sugarbeet and sugarcane. Elsevier, Amsterdam, pp 139–145
67. Sargent D, Philip EJ, Cubero MTG (1998) Lime reduction in juice purification. Proceedings of the sugar processing research conference, Savannah, Georgia, USA, pp 21–39
68. Tebble I (2002) Use of membrane filtration to reduce lime usage in purification. Proceedings of the sugar processing research conference, New Orleans, LA, pp 103–111
69. Schrevel G (2009) Beet thick juice degradation during storage. Zuckerind 134:9–13
70. Godshall MA, Vercellotti JR, Triche R (2002) Comparison of cane and beet sugar macromolecules in processing. Int Sugar J 104:228–233
71. (1998) Molasses desugaring. In: van der Poel P et al (eds) Sugar technology, beet and cane sugar manufacture. Verlag Dr. Albert Bartens KG, Berlin, Chapter 17, p 939
72. Kearney M (1990) Simulated moving-bed technology applied to chromatographic recovery of sucrose from sucrose syrups. Proceedings of sugar processing research conference, San Francisco, California, USA, pp 291–304
73. Paananen HA (1996) Trends in the chromatographic separation of molasses. In: Clarke MA (ed) Separation processes in the sugar industry. Proceedings of a workshop. Sugar Processing Research Institute, New Orleans, LA, pp 100–119
74. Anon (1998) BMA Information, 36/1998, BMA AG, Braunschweig, GERMANY, pp 8–11
75. Kearney M, Kochergin V (2001) Chromatographic applications in the cane sugar industry. Proc Int Soc Sugar Cane Technol 24:11–15
76. Bates F (1942) Polarimetry, saccharimetry and the sugars. Circular C440 of the National Bureau of Standards, pp 767–779
77. Theisen K-H, Senkowski K, Diringer T (2001) Concentration measurement for continuous sugar crystallization processes. Paper presented at 31st American Society of Sugar Beet Technologists, Vancouver
78. Anon (2010) Sugar and sweetener: data tables, Table 1—world production, supply and distribution, centrifugal sugar. Economic Research Service, USDA, update 1 June 2010. http://www.ers.usda.gov/briefing/Sugar/Data/data/.htm
79. Haley S, Bolling C (2005) Brazil: the emerging giant of the global sugar industry. Sugar and Sweetener Outlook, SSS-242. Economic Research Service, USDA, 28 Jan 2005, pp 21–30
80. F.O. Lichts world sugar statistics, Sugar per caput use, F.O.Licht GmbH, pp 56–57, 2005
81. Sugar and Sweetener Yearbook, Table 20a, US sugar deliveries for human consumption, Economic Research Service, USDA, updated 10 Feb 2010. http://www.ers.usda.gov/Briefing/Sugar/Data/data.htm
82. Hickson JL (1979) The potential for industrial uses of sucrose. In: Birch GG, Parker KJ (eds) Sugar: science and technology. Applied Science, London, pp 151–180
83. Kollonitsch V (1970) Sucrose chemicals. The International Sugar Research Foundation, Washington, pp 17–105, 173–221
84. Lichtenthaler FW (ed) (1990) Carbohydrates as organic raw materials. VCH, Germany
85. Blommel PG, Keenan GR, Rozmiarek RT, Cortright RD (2008) Catalytic conversion of sugar into conventional gasoline, diesel, jet fuel, and other hydrocarbons. Int Sugar J 110:672–679
86. Godshall MA (2001) Future directions for the sugar industry. Int Sugar J 103:378–384
87. Buzzanell PJ. U.S. corn sweeteners: recent trends and future prospects. F.O. Lichts World Sugar and Sweeter Yearbook, F.O. Licht GmbH, 1997/98, pp D20–D34

88. Sugar and sweetener situation and outlook report, SSS-232. Economic Research Service, U.S. Department of Agriculture, Sept 2001
89. Kooreman JA (1955) Physical and chemical characteristics of enzyme converted syrup. Manufac Confect June:35–90
90. MacAllister RV (1979) Advances in carbohydrate chemistry and biochemistry. In: Tipson RS, Horton D (eds), vol 36. Academic, New York
91. Olsen HS (2002) Enzymes in starch modification. In: Whitehurst RJ, Law BA (eds) Enzymes in food technology. CRC Press, Boca Raton, FL, pp 200–228
92. Renewable Fuels Association. http://www.ethanolrfa.org/pages/ethanol-facts-trade
93. Godshall MA (1988) The role of carbohydrates in flavor development. Food Technol 42:71–72, 74, 76, 78
94. Godshall MA (1990) Use of sucrose as a sweetener in foods. Cereal Food World 35:384–389
95. Calorie Control Council. http://www.caloriecontrol.org/erythritol.html
96. Embuscado ME (2000) Erythritol, a unique low-calorie alternative sweetener. 2000 IFT annual meeting, Abstract No. 1 72-2
97. Roh HJ, Kim P, Park YC, Choi JH (2000) Bioconversion of D-galactose into D-tagatose by expression of L-arabinose isomerase. Biotechnol Appl Biochem 31:1–4
98. Chaen H (1997) Production of trehalose by a novel enzymatic method and its applications. J Appl Glycosci 44:115–120
99. Ennen S (2001) Trehalose gets GRAS. Food Process August:44
100. Anon (2001) Sugar and sweetener: background. Economic Research Service, USDA, 10 May 2001. http://www.ers.usda.gov/briefing/sugar/background.htm
101. Sustainable sugarcane and sugar production technology. In: Yang-Rui L, Solomon S (eds) Proceedings of 2004 meeting. China Agriculture Press, Beijing, Nanning, China, Nov.29–Dec 2, 2004, 764 pp
102. Suhr M, Schulze BC (2003) Technology in use in the modern beet sugar factory. Int Sugar J 105(52–58):69–70
103. Burnquist WL (2005) Cane genome: what is the direction of research?" (Genoma da cana: Qual o Rumo das pesquisas?). In: Portuguese, paper presented at 26th annual meeting of encontros Fermentec, Sao Pedro, Brazil, 14–18 March 2005; Proc. on CD
104. Clarke ME (1991) Cane sugar process and byproduct streams. Paper presented at Iowa biotechnology consortium, San Diego, March 1991
105. F.O. Lichts world sugar statistics. F.O. Licht, GmbH, Sugar: World production by country, F.O. Licht GmbH, pp 17–21, 2010
106. Corn Industries Research Foundation (1965) Corn syrups and sugars. Corn Industries Research Foundation, Washington, DC

36

Soap, Fatty Acids, and Synthetic Detergents

Janine Chupa, Steve Misner, Amit Sachdev, Peter Wisniewski,
and George A. Smith

ABS	Alkylbenzene sulfonate
AE	Alcohol ethoxylate
AES	Alkyl ether sulfonate
AOS	Alpha olefin sulfonate
APE	Alkyl phenol ethoxylate
APG	Alkyl polyglycoside
AS	Alkyl sulfate
BAB	Branched alkylbenzene sulfonate
BAC	Benzalkonium chloride
CAPB	Coco amido propyl betaine
CPC	Cetyl pyridium chloride
CTAB	Cetyl trimethyl ammonium chloride
HEQ	Hydroxyethyl quat
LAS	Linear alkylbenzene sulfonate
MEE	Ethoxylated methyl ester
MES	Methyl ester sulfonate
NPE	Nonylphenol ethoxylates
PE	Phosphate ester

Soap and Fatty Acids

Introduction

The origin of the word “soap” is traced to sacrificial Mount Sapo of ancient Roman legend. The mixture of fat and wood ashes that reacted to form soap was carried by rain to the banks of the Tiber River and was found as a clay deposit useful for cleaning clothes [1].

The boiling of fats with ashes was recorded as early as 2500 B.C. Commercial soap-making was a widespread art in the Middle Ages in Europe. The invention of the soda ash process by LeBlanc in 1791, and the discovery by Chevreul in 1811 that soap was composed of a mixture of fatty acids paved the way to modern soap-manufacturing processes [1, 2].

Chemistry

The modern definition of soap relates to cleansing agents derived from fats, oils, and other fatty derivatives: the soaps are alkali and ammonium salts of fatty acids containing 8 to 22 carbon atoms. These fatty acids are generally a mixture of saturated, unsaturated, and polyunsaturated moieties:

$$
\begin{gathered}
\mathrm{CH_3(CH_2)_{10}COOM}\text{: saturated soap}\\
\mathrm{CH_3(CH_2)_6CH_2CH{=}CHCH_2(CH_2)_6COOM}\text{:}\\
\text{mono-unsaturated soap}\\
\mathrm{CH_3(CH_2)_3(CH_2CH{=}CH)_2CH_2(CH_2)_6COOM}\text{:}\\
\text{di-unsaturated soap}\\
\mathrm{CH_3(CH_2CH{=}CH)_3CH_2(CH_2)_6COOM}\text{:}\\
\text{poly-unsaturated soap}\\
[\mathrm{M = Na, K, (HOCH_2CH_2)_3NH^+}]
\end{gathered}
\tag{36.1}
$$

In general, two types of chemical reactions are utilized in the manufacture of soap: the saponification of triglycerides (fats and oils) and the neutralization of fatty acids (which themselves are produced from the triglycerides by a variety of methods, most notably splitting or hydrolysis of fats and oils with steam under high pressure). Sodium hydroxide (the predominant alkali employed in the manufacture of soaps), potassium hydroxide, sodium carbonate, and triethanolamine are the most commonly utilized alkaline moieties in these processes (Fig. 36.1) [3, 4].

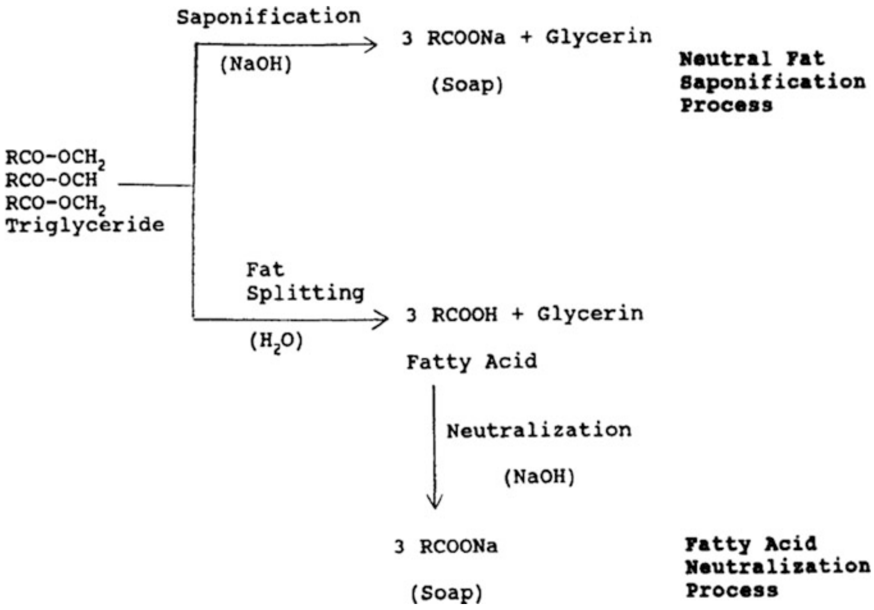
In recent years, soap manufacture by an alternate route, the saponification of fatty methyl esters, has been under development, most notably in Japan (Lion Corporation) and Italy (Ballestra). The fatty methyl esters are obtained from the methanolysis of triglycerides; inorganic alkali, quaternary ammonium salts, and enzymes (lipase) have

J. Chupa (✉) • S. Misner • A. Sachdev • P. Wisniewski
Colgate Palmolive Company, Piscataway, NJ, USA
e-mail: Steve_Misner@colpal.com; peter_wisniewski@colpal.com

G.A. Smith
Huntsman performance products, Austin, TX, USA

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_36, © Springer Science+Business Media New York 2012

Fig. 36.1 The chemistry of soap manufacture



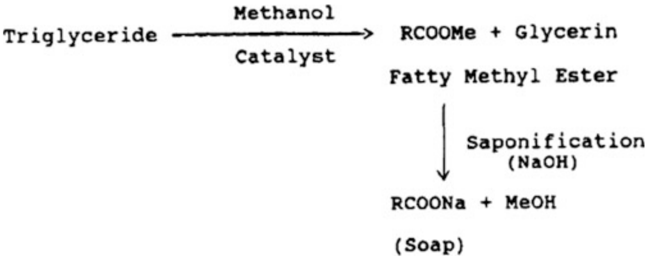


Fig. 36.2 Fatty methyl ester process for soap manufacture

been used as catalysts for methanolysis in commercially practiced processes (Fig. 36.2). The methyl ester process for soap manufacture is typically more capital and cost intensive vs. the more traditional fats/oils saponification and fatty acid neutralization (FAN) processes.

Manufacturing Technology

Several designs of commercial equipment, based on the two processes described in Fig. 36.1, are available for the manufacture of soap.

Neutral fat saponification	Fatty acid neutralization
Kettle processes	Mazzoni SSCT and Binacchi CHBS
Semi-boiled	Alfa Laval
Boiled	Sato
Modern processes	
Mazzoni SCNT-N, SCT	
Dial Saponiflex	
Binacchi CSWE, CSFA	
Alfa Laval Centipure Hitachi	

A critical review of these processes is beyond the scope of this chapter. However, the processes themselves are described in a subsequent section.

Raw Materials

The triglycerides (triesters of fatty acids with glycerin) are the basic raw materials for the manufacture of soap utilizing a variety of processes. These triglycerides, as a renewable resource, occur widely throughout the plant and animal kingdoms.

Soap products in the United States are produced almost exclusively from tallow (beef) fat and palm kernel oil or their fatty acid derivatives. The utilization of palm oils, and other all vegetable oil soaps, along with palm derivatives for soap manufacture is more commonly practiced in many other parts of the world, especially in Asia. More recent trends in lauric oil pricing have driven the purchase and use of palm kernel oil and palm kernel olein as cost-effective alternatives to the previously more commonly used coconut oil. The frequent substitution of coconut oil with palm kernel oil is done primarily for cost optimization of soap-manufacturing feedstocks. More recently, concerns have been escalating regarding the environmental impact of palm oil usage. Many manufacturers have active programs and have joined industry coalition groups working to ensure that the palm oil used in their products are grown and harvested in a sustainable manner. Among other triglycerides, cottonseed oil, rice-bran oil, castor oil, neem oil, sunflowerseed oil, fish oil, and olive oil have also been used as indigenous feedstocks in many developing nations. Palm stearin, the by-product of palm oil fractionation, is increasing in use as a substitute for tallow outside of tallow-producing markets. Depressed pricing in 2000 and continuing fears of Bovine Spongiform

Table 36.1 Fatty acid distribution and analytical characterization of soap base fatty acids

Fatty acid distribution	Coconut	Stripped coconut	Palm kernel	Palm	Tallow
Caprylic ($C_8H_{16}O_2$)	7	1	4		
Capric ($C_{10}H_{20}O_2$)	6	1	5		
Lauric ($C_{12}H_{24}O_2$)	48	55	50		
Myristic ($C_{14}H_{28}O_2$)	19	22	15	2	2.5
Palmitic ($C_{16}H_{32}O_2$)	9	11	7	42	27
Stearic ($C_{18}H_{36}O_2$)	2	2	2	5	20
Oleic ($C_{18}H_{34}O_2$)	8	7	15	41	42
Linoleic ($C_{18}H_{32}O_2$)	1	1	1	10	5
Linoleic ($C_{18}H_{30}O_2$)					1
C_{20}/C_{22} isomers			1		2.5
Characterization data					
Titer, C	26	29	25	40	40
Iodine value (IV)	10	5	18	55	55
Color, Gardner, max.	5	1			5
Acid value (AV)	270	255			205
Saponification value (SV)	270	255			205
Color, Lovibond, $5-\frac{1}{4}''$	10Y/1R	4.5Y			
		0.6R			

Encephalopathy (BSE) or Mad Cow disease have driven this use trend in Europe. Fatty acid composition and analytical characterization data of some of the key fats and oils utilized for soap production are given in Table 36.1.

Most of the fats and oils utilized for soap manufacture require a combination of steps: degumming, deodorization, refining, and bleaching, which are further discussed in the section on fatty acids.

For both saponification and neutralization reactions, sodium hydroxide is the most commonly used base. In some special situations, such as the manufacture of transparent and translucent soaps, potassium hydroxide and triethanolamine are also employed.

Functional Properties of Soap

Almost as a rule, the solubility of soap in water increases with an increase in the size of the monovalent cation (base); however, an increase in the size of a divalent cation (Mg^{+2}, Ca^{+2}) results in a decrease in the solubility. An increase in the chain length of the fatty acid moiety of soap results in a decrease in the solubility of soap, whereas the presence of unsaturation results in an increase in its solubility. In general, an increase in the water solubility is accompanied by an increase in the softness and use-up properties of soap. Adequate solubility results in good performance in soft water, although in hard water, lather decreases due to the formation of insoluble Ca^{+2} and Mg^{+2} salts. Cleansing bars formulated with various levels of synthetic detergents (combars or syndets) provide improved performance in hard water conditions.

Consumer-acceptable lather requires the consideration of a variety of lathering attributes, such as the speed with which lather is generated, volume, quality (i.e., loose or creamy), and the stability of the lather. In general, the use of fatty acids of C_{10}-C_{12} chain length provides voluminous lacey foam and fatty acids of longer chain length, from C_{16} to C_{18}, contribute to a richer, creamier foam.

Thus, a combination of lauric oil (C_{10}–C_{12}) and tallow (C_{16}–C_{18}) fatty acids is most suitable from the viewpoint of providing a balance in lathering attributes to deliver consumer-acceptable lather. Commonly accepted ratios of these soap feedstocks are 80–90% tallow and 10–20% lauric oil for standard purpose bars, and 60–75% tallow and 25–40% lauric oil for the higher lathering bar-soap products. Cleansing bars formulated with free fatty acids (superfatting) improves the quickness of lather generation and provides open bubbly foam. Commonly accepted levels of superfatting range from 1 to 7% for various types of bar soaps. Various methods are employed to evaluate bar lather attributes [5].

Manufacturing Processes

The following describes neutral fat/oil saponification and FAN as is commonly available from the two main equipment manufacturers (Binacchi and Mazzoni). These processes revolve around a continuous stirred tank reactor (CSTR) type of soap making reactor and neutralizer, and subsequent intermediate cooling and separating stages to produce soap and lye of a desired composition. There are many alternate processes available in the world today which, while operating with the same general theories, branch out to various plug flow reactions, improved efficiency high solids based reactors, optimized equipment and operating

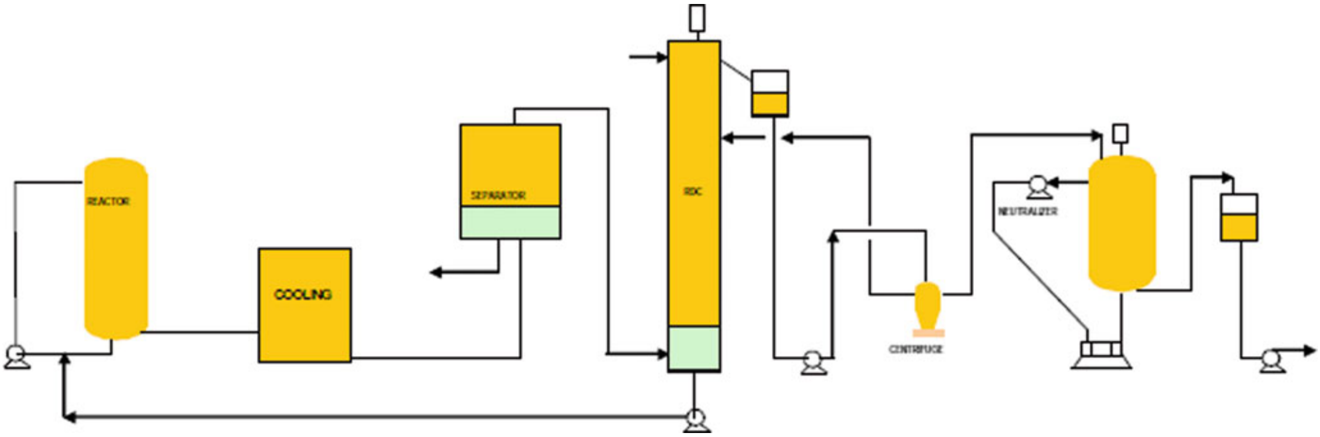


Fig. 36.3 Flow diagram of the general continuous saponification process with generic cooling area and standard neutralizer (Courtesy of Colgate Palmolive)

condition configurations, and combination processes which link the soap making and drying processes. This chapter provides the reader with knowledge of the more standard saponification process as it relates to the standard Binacchi and Mazzoni soap making designs.

Neutral Fat Saponification. In the saponification of triglycerides with an alkali, the two reactants are rather immiscible. Because saponification is a bimolecular nucleophilic substitution (SN2), the kinetic rate enhancement is achievable by both a high-shear mixing and an increase in the reaction temperature during the processing. Increasing the pressure of the reaction assures that vaporization does not occur at the high reaction temperatures and helps to accelerate conversion of the fatty feedstocks and of alkali to soap as well. The saponification reaction is self-catalyzed to some extent: the formation of soap product affects the emulsification of the two immiscible reactants, causing an acceleration of the reaction rate by providing increased surface area for the saponification reaction to occur.

The amount of alkali required to saponify a triglyceride blend is calculated by the following equations:

$$\begin{aligned}&\text{Triglyceride} + 3\text{NaOH} \rightarrow 3\text{ROONa} + \text{Glycerin}\\&\text{NaOH} = \text{weight oil} \times [\text{SV} \times 0.000713]\\&\quad \times 100/\text{NaOH}(\%)\ \text{in solution}\\&\ = \text{weight oil} \times \text{SV}\ (1\ \text{g}/1{,}000\ \text{mg})\\&\quad (\text{MW NaoH}/\text{MW KOH})(100\%/\%\text{NaOH})\end{aligned} \tag{36.2}$$

where SV = saponification value (mg KOH/g triglyceride) and MW = molecular weight. In general, the most commonly used soap feedstocks have these saponification values:

	(g NaOH/100 g triglyceride)
Beef tallow	14.09
Coconut oil	18.32
Palm kernel oil	17.81
Palm oil	14.19
Palm stearin	14.26

The use of these values simplifies the calculation of the stoichiometric alkali equivalent, for example:

$$\begin{aligned}&100\ \text{g beef tallow} \times 14.09\ \text{gNaOH}/100\ \text{g tallow}\\&\quad \times 1\ \text{g}\ 50\%\ \text{NaOH caustic}/0.5\ \text{g NaOH}\\&\quad = 28.2\ \text{g}\ 50\%\ \text{NaOH caustic required}\end{aligned}$$

The process for a typical triglyceride saponification plant is shown as a flow diagram in Fig. 36.3 [6]. In general continuous soap making processes have replaced the kettle process in most large scale soap making operations. The driving forces for this change include economies in utilities, simplification of manpower training, elimination of operator bias in process control, and more consistent, high quality soap base. In most locations, the continuous process of choice is the continuous saponification of fat/oil (neutral fat) with caustic soda. It must be understood however, that for a given level of finished soap color and odor quality, the "natural" color and odor quality of fats and oils to be continuously saponified must be better than the color and odor of materials for kettle saponification. Also, the saponification color must be lower. This is because in continuous saponification a significant portion of the oil and alkali soluble color components are not removed, representing roughly 5–10% of a typical fats/oils charge to saponification in the kettle process. This quantity of poorer quality soap is not reprocessed and returned as a cleaner base to the primary soap making system. In contrast, a small quantity (1–2%) of poorer quality soap is degraded from the kettle process as slurry soap thereby removing a greater amount of these color bodies as well as odor in the full boiled kettle process.

Also, smooth continuous saponification processing requires frequent standard analyses of all raw materials and, in addition, a number of simple analytical determinations of minor ingredients in raw materials.

Saponification Reactor

As the title of the process step implies, the purpose of the step is to saponify fat/oil. Fat/oil, caustic soda, aqueous salt solution (brine), water and Column Lye from a later process step, are metered into a pressurized reactor equipped with a recycle loop and a high-shear recycle pump. The point of injection of raw materials is at the pump suction, where they are immediately mixed with a large mass of recirculating reaction mass from the reactor, under conditions of elevated temperature and pressure, and high-shear mixing (from the pump). While the elevated pressure is important as it prevents vaporization, it is the elevated temperature and high shear mixing, combined with the effects of the recycling soap mass which cause the immediate homogenization and emulsification of the raw materials. This in turn results in the rapid saponification rate associated with the continuous soap making process.

Recycle takes place during the entire process. As long as the reactor and recycle loop are full, the metering of raw materials into the Reactor is continued as an equal weight of reaction mass is withdrawn and passed on to the next process step, Cooling.

The rate of addition (metering) of raw materials into the Reactor is governed by four requirements:

- That the fat/oil be almost completely saponified before it leaves the Reactor.
- That complete saponification takes place by the time the continuous process mass passes through the next process step, Cooling.
- That the Spent Lye by-product contains a minimum of alkali, and sufficient salt to keep the process mass in the Neat Soap/lye phase configuration, and to promote optimum separation of soap and Spent Lye during Separation.
- That the required soap production rate be met.

The design considerations of the saponification autoclave center around considerations for residence time, recycle rate, and mixing capability of the entire reactor design. Residence time and mixing efficiency determine the uniformity of the mixture exiting the reactor, as well as the required control efficiency and reaction time required to assure completeness of reaction and absorb various process deviations. The recycle ratio is studied to maximize the quantity of already reacted soap present during the saponification of the newly entering fats/oils and/or fatty acids as the reaction is autocatalytic, meaning that the more soap that is present, the faster the reaction can proceed.

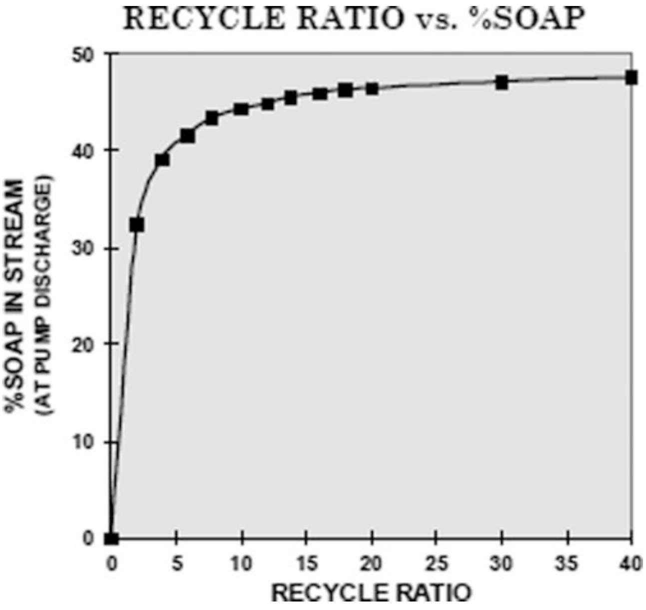


Fig. 36.4 Recycle ratio vs. per cent soap. This graph shows the change in the quantity of recycling soap in the reactor as compared to recycle ratio for a typical continuous saponification reactor (Courtesy of Colgate Palmolive)

A typical chart used to determine required recycle rate of a saponification reactor can be found in Fig. 36.4. This figure illustrates the dramatic effect of recycle ratio on emulsification, as measured in terms of % soap at the pump discharge of the recycle loop. Recycle ratios of about 13–25 have been found sufficient to achieve nearly complete saponification in the Reactor. Utilizing these design criteria, the small changes in process flow due to viscosity assure that the quantity of soap at the point of initial reaction can be held relatively constant.

Cooling

Because the temperature of the saponified soap/lye mixture is elevated, effective separation is made possible by passing this mass through a cooling stage and then to a static separator for separation of the unwashed soap from the spent lye solution.

The cooling process can be achieved in a number of ways depending on the desired process design. The most common in older plants is to pass the soap through a cooling vessel in which the soap is cooled utilizing water as the cooling medium. The most common form is the mixer type cooler where the soap is passed over a series of cooling coils while being constantly agitated. The water through these coils can be at various temperatures, and can have heat recovery so as to aid in reducing total energy required for process operation.

Improving on process optimization Mazzoni offers a flash cooler as its standard design which allows for process water to be flashed off from the saponification mixture. The theory

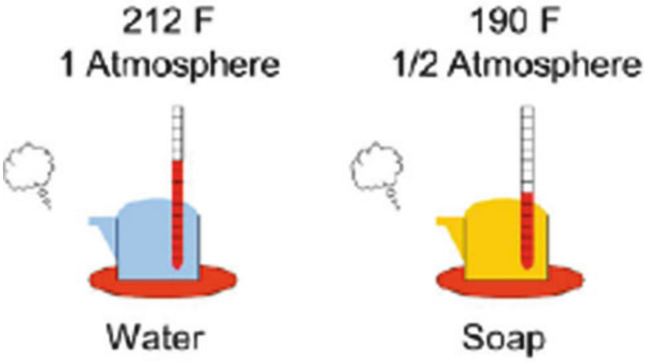


Fig. 36.5 Relative boiling point diagram showing boiling points for water at atmospheric pressure and soap at typical process pressure. The teapot provides reference to the constant temperature regardless of the amount of heat applied (Courtesy Colgate Palmolive)

behind the operation is like that of a pot boiling on the stove, where for a given pressure the soap will always boil at a constant temperature. At atmospheric pressure the pot of water on the stove will maintain 100°C regardless of how much water is in the pot, or how much heat is applied. At an absolute pressure near 0.5 atmospheres a flash cooler will maintain soap between 85 and 92°C with near constant temperature, an ideal condition for soap and lye separation. Exact pressures required will depend on solids and electrolyte content (Fig. 36.5).

The flashing of steam also increases the concentration of the components in the remaining soap and lye, including the glycerin. This allows the Mazzoni process mass balance to be altered so as to allow spent lye glycerin concentrations in excess of 40% glycerin where the solubility of soap in lye can often be lower, further increasing total process efficiency, reducing total required raw materials required per ton of soap produced as well as costs associated with subsequent glycerin recovery. The flashed steam can also be recovered to reduce water feed required, and can at the same time be used to preheat raw materials or other process streams to save on total energy requirements.

Another option to consider is a process design offered by Binacchi, which seeks to reduce total process capital by eliminating the cooler. This can be achieved by directly injecting cooled lye into the saponification mixture after it exits the saponification reactor. The lye can be sourced from the separator, which is the next process step. While making for a slightly bigger separator to accommodate the increased recycling lye flow, the reduction in capital can be quite attractive. The lye can be withdrawn from the separator and cooled, with a portion of the heat removed used for reheating a portion of the raw materials, further increasing process efficiency. Finally in eliminating the need for the cooling vessel Binacchi has eliminated the need for the cooler discharge pump and has reduced process footprint, further reducing capital and energy costs.

Soap Separator

No matter which cooling route is chosen, the process mass exiting the soap cooling process will require separation of the soap from the spent lye. This is accomplished in a soap separator.

The purpose of this process step is to provide sufficient residence time and sufficiently lowered process flow velocities to promote the gravity separation of the cooled process mass into an upper layer of crude Neat Soap and a lower layer of Spent Lye.

Static separation of the soap mass from the cooling process is accomplished in a simple insulated, cylindrical tank. However, because of the small density difference between Neat Soap and Spent Lye, approximately 0.97 vs. 1.05 g/cc, phase separation during this process step is far from complete; from 15 to 25% of the mass of the soap exiting the separation process is entrained spent lye. This entrainment is not a serious problem because in the next process step, Glycerin Extraction, entrained lye combines with lye passing down through the extraction column, and is recycled back to the saponification reactor as part of the Column Lye.

The Spent Lye that does separate from the soap mass must be clear, that is, free of visible soap, suspended or otherwise.

The soap mass, crudely separated from lye in the separation process is then fed to Glycerin Extraction. The clear separated Spent Lye is sent to glycerin recovery.

The soap separator is a vessel where the soap and lye mixture enter near the center of the vessel, with soap exiting from the top, and lye being withdrawn from the bottom. Spent lye discharge rates are then controlled so as to maintain a constant-level soap/lye interface, which provides for a constant, specified flow rate of soap from the top of the Separator to a holding tank at a specified temperature, and on to the Glycerin Extraction Column. Interface position is generally monitored using a sight glass with some processes also expanding to electronic level detection for improved process control.

The separator itself is most commonly an atmospheric tank with dimensions determined based on soap velocity and residence time for the soap and lye phases. In this design the soap and lye are fed to level control tanks to aid in maintaining a stable process, and in keeping discharge rates constant for both soap and lye. Alternatively the separator can be a pressurized device where the soap continues to separate as before, but the driving force for pushing the soap and lye is from a pump such as the Mazzoni cooler discharge pump, or the dosing pumps feeding the Binacchi reactor. In this case the pressure in the separator eliminates the need for the constant level tanks, the soap discharge pump, and in many cases the spent lye discharge pump as well.

Significant to the separation process in the temperature of the process mass to be separated, and the electrolyte composition. Lower temperatures generally increase separation efficiency, but usually required increased operating cost, and can in turn increase the viscosity of the process mass, making flow and pocketing of the lye within the soap an issue.

The electrolyte content (or content of salt and alkali) in the soap and lye is kept relatively high (as compared to the rest of the process) from the saponification stage up until separation. Electrolyte is defined in reference to a lye limit concentration (LLC), and the percentage above the LLC that the spent lye electrolyte is maintained. The Reactor Lye is usually tested to make sure the correct electrolyte content is maintained for good separation and flow, but will may analyze slightly higher than target because water flashes off analysis samples taken as the reactor is at elevated temperature.

From a practical viewpoint, the required total electrolyte levels of Spent Lye are functions of:

- The viscosity of the soap mass in the Reactor.
- The soap content of Spent Lye.
- The clarity of Spent Lye.

Changes in the tallow/lauric oil ratio affect the exact electrolyte requirement but all soaps generally have an elevated viscosity and poor separation with too high or too low an electrolyte, with an ideal intermediate electrolyte level which maintains the resulting wet soap in the Neat or Wax phase.

Also important is a general thought that Spent Lye should be as low as possible in caustic soda to minimize the quantity of acid required to neutralize the lye during glycerin recovery, while at the same time assuring near complete reaction. Salt concentrations therefore are the only electrolyte variable that can be adjusted to minimize the concentration of soap in the Spent Lye and to clarify it “to the appearance of fresh-brewed tea.” The color intensity of this lye (and its glycerin content) is a function of Lye Bulk (and of course raw material quality, especially fat/oil).

Soap Extraction

The purpose of this process step is to extract glycerin from the soap mass that is fed continuously from the separation process. The glycerin extracted is the glycerin from the soap which is also the glycerin that did not partition into the Spent Lye during the previous process steps.

Extraction is accomplished in a countercurrent extraction column in which the soap mass from separation and Fresh Lye is brought into intimate continuous countercurrent contact. The soap mass enters the bottom of the column while Fresh Lye enters the top. Droplets of lye trickle downward through a continuous phase of soap, become enriched in glycerin (more soluble in lye than in soap), and form a continuous lye phase at the bottom of the column. The lye from the bottom of the column, Column Lye, is then metered back to the saponification reactor so as to maintain both column interface level and reactor composition.

The soap mass that collects at the top of the column overflows into a constant level tank from which it is pumped to Centrifuging.

Also during Glycerin Extraction the lye-soluble color and odor bodies are extracted from the soap mass.

The column lye, as stated previously, is a mixture of the lye that has passed through the soap in the column, as well as the entrained spent lye which has entered the bottom of the column with the separated soap. The entrained lye has a density greater than the soap in the column, and also is contacted by the lye falling down the column, preventing the entrained spent lye from rising through the column.

Electrolyte composition for the extraction process is very important. The concentrations of salt, caustic soda, and glycerin in the Fresh Lye provide for:

- The final adjustment of the alkali and salt concentrations of the soap so as to meet desired soap electrolyte composition.
- The concentration of caustic soda required in the Column Lye (the lye exiting the bottom of the Glycerin Extraction Column and pumped back to Saponification in the Reactor).
- The glycerin extraction efficiency of the Fresh Lye in the Extraction Column (it must be noted also that the glycerin content of the Fresh Lye must be low enough so that the soap in equilibrium with the lye can satisfy its glycerin specification).
- Separation efficiency at the Centrifuge, where lye is separated from soap.

The concentrations of salt and caustic soda in this process step also must be such that the soap phase maintains its required phase configuration—Neat Soap.

With attention paid to all this detail, a multi-equilibrium, glycerin extraction equilibrium process is established within the Column between the two phases (soap and lye) as they pass through each other—the glycerin-laden crude soap from Static Separation passing up through the Column, and the Fresh Lye passing down.

The lye flowing down through the Column joins with the entrained Spent Lye from soap entering the bottom of the Column and exits as Column Lye. This lye flows back to Saponification. This resulting final Column Lye mixture is elevated in glycerin, and as the continuous saponification plant has a glycerin by-product stream, the glycerin rich Column Lye stream is sent to the reactor. Within the reactor the reaction of the fats/oils produces added glycerin further enriching the glycerin content of the lye. Also important to note is that column lye is also termed Half lye, where the caustic content is at an intermediate level between the Fresh and Spent Lye. The Column Lye or Half Lye is sent to the reactor to also utilize the remaining alkali in the lye to saponify a portion of the incoming oils.

The soap at the top of the column first passes through a small separation zone within the column to allow a portion of the lye to separate from the soap in a manner similar to the separation process, to minimize the quantity of lye sent to the next step, centrifugation. The soap, having passed through the separation zone at the top of the column, now stripped of most of its glycerin, flows from the top of the extraction column at specified temperature, composition, and rate to a constant level holding tank and on to Centrifugation. This lye and soap are at near-equilibrium with respect to salt, caustic soda, and glycerin.

While temperature of the separation process is governed by the previous soap cooling process, temperature within the column is controlled by two factors. The first is the incoming soap temperature and the second is by the temperature of the incoming Fresh Lye to the top of the extraction. The Fresh Lye controls not only the temperature of the soap at the top of the extraction column, but also the temperature of the soap at the next process, Centrifugation, where input soap temperature is a critical parameter in maintaining proper machine performance.

Flow rates to and from the Column are controlled to maintain a constant-level soap/lye interface near the bottom of the Column. This provides for a constant, specified overflow rate of soap from the top of the Column for optimum centrifuge performance, and a constant specified flow rate of Column Lye from the bottom of the extraction column to minimize required electrolyte adjustments at the Reactor, and to maintain a constant desired spent lye glycerin concentration.

Centrifugation

The purpose of this process step is to separate as much entrained Fresh Lye as possible from the soap mass exiting from the extraction column, thus producing Neat Soap ready for Neutralizing or direct consumption by the soap plant. This is accomplished via the elevated gravitational forces produced in a centrifugal separator.

The separated Centrifuge Lye exiting the centrifuge process is returned back to the extraction column where it enters near the top of the column. The centrifuged Neat Soap is sent for Neutralizing. Centrifugal separators remove entrained lye from the soap directly after extraction. This centrifuging step, accomplished from one or several centrifuge units operating in parallel, reduces moisture and electrolyte content from the soap that is required in downstream processing stages.

Pressures within the centrifugation system are manipulated to adjust for interface position within the machines as well as for downstream pressure so as to optimize separation efficiency.

Important factors in the centrifugation process include:

- A controlled electrolyte level in Fresh Lye flowing into the extraction column, and thus the electrolyte level of the extracted soap flowing from the extraction column into the Centrifuge.
- Constancy of temperature and feed rate of soap into the Centrifuge.
- Centrifuge operations within prescribed parameters to maintain a stable soap–lye interface utilizing correct and constant soap and lye discharge pressures.

Given equilibrium in centrifuge operations, constant-composition Centrifuge Lye flows from the Centrifuge and back to the Extraction Column, and high quality centrifuged soap flows out and on to Neutralization.

Neutralization

In the production of most toilet soap, residual-free alkali in Neat Soap from Centrifuging is adjusted downward continuously in a CSTR type neutralization reactor equipped with a recycle loop. This reactor is generally designed to operate at or below 1 bar pressure, and at 80–90°C, so as to produce soap at a temperature suitable for atmospheric operations or soap storage.

The neutralizing agents are various, all capable of reacting with the residual alkali, and are often combined with a preservative to also start the preservation of the neutralized soap as early as possible to minimize degradation of the newly formed soap. The degree of neutralization is determined by measuring pH in-line or alkalinity by direct titration of a representative sample. Ultimately the need is to determine the concentration of caustic soda and salt. Salt concentration is also determined to make sure that the fresh lye is providing for a soap/lye equilibrium satisfactory for properly satisfying specifications for salt in soap drying, and that potential changes in brine to not subtract from the separation efficiency of the centrifugation process.

This finished base soap, once confirmed as meeting desired composition is then dried (typically on vacuum spray-type dryers) to produce soap as pellets (70–82% total fatty matter), ready for processing into finished product on finishing lines.

Soap Drying. The soap mass obtained after the completion of saponification (neat soap) is most commonly dried nowadays on vacuum spray dryers (Fig. 36.6). The moisture content of soap is thus reduced from 30 to 35% in neat soap to 8–18% in soap pellets (soap flakes or chips are produced on APV-type dryers, after milling). A variety of vacuum spray-dryers, from single-stage to multistage designs, are available from several manufacturers. The operation of a single-stage

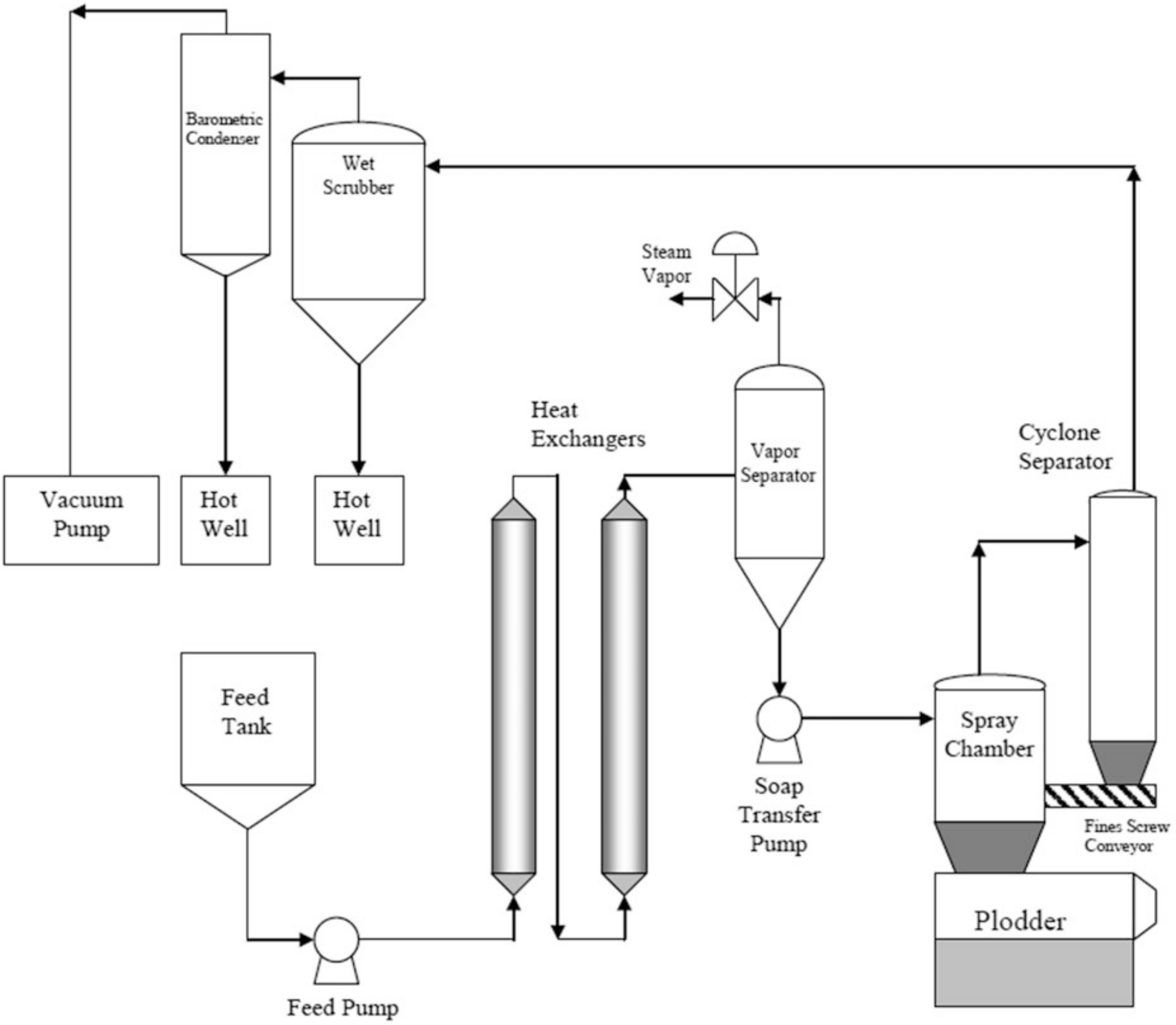


Fig. 36.6 Typical soap dryer configuration including vapor separator and wet scrubber (Courtesy Colgate Palmolive)

vacuum spray-dryer involves the pumping of neat soap at 80–90°C through a shell-and-tube heat exchanger where the soap is heated by high-pressure (6–10 bar) steam passing through the outside tube in a countercurrent manner. The soap is preheated to 135–155°C and sprayed onto the walls of a cylindrical vacuum chamber through either a stationary or revolving nozzle. The thin layer of dried and partially cooled soap deposited on the walls of the vacuum chamber is removed by rotating scrapers and falls to the bottom of the chamber onto a discharge plodder. This plodder extrudes the dried soap mass in the form of noodles or pellets.

The excess moisture from the incoming soap is either flashed off directly in the spray chamber of the dryer or first sent to a vapor liquid separator following the final soap heat exchanger. In the case of the vapor liquid separator the steam generated by the heating of the soap is removed under controlled pressure prior to the spray chamber so as to reduce the quantity of fines generated, the load on the barometric condenser, as well as to potentially create a valuable source of steam that can be used to drive the dryer booster, or provide heating to an intermediate process stream. In certain circumstances the steam can also be atmospherically vented to gain the fines reduction and barometric condenser benefits if the steam generated is not needed in the booster operation. With or without the vapor liquid separator the heated soap will be sent to the spray chamber, where it will be under 40–60 mmHg vacuum and the remaining moisture will be removed as vapor. This vapor will be passed through one or two cyclone separators to remove the majority of soap fines produced, and then potentially to a wet scrubber for added fines removal, and then on to the booster (if installed) and barometric condenser system. Proper control of moisture, temperature, vacuum, flow rate, vapor liquid separator level, and pressure if installed, spray chamber entry pressure and scraper function ensure the drying of acceptable quality soap pellets. Typically these soap pellets are conveyed pneumatically to storage bins/silos that act as surge for the soap finishing process. Soap storage silos are constructed of stainless steel and have angled bottoms that direct soap pellets toward centrally located single or twin discharge screw conveyers for efficient transfer to the finishing lines.

Fatty Acid Neutralization. The hallmark of soap making by the continuous FAN process is simplicity relative to the complexity of soap making by the continuous saponification of fat/oil. Consider for example the fundamental steps of the two processes:

- In FAN (the Direct Neutralization process variant) there are three process steps: raw material dosing, neutralizing with caustic soda, and post-addition of minor finished soap ingredients.
- In fat/oil saponification, there are seven process steps: raw material dosing, saponification with caustic soda, cooling, static phase separation, glycerin extraction, centrifuging, and neutralizing.

Four of the fat/oil saponification steps have to do with the requirement to separate large quantities of by-product glycerin from the process mass—cooling, static phase separation, glycerin extraction, and centrifuging.

The continuous neutralization process mass need not be cooled to separate phases; the process mass is a single phase, Neat Soap. The process mass need not be allowed to quiescently separate into two phases, Spent Lye and Neat Soap; again, it is a single phase, Neat Soap. Large quantities of glycerin need not be extracted from the process mass and the process mass does not need to be centrifuged.

A consequence of the process simplicity of FAN is its contribution to the relative economics of soap making through reduced initial required capital and floor space.

There are two FAN processes, namely:

- The so-called Direct Neutralization Process in which fatty acids are neutralized with sodium hydroxide (caustic soda).
- The Carbonate Neutralization Process in which fatty acids are first partially neutralized with sodium carbonate and then the neutralization process completed with sodium hydroxide (caustic soda).

The chemical reactions of the above-mentioned processes are:

- Direct Neutralization:

$$\mathrm{RCOOH} + \mathrm{NaOH} \rightarrow \mathrm{RCOONa} + \mathrm{H_2O} + \mathrm{heat}$$

- Carbonate Neutralization:

$$\mathrm{RCOOH} + \frac{1}{2}\mathrm{Na_2CO_3} \rightarrow 70 - 85\%\ \mathrm{RCOONa} + \mathrm{CO_2} + \mathrm{H_2O} + \mathrm{heat}$$

$$\mathrm{RCOOH} + \mathrm{NaOH} \rightarrow 30 - 15\%\ \mathrm{RCOONa} + \mathrm{H_2O} + \mathrm{heat}$$

where NaOH for reaction = weight of fatty acid × AV× 0.000713

with AV (acid value of fatty acid blend) = mg of KOH required to neutralize 1 g of fatty acid.

AV = acid value (mg KOH/g of fatty acid), and the remaining portion of the equation being derived as in (36.1).

The average molecular weight of a fatty acid blend is calculated from the following equation:

$$\begin{aligned}\text{MW fatty acid} &= 56.1 \times 1{,}000/\text{AV}\\ &= (\text{MW KOH/AV}) \times (1{,}000\ \text{mg/g})\\ &\quad \times (1\ \text{mole KOH}/1\ \text{mole fatty acid})\end{aligned} \tag{36.2}$$

A typical Mazzoni FAN soap plant is illustrated in Fig. 36.7 [6]. The principal raw materials for the Continuous Neutralization processes are fatty acids, alkali (caustic soda or soda ash/caustic soda), and brine (salt and water). These raw materials must be of rather high quality because:

- The soap mass is not "fitted" as in kettle saponification—to remove a quantity of lower quality soap from the in-process soap mass. There is no lye for "fitting."
- The soap mass is not in contact with lye as in both kettle saponification and continuous saponification—to remove water-soluble impurities from the in-process soap mass.

The impurities that go into the process in raw materials exit the process in the product, Neat Soap, ready for drying.

Direct Neutralization

In Direct Neutralization, caustic soda, brine, and any required additional water are premixed in-line and dosed into the high shear mixer simultaneously with fatty acid. This procedure produces heat (the heat of dilution of sodium hydroxide with water and the heat of neutralization of fatty acids with sodium hydroxide). Depending on raw material temperatures and required throughput rate, it may be necessary to preheat or precool one or more raw materials to maintain the process mass at 85–98°C during the entire short process. In both instances, only caustic soda/water mixtures should be directed through the heat exchanger. Brine should be added after heat exchanging to minimize heat exchanger corrosion (and required metallurgy) and fouling.

Neutralizing takes place in the high shear mixer and draft tube mixer shown in Fig. 36.7 With the positive displacement recycle pump, the equipment is a closed recycle loop neutralizer operating under sufficient agitation and recycle rate to ensure homogeneity in the process mass. Dosed raw

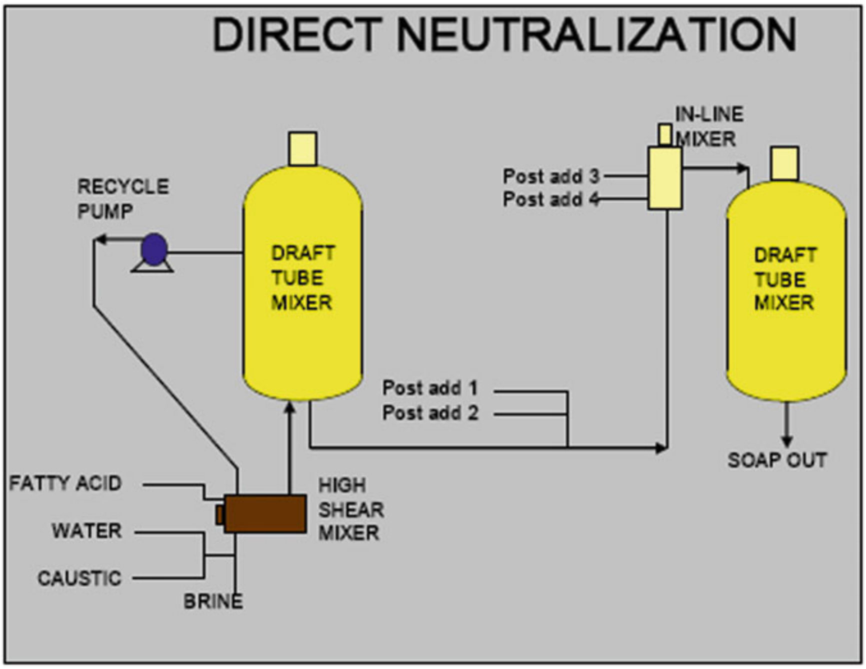


Fig. 36.7 Flow diagram of the general Direct Neutralization process with post-addition option and standard neutralizer (Courtesy of Colgate Palmolive)

materials enter the recycling process stream just prior to the shearing zone of the high shear mixer. Neutralization is essentially instantaneous.

Relative to the conditions for the previously described Continuous Saponification Reactor, operating conditions for Direct Neutralization are considerably less severe.

In summary, the paths of raw materials and process mass in the Neutralizer during this second process step are as follows:

- Raw materials are dosed into the recycling soap stream just prior to the shearing zone of the high shear mixer. Within the mixer the raw materials are almost instantaneously emulsified and the fatty acids and almost instantaneously neutralized by caustic soda.
- From the high shear mixer, the soap mass passes on to the draft tube mixer where neutralizing is brought to completion.
- From the draft tube mixer, the process mass moves on to a positive displacement pump for pumping back to the high shear mixer where it receives fresh raw material and is recycled through the several elements of the system—the high shear mixer, the draft tube mixer, and the positive displacement recycle pump.

A pH probe can be mounted in the recycle loop prior to raw material entry to monitor the alkalinity of the recycling soap mass. Alternatively, samples can be taken, usually just prior to raw material entry, and titrated with acid (i.e., phenolphthalein indicator). As necessary, adjustments are made in the quantity of either caustic soda or fatty acid being dosed into the system.

Neutralizing is brought to completion as the process mass passes through the neutralizing equipment in the above-described path. A typical Na_2O specification for process mass exiting the Neutralizer is 0.03–0.07%.

Carbonate Neutralization

The Carbonate Neutralization Process consists of five process steps—Raw Material Dosing (for Carbonate Neutralizing); Partial Neutralizing With Soda Ash; Carbon Dioxide Separation; Raw Material Dosing (for Caustic Soda Neutralizing); and Completion of Neutralizing With Caustic Soda; See Fig. 36.8 for a process flow diagram.

It might seem that, with five steps, the Carbonate Neutralization process could never be preferred over Direct Neutralization with only three steps. However, in some locations the availability, cost and/or purity of soda ash relative to caustic soda may make Carbonate Neutralization the preferred process. Appreciate however that operationally and with currently available equipment Carbonate Neutralization is relatively problematic. The viscosity of the process mass is the major issue. See later remarks on viscosity.

Soda ash alone cannot be used effectively in FAN processes for two reasons.

- First, after 70–85% of the fatty acid has been neutralized with soda ash, the viscosity of the process mass becomes so high that pumping the mass becomes difficult.
- Second, since fatty acids are weak acids and soda ash is a weak base, the neutralization reaction rate falls off dramatically as the process proceeds, making impractical a process based on soda ash alone.

Hence, the need for the fourth and fifth process steps of Carbonate Neutralization.

Raw material dosing equipment (proportioning pumps and mass flow meters) and dosing processes are similar to those for Continuous Saponification.

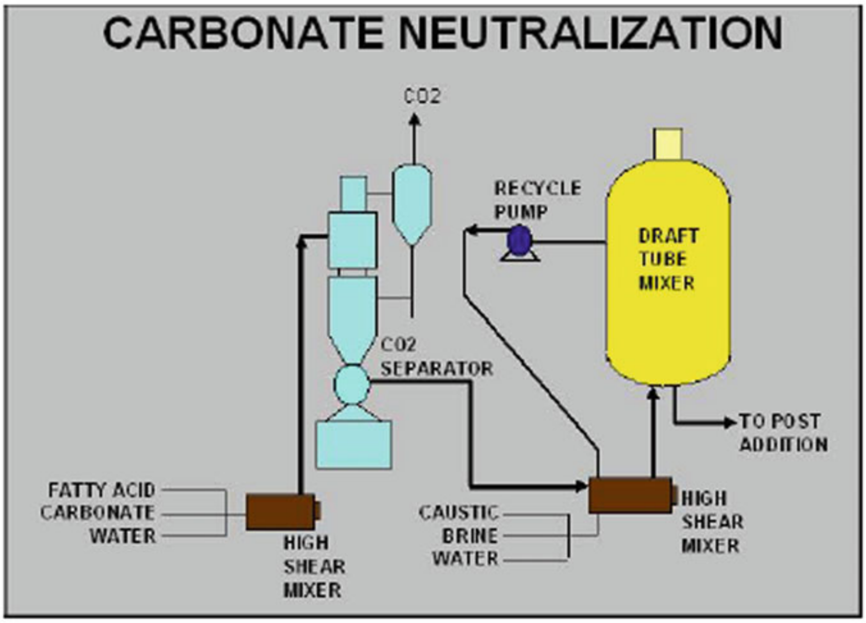


Fig. 36.8 Flow diagram of the general Carbonate Neutralization process with post-addition option and standard neutralizer (Courtesy of Colgate Palmolive)

Soda ash purchased in granular form is pre-dissolved in water to make a 20–22% solution. The aqueous solution of soda ash, fatty acid, and any extra required water are dosed into the high shear mixer at whatever temperature is required to maintain the neutralizer process mass at 85–98°C.

Partial Neutralizing with Soda Ash

As shown in Fig. 36.8 above, this partial neutralizing is accomplished in a high shear mixer, without recycle. The progress of the reaction is estimated by analysis of the process mass for unreacted fatty acid and by measuring pressure drop in the process mass transfer pipe between the high shear mixer and the CO_2 Separator. Alternatively, and crudely, it can be estimated by monitoring the pump speed required to transfer the process mass through the pipe at the required throughput rate.

The viscosity increase is very large, the result of the intimate dispersion of CO_2 into the process mass.

Viscosity can be reduced if necessary by diluting with water the soda ash solution being dosed into the Neutralizer or by reducing the targeted degree of Neutralization with soda ash. Both of these process modifications of course negatively affect production rate and efficiency.

As shown in Fig. 36.8, the continuous stream of process mass exiting the high shear mixer passes on to the CO_2 Separator.

Carbon Dioxide Separation

The Carbon Dioxide Separator is a tall, jacketed column equipped with a high-viscosity pump at the bottom. CO_2 is eliminated from the process mass and passes out of the top of the Column as the reduced-viscosity CO_2-free process mass moves downward in the Separator and is pumped from the bottom and onward to the second Neutralizer for completion of neutralizing with caustic soda. In some columns a screw impeller is mounted at the bottom of the Column to force process mass into the pump. As with process equipment for the other process steps, the Carbon Dioxide Separator is jacketed and insulated so as to maintain the process mass temperature within the range 85–98°C.

Raw Material Dosing, Caustic Soda Neutralizing

The raw materials are caustic soda in sufficient quantity to complete the neutralization of fatty acid and brine in sufficient quantity and concentration to bring the process mass into the desired Neat Soap phase configuration.

Completion of Neutralizing with Caustic Soda

At this point the process mass passes through a neutralizer identical to Direct Neutralization, and once combined with the correct raw materials in the desired flow ratios, the process mass consists of a single phase, Neat Soap.

Neutralization Soap Viscosity

The viscosity (flowability) of the soap through the neutralization process, be it direct neutralization or carbonate neutralization, is highly important as it provides insight as to composition and governs pressure drop through the process. Three factors affect viscosity—soap phase configuration, temperature, and electrolyte content. To elaborate:

- The process must always be Neat Soap as the target. None of the alternative phase configurations is satisfactory.

Slurry Soap is very low in soap content (and high in water). Middle Soap and Kettle Wax are too high in viscosity for satisfactory throughput rates.

- Process mass temperature must be suitable for required throughput rates. The lower the temperature the higher the viscosity and vice versa. The solidification temperature of the soap in process is about 70°C. However, at temperatures approaching the boiling point of water, steam forms and becomes dispersed in the soap mass, thus increasing viscosity.
- The process mass must contain a concentration of electrolyte within prescribed limits. Concentrations both above and below the limits increase viscosity.

Having kept the soap at the correct composition and temperature, a uniform quality soap product can be produced ready for drying or for use in other areas of the soap making system. The soap drying process can be accomplished by vacuum spray-dryers, as described earlier, to produce soap pellets ready for finishing into soap bars.

Soap Finishing. The starting point for soap finishing is usually the amalgamator, a horizontal mixer that coarsely blends soap pellets with minor ingredients such as color and fragrance. The soap mix from amalgamation is conveyed to subsequent processing stages like milling and refining to provide satisfactory homogenization and working of the soap before extrusion. A mixer/refiner process is also available for blending of soap pellets with minor ingredients on a more continuous basis. Sufficient working of the soap through fine-mesh screens in the refining stages provides uniformity and frictional heat necessary to produce acceptable-quality billets for pressing into bars. The finished product is extruded from the last plodder as a continuous slug. Vacuum is provided at the final extrusion chamber to evacuate entrained air. This step ensures the extrusion of a uniform slug that will not separate or exhibit surface bubbles/deformities. A multiblade chain cutter or uniblade cutter with blades separates the extruded slug into individual “billets” that are stamped with a press into soap bars of the desired size and shape.

The wrapping, cartoning, and end packaging machinery complete the bar finishing lines. The schematic diagrams of a variety of soap finishing configurations are illustrated in Fig. 36.6.

Soap Aesthetics. The soap finishing stage is also critical in defining the final appearance of the bar, an important factor in the consumer appeal of the product. Aside from the bar color, several options are available to the soap manufacturer in developing unique soap aesthetics. A widely used look, pioneered by Colgate-Palmolive and others, is bar striation. Here, a liquid mixture of contrasting color is injected into the soap pellets in the final finishing stages just prior to extrusion. By preventing the liquid mixture from fully blending into the soap, a randomized striping effect is created. A similar process with minor modifications can be used to achieve a marbleized look. Another example of multiphase look has recently been introduced by Unilever’s Lux in Latin America. Here relatively large, discrete particles of a second soap of slightly different color and composition are dispersed within the bar. Sometimes it is possible to incorporate soap performance enhancers such as skin conditioners into the aesthetic features of the bar. Other commercial examples of soap-aesthetics are exfoliating soaps where micron-sized particles of exfoliants such as polyethylene beads are added to the soap pellets in the plodder hopper. A bar with a shimmery of pearlescent effect can also be produced by dosing a low level of mica into the soap pellets generally in a mixing device such as an amalgamator.

Fatty Acids

The fatty acids are produced industrially from triglycerides. This process involves the steps shown in Fig. 36.9 [7, 8].

Refining involves the purification of triglyceride to remove impurities (phosphatides, polyethylene, chlorophyll, heavy metals, off-odors, color bodies) by a combination of acid/alkali washing, clay/activated silica bleaching, deodorization, and hydrogenation steps.

Splitting, or the hydrolysis of triglycerides, is usually performed with high-pressure steam, resulting in the formation of split crude fatty acids and glycerin. The production of fatty acids by more sophisticated “splitting” processes, such as hydrolysis of fatty methyl esters, ozonolysis of unsaturated fatty acids, and chemical oxidation is practiced in special situations.

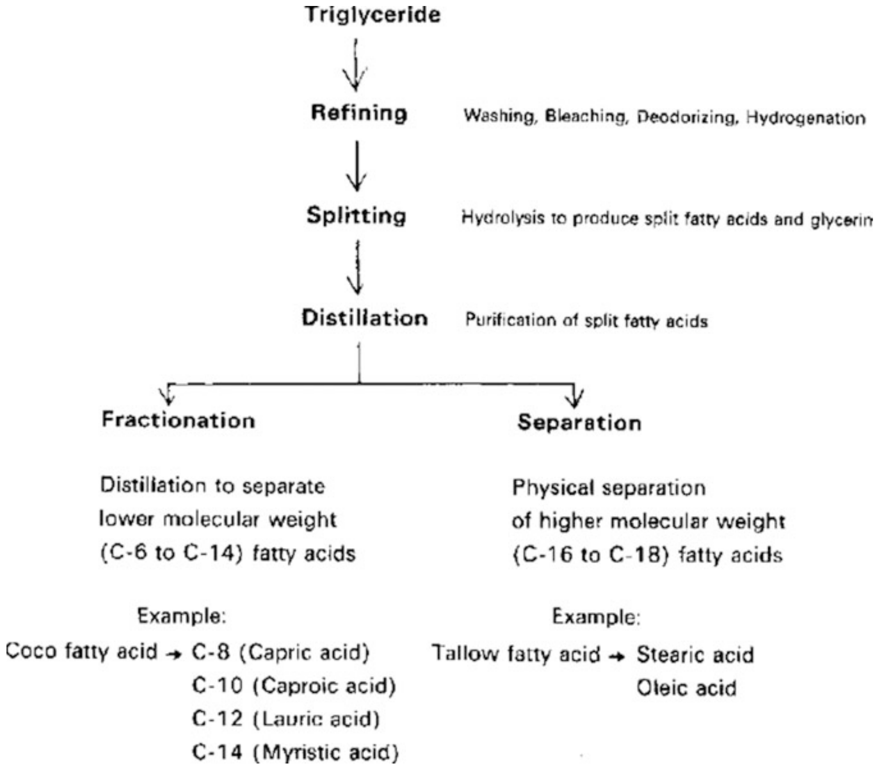
Distillation of split fatty acids results in the improvement of chemical, color, and odor quality of the finished product.

Fractionation and Physical Separation

Distilled fatty acid streams are separated to obtain individual components of those fatty acid streams in a higher state of chemical purity. Fractionation is performed to separate the more volatile lower molecular weight fatty acids (C_8-C_{14} chain length). Physical separation is practiced to obtain higher molecular weight, less volatile, or thermally unstable fatty acids (C_{16}-C_{18} chain length).

Hydrogenation is utilized to reduce the unsaturation and poly-unsaturation of fatty acids. Hydrogenated fatty acids can also be produced by the hydrogenation of triglycerides followed by splitting of the resulting fat/oil blend. The

Fig. 36.9



selective reduction of di- and poly-unsaturation is accomplished frequently by the use of nickel-based specialized catalysts in a process often called “touch” hydrogenation.

In the partial hydrogenation of the unsaturated moities the isomerization of the double bonds of mono-unsaturated components is a common occurrence; it results in the transformation of the *cis* to the corresponding *trans* stereoisomer. As the physical properties of these isomers may vary significantly, their impact on the properties of the final product should be evaluated rather carefully, for little information is available on this subject at present.

Fat Refining

A typical fat refining plant (Alfa-Laval process) for the acid washing of fats and oils is illustrated in Fig. 36.10 [9]. The triglyceride is degummed to remove phosphatides and other impurities such as mucilage, proteinaceous matter, and trace metals by acid washing with citric or phosphoric acids, and then is sent to a bleaching plant where it is dehydrated and treated with an activated clay for the removal of color bodies, heavy metals, chlorophyll, and polyethylene (Fig. 36.11) [6].

Because odor is a key quality attribute in bar soaps, a deodorization step is commonly used following earth bleaching to remove objectionable odor bodies from fats/oils. The base odor of the refined triglyceride affects the fragrance character of the finished bar and needs to be minimized for optimum fragrance delivery in product use. Deodorization is typically accomplished with 0.2–0.5% stripping steam at 1–3 mmHg and 200–250°C. The material of construction for refining processes should be stainless steel, 316L preferred, to avoid contamination with trace metals that are common to the use of mild steel.

Fat Splitting Processes

The direct reaction of triglycerides with alkali followed by acidification or the treatment of triglyceride with water (with or without an acid catalyst) results in the splitting of fats/oils into fatty acids and glycerin. The splitting, or hydrolysis of fats/oils with water, can also be performed by using various lipase enzymes as catalysts. The enzyme process is still in the developmental stages. The inter-esterification of triglycerides with methanol followed by acid or alkali cleavage of the resulting fatty methyl ester is used for the production of certain molecular weight (C_8-C_{10}) fatty acids; those methyl esters are generally the by-products of fatty alcohol production (Fig. 36.12).

The splitting of fats/oils can be done either in a batch (Twitchell process) or in a continuous (Colgate-Emery process) mode; the enzyme-based processes are still under development for commercial applications. A comparison of these processes follows [10, 11].

Process	Hydrolysis time (h)	Conversion (%)

(*continued*)

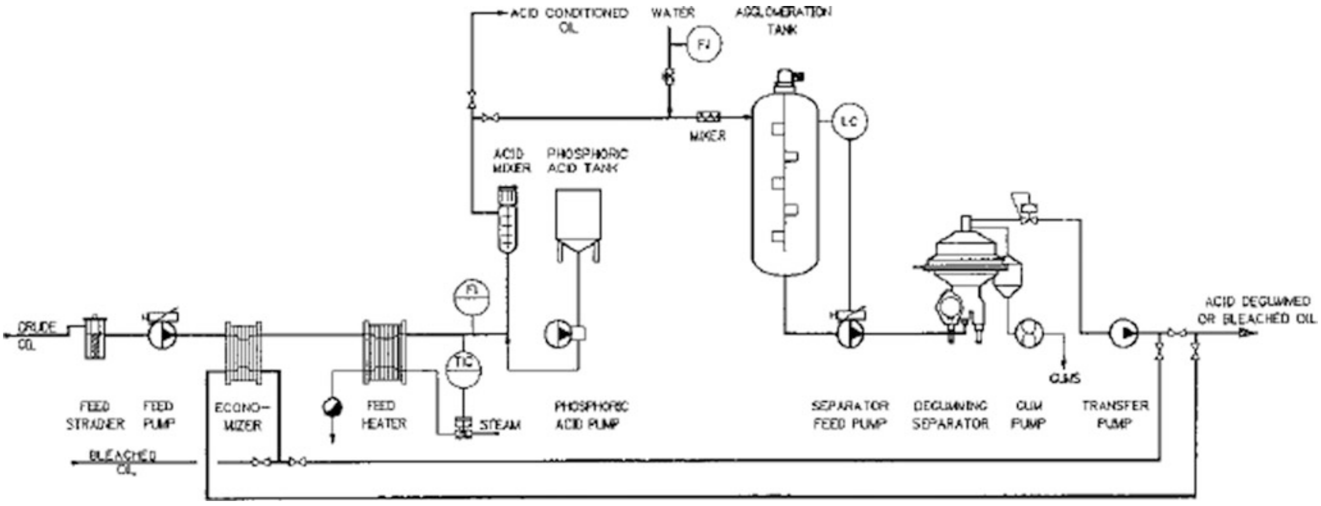


Fig. 36.10 Acid degumming plant for acid washing of fats and oils (*Courtesy Alfa-Laval, Tumba, Sweden*)

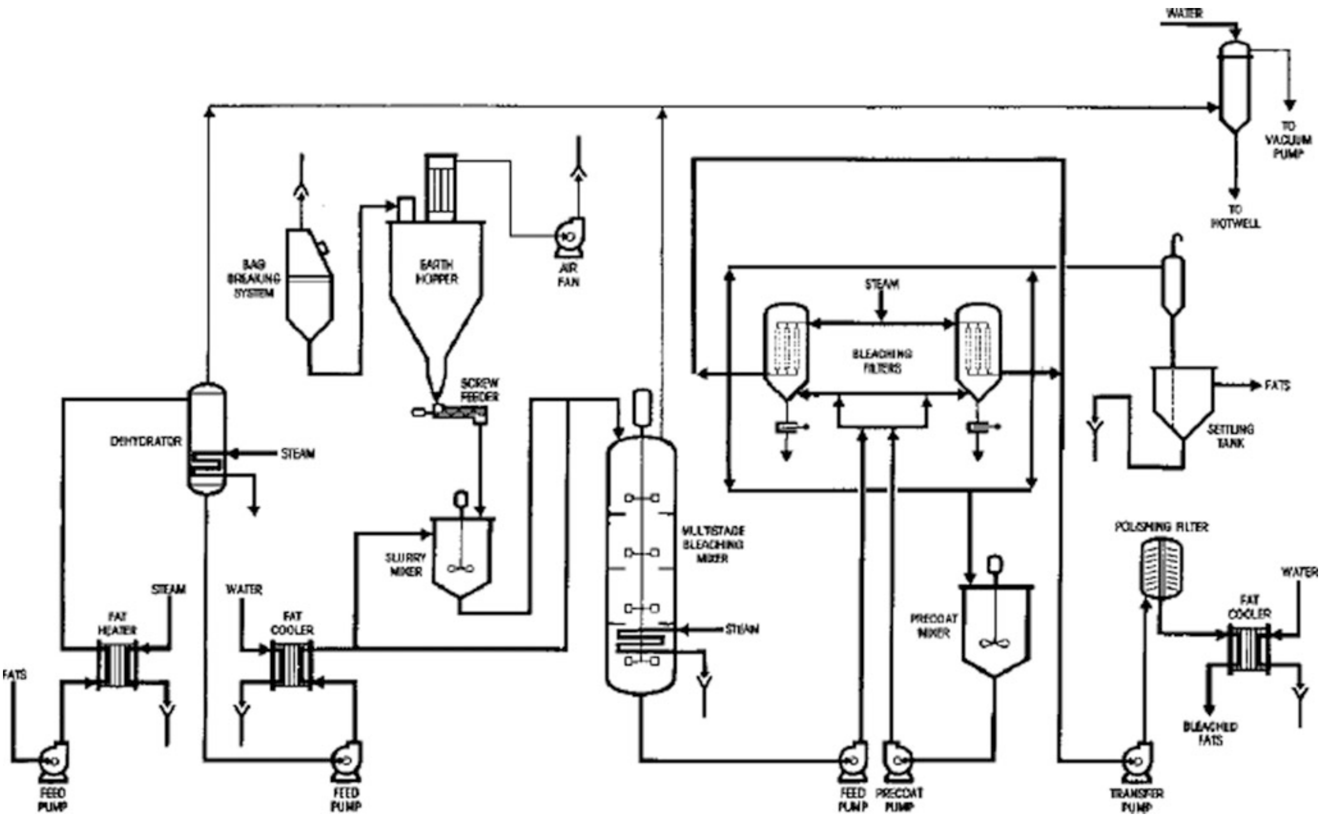


Fig. 36.11 Continuous fatty matter bleaching plant (*Courtesy G. Mazzoni S. p. A., Busto Arsizio, Italy*)

Process	Hydrolysis time (h)	Conversion (%)
Twitchell batch	8–24	95
Continuous	1–3	98–99
Enzyme	72	98

The manufacture of fatty acids from triglycerides via their reaction with alkali followed by acidification is not used commercially. However, the recovery of fatty acids from alkali refining of fats/oils (to produce acid soaps) and soap reboiling operations (the recovery of fatty acids from the acidification of soap waste streams) are practiced still in special situations, for example, in edible oils refining where soap-stock that is generated from chemical (alkali) refining

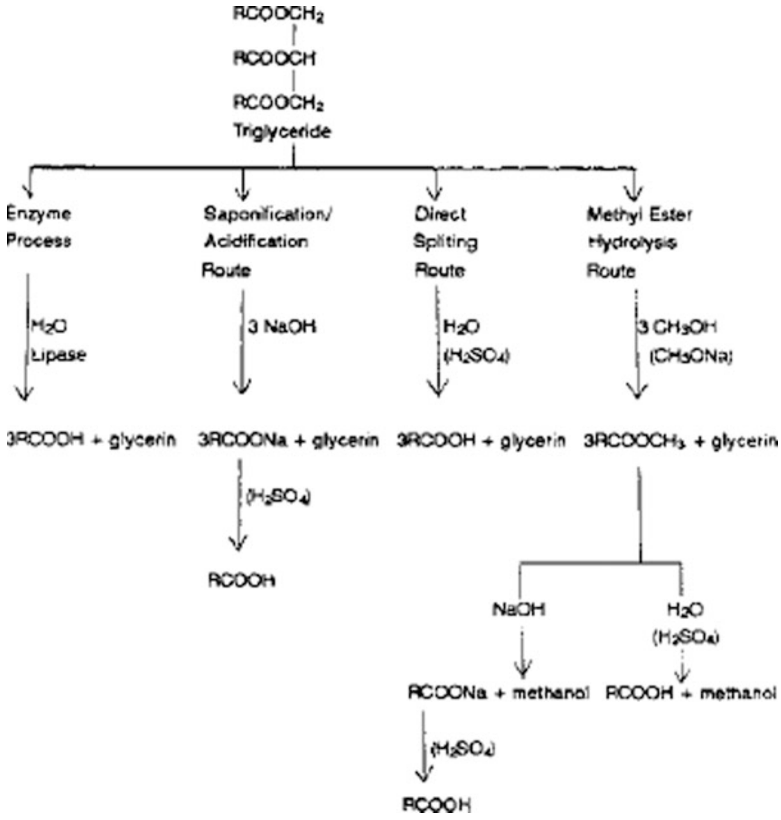


Fig. 36.12 Fat splitting processes

requires acidulation to produce acid oils with commercial value.

Industrially, the fatty acids are produced predominantly by the high-pressure, high-temperature splitting of triglycerides. Both batch and continuous processes are available, the continuous process being in greater worldwide use today.

Batch Process. The triglyceride is heated with water at 150–250°C and 10–25 bar pressure. The splitting is achieved in 6–10 h, depending on the nature of the triglyceride feed, for example, the higher the molecular weight of the triglyceride, the slower the splitting reaction. The reaction can be accelerated by increasing the temperature; under high pressure and at 225–260°C, no catalyst is required.

A typical commercial batch splitting plant (Lurgi) operates at 230°C/400 psig with over 96% split yields after two changes of water to remove the glycerin by-product. The sweetwater (glycerin/water mixture) is circulated continuously from the base to the top of the autoclave in this process [12].

A number of other batch fat splitting plants are also available.

Continuous Fat Splitting Process. Currently, continuous fat splitting is the most widely used process in the industry. In the single-stage splitting process, fat and water are fed countercurrently into a continuous flow column at up to 260°C/60 bar pressure conditions. The glycerin by-product generated in the splitting process is removed by water (sweetwater), thus accelerating the reaction equilibrium toward completion (Fig. 36.13) [12]. The high splitting temperature ensures

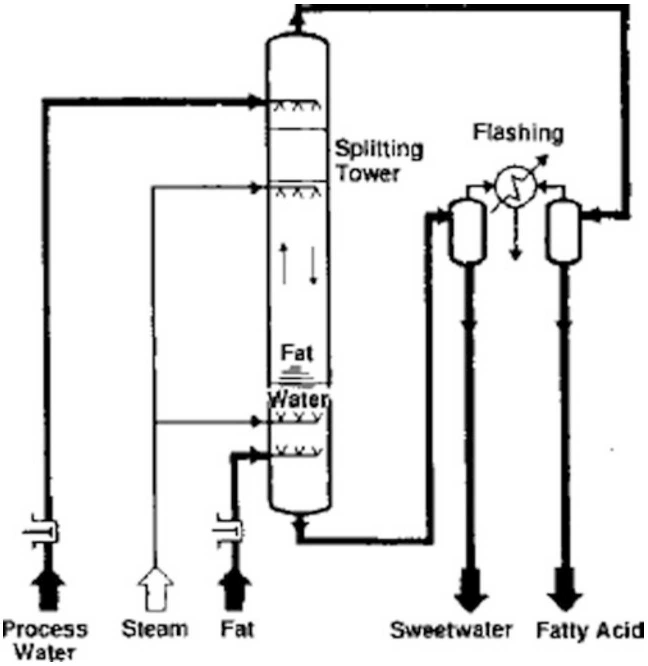


Fig. 36.13 Single-stage countercurrent fat splitting plant (*Courtesy Lurgi GmbH, Frankfurt am Main, Germany*)

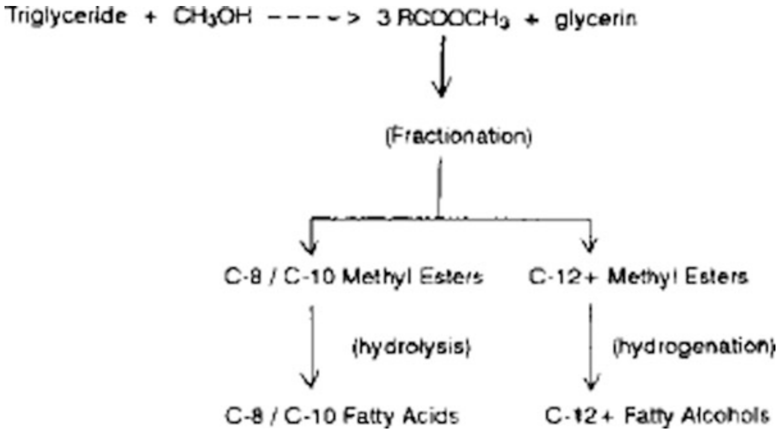


Fig. 36.14 Fatty acid production via methyl esters

adequate dissolution of the water phase in the fat to the extent that mechanical forces for contacting the two reactants are not required. The fat phase passes from bottom to top through the tower, with the void volume of the tower acting as a reaction chamber. The water phase travels downward through the mixture of fat and fatty acid, concentrating the glycerin by-product in the sweetwater phase. The top and bottom sections of the tower are also designed to effect the transfer of heat from the products to the incoming reactant streams. Split yields of up to 99% can be achieved by the use of prepurified fat/oil feedstocks.

In summary, the continuous high-pressure splitting of fats and oils involves:

1. The use of excess water to drive the reaction to completion and for the removal of glycerin from the reaction zone.
2. The use of a proper combination of temperature and pressure conditions to increase the solubility of water in the fat phase while maintaining the boundaries of the fat/water interface. These conditions also result in an increase in the rate of splitting.
3. The possible use of an acid catalyst to accelerate the reaction rate.

Saponification/Acidification Process. The recovery of fatty acids from soapstocks and soap reboil operations is of commercial value:

$$\underset{\text{(unrefined)}}{\text{Triglyceride}} \xrightarrow{\text{caustic refining}}$$

$$\text{Triglyceride} + \underset{\text{(soapstock)}}{\text{RCOONa}} \xrightarrow{(H_2SO_4)} \text{RCOOH}$$

$$\underset{\text{(impuresoap)}}{\text{RCOONa}} \xrightarrow{\text{reboiling}}$$

$$\underset{\text{(purifiedneatsoap)}}{\text{RCOONa}} \xrightarrow{(H_2SO_4)} \underset{\text{(purifiedfattyacid)}}{\text{RCOOH}}$$

The recovery of fatty acids from soapstocks by a continuous process has been described: soap-stocks obtained from degumming and alkali refining operations are subjected to a saponification step followed by controlled acidification for cost efficiency and pollution control [13, 14].

The saponification/acidification route is used for the manufacture of fatty acids that are sensitive to excessive heat. The splitting of castor oil with NaOH at 100–220°C followed by acidification has been used to produce castor fatty acids [15].

Enzymatic Processes. The enzyme hydrolysis of triglycerides is catalyzed by lipases:

$$\text{Triglyceride} + 3H_2O \xrightarrow{\text{enzyme}} 3\text{RCOOH} + \text{Glycerin}$$

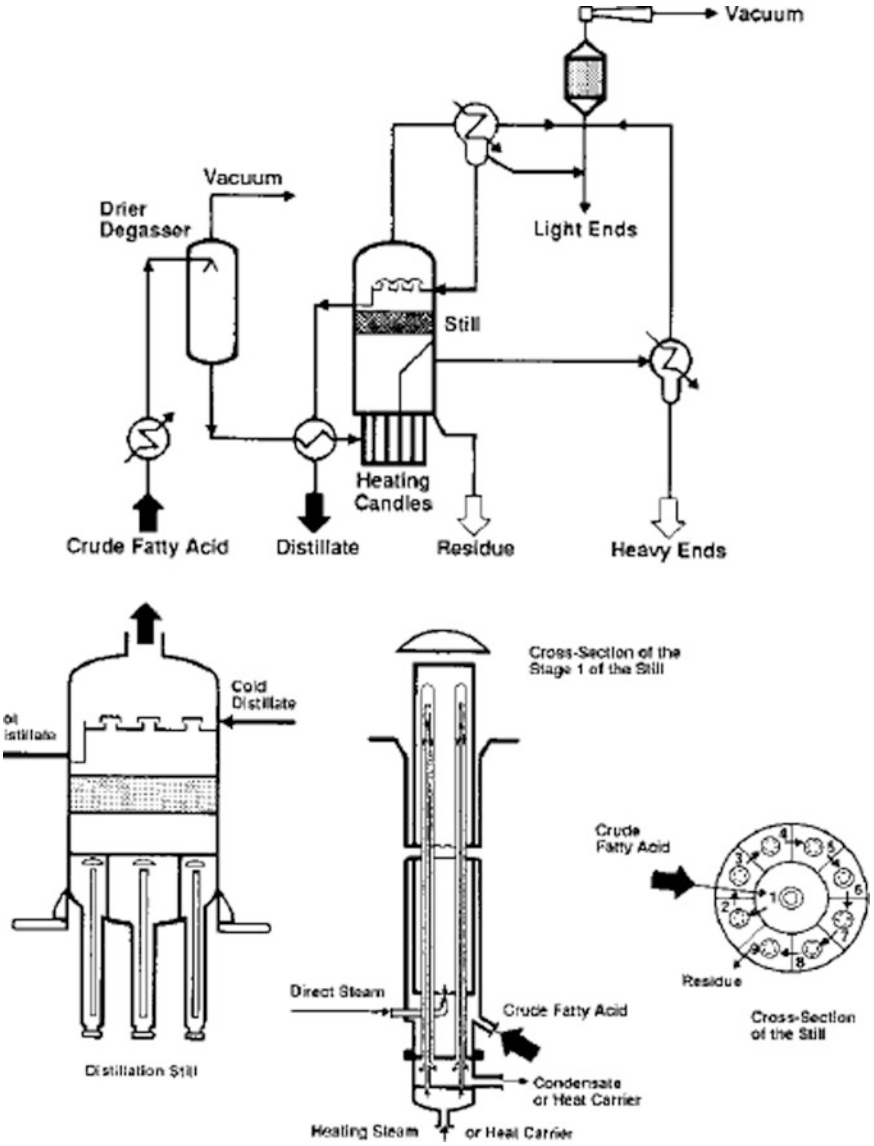
Enzyme-based splitting of triglycerides is under active development and offers the following advantages over modern steam hydrolysis plants.

- The hydrolysis occurs under mild conditions (35–37°C, ambient pressures).
- The color quality of the products is superior.
- The capital investment may be lower.
- The split yields are equivalent.

The selection of a suitable lipase for a commercially viable process is based on the following criteria.

- It should be nonspecific and nonselective as to the structure of the triglyceride.
- It should be stable at above ambient processing temperatures, typically 45–65°C.
- It should be economically viable. (This aspect has initiated a number of ongoing studies on the recycling and immobilization issues of lipase utilization for triglyceride hydrolysis.) [16, 17]

Fig. 36.15 Continuous fatty acid distillation plant (*Courtesy Lurgi GmbH, Frankfurt am Main, Germany*)



Methyl Ester-Based Processes. The fatty methyl esters are produced predominantly by the transesterification of fats and oils with methanol in the presence of an alkaline catalyst under very mild reaction conditions [18, 19]. They are used in the production of lauric-type (C_{12}) alcohols. The short-chain fatty methyl esters (C_8-C_{10}), produced as by-products via the fractional distillation of crude lauric-type (coconut, palm kernel) methyl esters, are converted to fatty acids via acidic or alkaline hydrolysis (Fig. 36.14). The hydrolysis of short-chain fatty methyl esters by stream splitting or Twitchell-type processes is not very efficient because of unfavorable equilibrium constants [20, 21].

Fatty Acid Distillation

The removal of contaminants (unreacted triglycerides, odor and color bodies, polymerized matter, and decomposition products) from split fatty acids is achieved by simple distillation [22, 23]. The products obtained from such straight distillations are called "whole cut" fatty acids (e.g., whole coconut fatty acid).

The separation of individual fatty acid components from whole fatty acid distillates is achieved by fractional distillation [24, 25]. However, fractional distillation cannot efficiently separate C_{16}-C_{18} saturated from the corresponding unsaturated

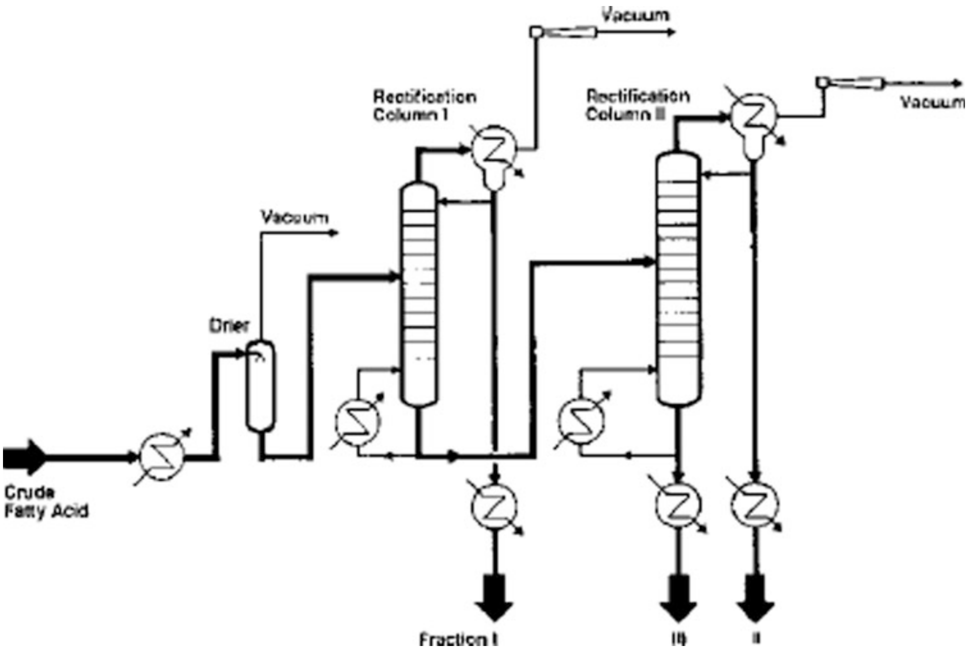


Fig. 36.16 Fatty acid fractionation (*Courtesy Lurgi GmbH, Frankfurt am Main, Germany*)

fatty acids. In such cases, the melting points vary considerably with the degree of unsaturation and this property is used for their separation via a number of industrial processes [26, 27].

Distillation. Straight distillation is used for the removal of low-boiling (top cuts) and high-boiling (pitch or bottoms) impurities from the crude split fatty acid stock. The distillation can be done either in a batch or a continuous manner. Generally batch distillation is employed for small-capacity manufacturing; continuous distillation is utilized for large-scale manufacturing and is highly automated (Fig. 36.15) [12].

The split fatty acid (deaerated and preheated) is sent to a distillation column (Fig. 36.15) where it is vaporized rapidly by flowing over heated trays under high vacuum (10 mbar) with the addition of live steam (for fatty acid circulation and reduction of the partial pressure). The high boilers are removed by two stages of entrapment: the high boilers and bottoms from the distillation column pass through additional heated trays and a reboiler section to vaporize residual fatty acids from the bottoms; the light end condenser traps the low boilers and the distilled product. The heavy "bottoms" distillate is collected at the bottom section of the still.

The distillation process is not used for the narrow-cut fractionation. However, by passing a second time through the distillation tower the main-cut distillates may be separated into a light and heavy fraction; the light end is collected at the top products section and the heavy end is collected at the bottom residue section of the still. The separation of light and heavy fractions in the first pass is possible for high lauric-type split fatty acids, for example, coconut and palm kernel fatty acids where the lower molecular weight C_6–C_{10} fraction is "topped" from the heavier C_{12}–C_{18} fraction.

The operation of a Lurgi fatty acid distillation system [12] proceeds as follows. The crude fatty acid is dried and degassed under vacuum. This dried and deaerated crude fatty acid is pumped into the distillation chamber (vacuum of 10–12 mbar) that is divided into 5–10 chambers by vertical plates. Each chamber is equipped with a heating element and a baffle. Live steam is passed into the system to effect recycling of the liquid between the heating elements and the baffles by the principle of an air-lift pump. The fatty acid is sprayed on the baffles as a fine mist to allow it to evaporate easily; non-vaporized liquid is recycled from the bottom. The fatty acid passes through each chamber, from which it is discharged continuously. The fatty acid vapor is condensed in the surface condensers. Typical yields from the fatty acid distillation process are:

Tops or head cut	1–5%
Refined product	85–91%
Residue	7–11%

Fractionation. The separation of individual components of a fatty acid mixture into high purity fractions (98% +) is achieved by fractional distillation. In a continuous fractionation still the fractionation column can be equipped with trays, in the "Thormann" system, or with the traditional packing to effect contact between the upward-moving vapors and the downward-flowing condensing liquid states.

A typical high-efficiency fractionation system (Fig. 36.16) consists of one continuous straight distillation and one, or preferably two fractionation columns [12, 22, 23]. This system is very efficient for the separation of low-boiling fatty acid components [6, 12], a C_{12} fraction of 99.5% purity from coconut fatty acid or a C_{22} fraction of 95% purity from rapeseed fatty acid is readily obtained.

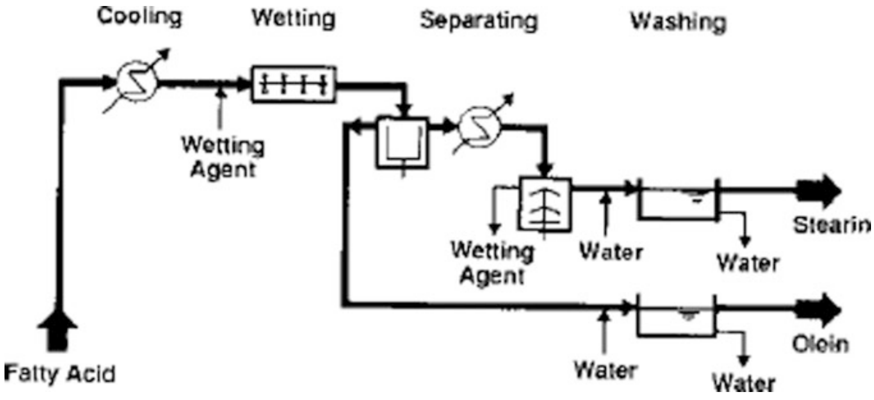


Fig. 36.17 Hydrophilization process for separation of olein and stearin (*Courtesy Lurgi GmbH, Frankfurt am Main, Germany*)

From an operational perspective, in rectification column 1 of Fig. 36.16, crude fatty acid is dried and vaporized with the addition of live steam under a vacuum of 8–40 mbar. A mixed fraction is withdrawn from the column top. The product from the bottom of column 1 is vaporized in rectification column 2 with live steam under a pressure of 6–8 mbar. The low-boiling fraction is collected in downstream scrubbers after the vapor has been condensed in surface condensers.

The separation scheme for coconut fatty acid is shown in Fig. 36.9. A variety of alternate fractionation schemes can readily be devised.

Fatty acid distillation has been done for over 100 years [22, 23]. However, most of the currently practiced art has originated in the past several years [24, 25]. The majority of the recent work falls in the following areas.

- Designs with gentle heating
- Improvements in the column design to reduce residence time
- Enhancements of system design to minimize pressure drops
- Demister designs for better air/effluent water pollution control

Special Separation Methods

The separation of long-chain saturated fatty acids (C_{16}–C_{18}) from the corresponding unsaturated fatty acids by distillation is not practical because of the proximity of their boiling points. However, both the melting points and the solubility (in organic solvents) of these two types of fatty acids are vastly different [28]. These properties are used to advantage for separating mixtures of saturated and unsaturated fatty acids into fractions that are rich in either saturated (stearin) or unsaturated (olein) components [26, 27].

$$\text{Tallow fatty acid} \xrightarrow{\text{separation}} \underset{(\text{Stearin})}{\text{Stearic acid}} + \underset{(\text{Olein})}{\text{Oleic acid}}$$

The following processes are used commercially for the separation of such fatty acids: mechanical pressing, solvent crystallization, and hydrophilization.

Mechanical Pressing. This method, also called panning or pressing, is the simplest and oldest method, and is still practiced in many developing countries. In this method, fatty acids are melted in shallow pans and then allowed to cool. The material transforms into a crystalline solid with a liquid phase trapped between the crystals of the solid phase. The liquid portion (olein) is pressed, or squeezed out by the application of mechanical force, separating it from the solid (stearin) component.

The pressing can be done by hydraulic presses in either a batch or a continuous mode. A multiple number of melting/crystallization/pressing cycles are utilized for obtaining fractions of increased purity:

$$\text{Tallow fatty acid} \xrightarrow{\text{first pass}}$$

$$\underset{(\text{stearicacid})}{\text{Single press}} \xrightarrow{\text{second pass}}$$

$$\underset{(\text{Stearicacid})}{\text{Double press}} \xrightarrow{\text{third pass}} \underset{(\text{Stearicacid})}{\text{Triple press}}$$

The unsaturation, as determined by the iodine value (IV), decreases from an IV of 10.0 for a single press to an IV of 1.0 for a triple press stearic acid grade.

Solvent Crystallization. Two processes, one utilizing acetone (Armor) and the other employing methanol (Emery), are well known. Using the latter, fatty acid is dissolved in 90% aqueous methanol in a 1:2 acid/solvent ratio by the application of heat. The resulting solution is then cooled to −15°C in a multi-tubular crystallization chamber equipped with scrapers for efficient heat transfer. The crystallized fatty acids are removed by filtration. The filter cake is melted and stripped

Table 36.2 US consumption of soap and synthetic detergent[a]

Year	Soap (MM lbs)	Synthetic detergent (MM lbs)	Total
1940	2,306	30	2,336
1950	2,882	1,443	4,325
1960	1,230	3,940	5,170
1970	1,050	5,650	6,700
1980	1,300	6,400	7,700
1990	1,625	9,000	10,625
1996	–	18,214[b]	0

[a] CEH Marketing Report, "Specialty Chemicals," Chemical Economics Handbook, SRI International, Aug. 1988. A personal Communication from SRI (08/14/02) states that because of the variety of forms and concentrations in which detergent products now appear, such data are no longer available

[b] Colin A. Houston & Associates, "Surfactants for Consumer Products, North America Forecast for 2008"

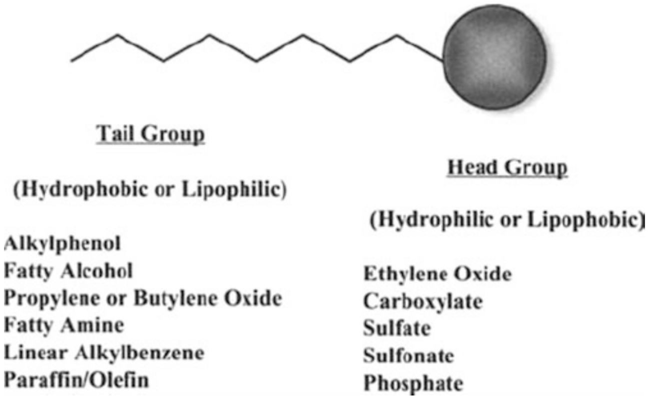


Fig. 36.18 Representation of surfactant amphipatic structure

of any residual solvent to yield the refined stearin fraction, and then the liquid stearin is converted to flakes or powder by a variety of processes, for example, chill roller, and the like. The mother liquor from the filtration is stripped to obtain the olein fraction. The separated stearin and olein fractions have a variety of commercial applications in both the chemical and food processing industries.

Hydrophilization Process. In this process, olein and stearin are separated by forming a suspension of the stearin moiety in water with the aid of a surfactant followed by filtration to remove the stearin (Fig. 36.17). The mixture of fatty acids is melted and then partially crystallized by cooling. The resulting slurry, which is a mixture of liquid olein and crystalline solid stearin, is mixed with water containing a wetting agent (sodium dodecylbenzene sulfonate, sodium lauryl sulfate). After mixing for a specified period, the slurry is centrifuged; the light fraction provides the liquid olein moiety, the middle fraction is a suspension of solid stearin in water and the lower fraction is a water/surfactant solution [12].

Synthetic Detergents

In the beginning there was soap, and soap was used for thousands of years to wash virtually everything. After the First World War, much of the European economy was destroyed. Edible fats and oils, the basic raw material for soap production, were in very short supply. Synthetic detergents were developed to replace soap but it was soon discovered that the new synthetic compounds were more surface active and provided better hard water tolerance and improved cleaning performance. The term "synthetic detergent" has been shortened to "syndets" to describe detergent compositions comprising surfactants, hard water control agents (builders), and other detergent additives. Since their introduction in the 1940s, synthetic detergents have continued to grow in both total volume and the percentage of detergent use as shown in Table 36.2.

Modern detergent products are complex mixtures of many different ingredients. Typical formulations consist of surfactants, builders, and other additives designed to maximize performance for the consumer while maintaining reasonable raw material and manufacturing costs. Typical detergent formulations contain multiple surfactant types to optimize performance and product stability [29, 30]. Performance additives such as bleaches, bleach activators, enzymes, enzyme stabilizers, antiredeposition polymers, fluorescent whitening agents, dye transfer inhibitors, and fabric softeners are commonly used to increase overall performance. Regardless of the product format or formulation type, surfactants play a critical role in the cleaning process.

Characteristic Features of Surfactants

Surfactants have a characteristic molecular structure consisting of one structural unit that has very little attraction for the solvent (lyophobic group) combined with a group that has a strong attraction for the solvent (lyophilic group). This duality is known as an amphipathic structure. When a surfactant is dissolved in a solvent, the presence of the lyophobic group distorts the solvent liquid structure, resulting in an unfavorable increase in the Gibbs free energy of the system. This acts to force the molecules out of solution in an attempt to decrease the free energy of the system. However, the presence of the lyophilic group prevents the surfactant from being completely expelled from the solvent. The opposing forces give the surfactant molecule their surface active properties. In aqueous solution, the lyophilic group is known as the hydrophile and the lyophobic group is called the hydrophobe.

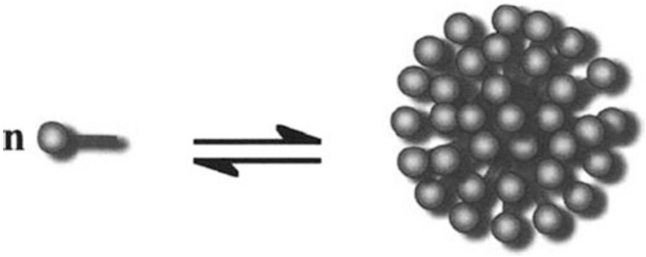


Fig. 36.19 Micelle formation in aqueous solution

Table 36.3 Typical HLB values for different applications

HLB range	Typical application
1–3	Reverse micelles
3–6	W/O emulsifier
7–10	Wetting agent
8–18	O/W emulsifier
11–15	Detergent
>15	Solubilizing agent

Surfactant molecules are typically depicted using a stick and ball representation as shown in Fig. 36.18. The stick portion represents the hydrophobic group or "tail" whereas the ball represents the hydrophilic "head group." Examples of common hydrophobes include alkylphenol, fatty alcohol, paraffin, olefins, and alkylbenzene. Examples of common hydrophiles include ethylene oxide (EO), sulfates, sulfonates, phosphates, and carboxylic acid groups. Depending on the charged nature of the head group, surfactants are classified as:

1. Anionic. The head group bears a negative charge. Examples: linear alkylbenzene sulfonate (LAS), AS, alkyl ether sulfonate (AES), PE, alpha olefin sulfonate (AOS), methyl ester sulfonate (MES)
2. Cationic. The head group bears a positive charge. Examples: cetyl trimethyl ammonium chloride (CTAB), cetyl pyridium chloride (CPC), hydroxyethyl quat (HEQ), benzalkonium chloride (BAC)
3. Zwitterionic. The head group has either a positive or negative charge depending on solution pH. Examples: betaines, amine oxides, amine propionates, amphoacetates
4. Nonionic. The head group bears no formal ionic charge. Examples: AE, alkyl phenol ethoxylate (APE), alkyl polyglycoside (APG), EO/PO Block Copolymers, ethoxylated methyl esters (MEEs)

Aside from their ability to adsorb at interfaces, the most important aspect of surfactants is their ability to form colloidal-sized aggregates in solution. In dilute solution, the surfactant is present as individual molecules. Increasing the concentration promotes the formation of surfactant aggregates or "micelles" as shown in Fig. 36.19. The concentration at which micelles start to form is referred to as the critical micelle concentration (CMC). Micelle formation is an important phenomenon because interfacial effects such as detergency and solubilization depend on the existence of micelles in solution. Over the years, the term "surface active agent" has been shortened to "surfactant."

Surfactant molecules can be tailored for use in either aqueous or nonaqueous systems, depending on their solubility characteristics. The most common way to describe surfactant solubility is the hydrophile–lipophile balance (HLB). In the HLB method, a number is assigned (0–40) related to the balance between the water-loving (hydrophile) and lipid-loving (lipophile) portions of the molecule. In some cases the HLB number is calculated from the structure of the molecule, in others it is based on experimental emulsification data. In general, there exists an optimum HLB for different applications as shown in Table 36.3. Matching the HLB of the surfactant to that of the application generally gives the best performance.

For example, alkylbenzene is slightly surface active in nonaqueous media but is insoluble in water. Addition of a single SO_3H group to the molecule imparts solubility and excellent surfactant characteristics in water but insolubility in petroleum solvents. If alkylbenzene is di- or trisulfonated, the compound becomes more and more water-soluble and loses much of its surface activity. In both cases, the hydrophile balance is exceeded with insufficient lipophilic influence to cause micelle formation. However, neutralization of the SO_3H group of a highly water-soluble alkylbenzene sulfonic acid with a long chain amine produces an effective nonaqueous surfactant. A similar effect of decreasing water solubility and increasing nonaqueous detergency can be achieved by increasing the alkylbenzene carbon chain length from C_{12} to C_{18} or higher.

Commercially available surfactants typically consist of a complex mixture of homologues. For instance, "lauryl sulfate" is a generic designation for a mixture of sulfates whose largest fraction is derived from the C_{12} alcohol but also contains higher and lower alkyl chain lengths. The homologue distribution depends upon the alcohol source and degree of fractionation. Alcohol ethoxylates are also complex mixtures of homologues. Not only is there a distribution of different alkyl chain lengths, but EO adds to the alcohol according to a normal distribution of adducts. Although it might seem that highly purified products would show a competitive advantage, mixtures are frequently more effective than pure compounds.

Raw Materials for Surfactant Production

Surfactants are produced by various chemical routes, using either natural or synthetic raw materials as shown in Fig. 36.20. Natural raw materials include fats and oils

Fig. 36.20 Raw materials for surfactant production

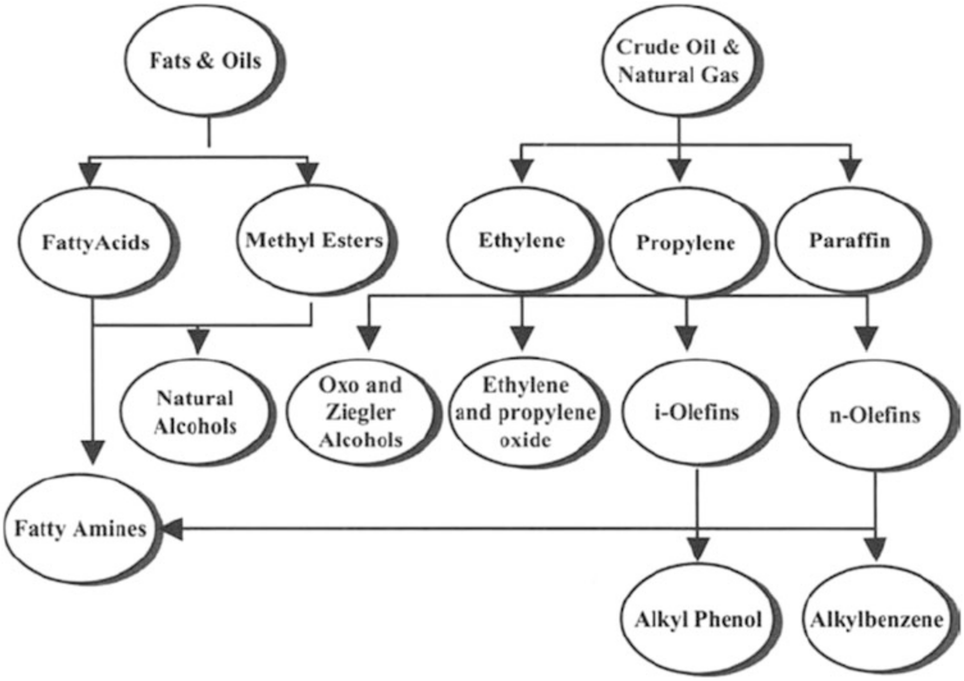


Table 36.4 Carbon chain distributions

	Coconut	Palm kernel	Soybean	Canola/ rapeseed	Palm stearin	Inedible tallow
C6	1	1				
C8	8	4				
C10	7	4				
C12	48	48				
C14	18	16	1		1	3
C16	8	8	12	5	57	26
C16 = 1						3
C18	2	2	4	2	5	21
C18 = 1	6	15	25	60	30	41
C18 = 2	2	3	52	22	7	3
C18 = 3			7	10		1

derived from vegetable and animal sources. Synthetic raw materials include coal, crude oil, and natural gas. Regardless of their origin, raw materials are chemically modified to produce a wide variety of surfactant intermediates. Because chemical processing is required to produce the surfactant intermediate, the distinction between natural and synthetic is a bit fuzzy. In general, raw material prices increase with the number of processing steps.

The current world production of fats and oils is around 65 MM metric tons of which 30% is from animal sources and 70% is of vegetable origin. Fats and oils are composed of long chain fatty acid esters of glycerin called triglycerides. Plant-derived oils are composed of fatty acids with an even number of carbon atoms from 8 to 22 and can be either saturated or unsaturated. Oils with a high content of C_{12} and C_{14} fatty acids are most commonly used by the detergent industry and are readily obtained from coconut and palm kernel. Other important oils include soya, rapeseed, and palm stearin, The most important animal fats are lard and tallow which are rich in C_{16} and C_{18} fatty acids. The carbon chain distributions of common fats and oils are shown in Table 36.4.

Common natural-based raw materials for surfactant production include fatty acids and methyl esters. After purification to remove phospholipids and sterols, fats and oils are treated with alkali to produce soap which is acidified with inorganic acids to give the corresponding fatty acid. To produce methyl esters, fatty acid can be reacted with methanol. Alternatively, the oil is treated with methanol using a base catalyst to give the corresponding methyl ester. Both processes generate one mole of glycerin for every three moles of fatty acid or methyl ester.

Common synthetic-based raw materials for surfactant production include ethylene, and propylene. Crude oil consists of a complex mixture of long chain hydrocarbons and aromatic molecules. Natural gas is a mixture of short chain hydrocarbons rich in methane, ethane, propane, and butane. The exact composition of both depends on its source and how it has been processed. Ethylene and propylene are produced by thermal or catalytic cracking of natural gas or aromatic rich petroleum streams.

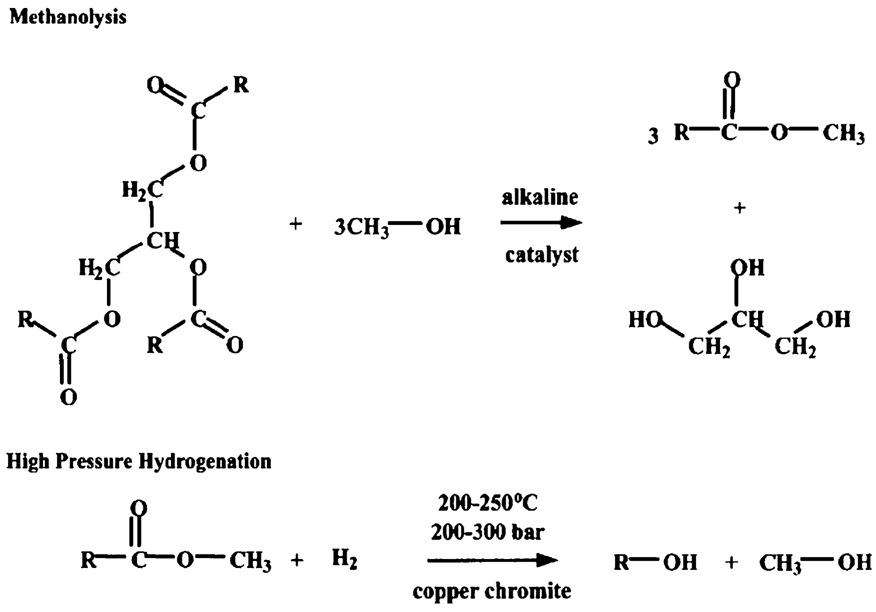


Fig. 36.21 Methyl ester process for production of natural fatty alcohols

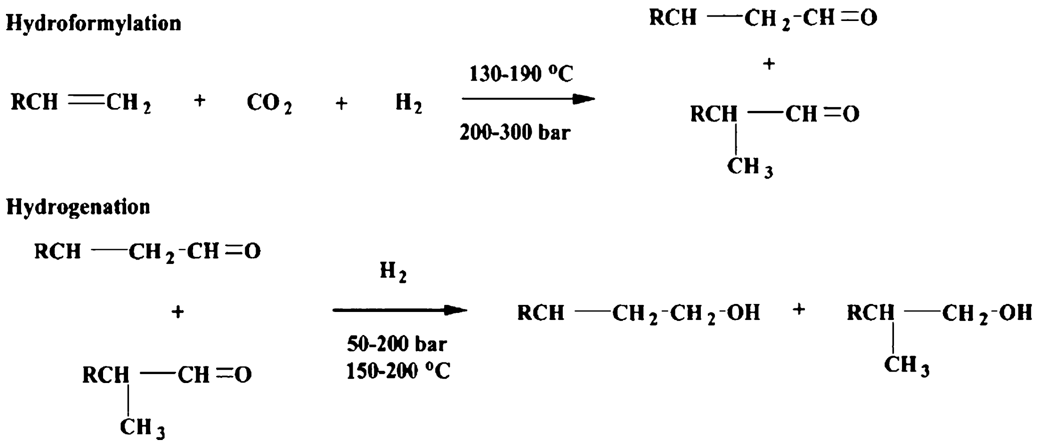


Fig. 36.22 Oxo-process for production of synthetic fatty alcohols

Intermediates for Surfactant Production

Fatty Alcohols. Surfactant raw materials and intermediates are used to produce a wide variety of different surfactant types. One of the major factors leading to the development of modern synthetic detergents was the industrial manufacture of fatty alcohols. This was first achieved in Germany in the 1930s using the sodium reduction process. Although this process was commonly used until the 1950s, it has been largely replaced by less costly processes. Today, there are three important technical processes for the manufacture of detergent grade fatty alcohols: high-pressure hydrogenation, the Ziegler process, and the oxo-synthesis [31].

The most important process for the manufacture of natural fatty alcohols is the high-pressure hydrogenation of fatty acids or their corresponding methylesters. Common raw materials for the production of detergent-range alcohols include coconut, palm kernel, palm stearin, soya, tallow, and lard. In the methyl ester process, the crude oil is reacted with an excess of methanol at 50–70°C to give three moles of methyl ester and one mole of glycerin as shown in Fig. 36.21. The reaction can be accelerated using sodium methoxide as a catalyst. The glycerol separates at the bottom of the reactor after an appropriate reaction time. The transesterification is complete when the methylester no longer contains glycerol.

The hydrogenation reaction is run at 200–350°C and hydrogen pressures of 250–300 bar using an appropriate catalyst. A variety of different metal oxide catalysts can be used; however, copper chromite is the most common. Under these relatively extreme reaction conditions, the ester group is reduced to the primary alcohol. Any unsaturation in the

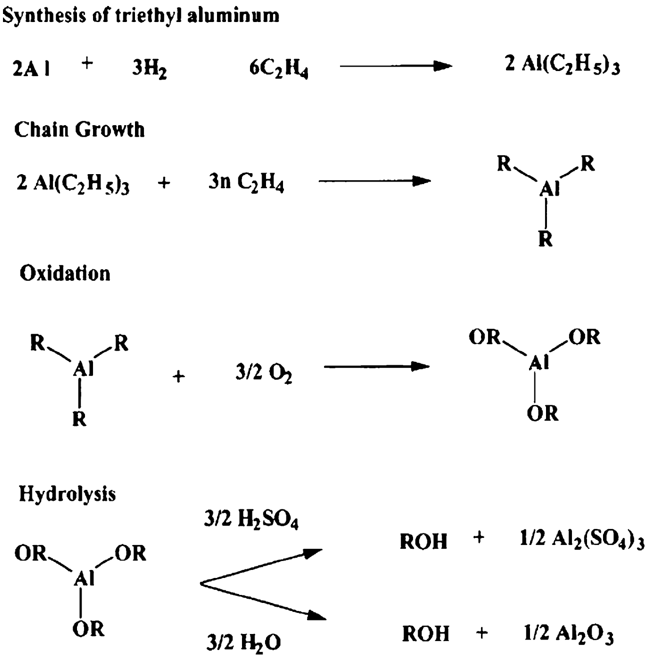


Fig. 36.23 Ziegler process for production of synthetic fatty alcohols

starting raw material is also destroyed under these conditions although it is possible to produce unsaturated fatty alcohols using various zinc catalysts. A number of different process designs have been developed including the fixed bed, the trickle bed, and the suspension process which differ primarily in the way in which the catalyst contacts the feedstock.

The oxo-process is based on petrochemical feedstocks. The oxo- or hydroformylation reaction was first discovered in Germany in the late 1930s, but was not used commercially until 1948, by what is now Exxon Chemical. The oxo-process uses α-olefins obtained from oligomerization of ethylene, which are reacted with carbon dioxide and hydrogen at 200–300 bar and 130–190°C, to give a mixture of linear and branched aldehydes as shown in Fig. 36.22. The resulting aldehydes are hydrogenated at 50–200 bar pressure and 150–200°C using copper, zinc, or nickel catalyst to give the primary alcohol. Oxo-alcohols contain both even and odd carbon numbers and are more highly branched than natural or Ziegler alcohols. Using a conventional cobalt catalyst, the amount of branching exceeds 50%.

The Ziegler process based on chemistry discovered in the 1950s produces linear primary alcohols with an even number of carbon atoms. The main reaction steps are shown in Fig. 36.23. Aluminum metal is reacted with hydrogen gas and ethylene to form triethyl aluminum which is then reacted with additional ethylene to form trialkyl aluminum using a one-step or two-step process. In practice, the one-step process yields more by-products and is seldom used. The distribution of chain lengths follows a Poisson distribution with only slight deviations due to competing side reactions. The aluminum alkyls are oxidized to the corresponding alkoxides using dry air and hydrolyzed using water or sulfuric acid to yield fatty alcohol. The Ziegler process produces alcohol similar to that obtained from natural sources.

Olefins. Olefin is a key surfactant intermediate produced by thermal or catalytic cracking of paraffin or alternatively from oligomerization of ethylene. Olefins suitable for surfactant production are primarily linear with either terminal (α-olefins) or internal (i-olefins) double bonds.

Thermal cracking is not very selective and produces a mixture of products including *n*-olefin, i-olefin, di-olefin, aromatics, and paraffin wax. The major challenge associated with the cracking process is separation of the desired olefins from the other by-products.

Since the mid-1960s, *n*-paraffins have been used as starting materials for the manufacture of *n*-olefins. Today, *n*-paraffins are obtained from the distillation of paraffin rich oils with the use of molecular sieves.

The PACOL® process (paraffin conversion to olefin) produces *n*-olefins by dehydrogenation of paraffin over a heterogeneous platinum catalyst. The Pacol process is more selective than thermal cracking and produces smaller amounts of by-products.

Oligomerization of ethylene can be performed in a number of different ways. Reaction of ethylene with triethyl aluminum using the Ziegler process produces normal olefins by a three-step process. The first step consists of reacting ethylene with triethyl aluminum to give long chain trialkyl aluminum. Transalkylation of the trialkyl aluminum with ethylene (Alfol process) or a long chain olefin (Ethyl process) produces a mixture of olefins and regenerates the triethyl aluminum. The third step consists of separating the olefin from the undesired by-products.

Another commercial process to produce *n*-olefins is based on the Shell Higher Olefins Process (SHOP). The oligomerization reaction is carried out on a nickel phosphine complex which is separated and recycled by means of a solvent. The reaction product consists of α-olefins and i-olefins of different carbon chain lengths.

Recently there has been renewed interest in the use of the Fisher–Tropsch process to produce olefins. The process is named after F. Fischer and H. Tropsch, the German coal researchers who discovered it in 1923. Detergent-range olefins are produced by treating white-hot coal with a blast of steam to produce syngas, a mixture of carbon monoxide and hydrogen. Syngas is reacted in the presence of an iron or cobalt catalyst to produce a mixture of products including methane, synthetic gasoline, olefins, and waxes.

Ethylene and Propylene Oxide. Ethylene oxide (EO) and propylene oxide (PO) are commonly used raw materials for surfactant production. Both are highly reactive compounds

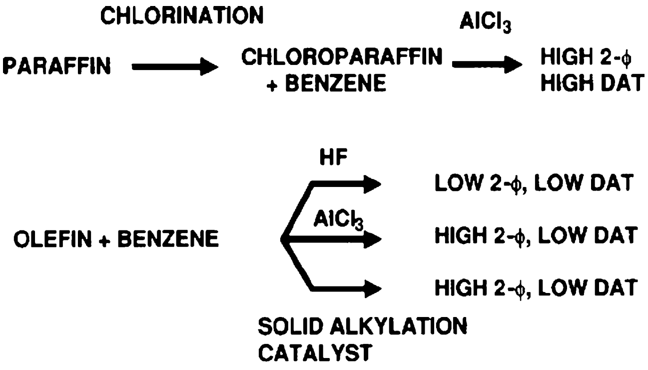


Fig. 36.24 Commercial alkylbenzene production

Table 36.5 2003 world surfactant consumption[a]

Surfactant	Billion lbs
Soap	19.8
LAS	6.4
BAB	0.4
AES	1.8
AS	1.3
AES	2.4
APE	1.3
Quats	1.1
Amphoterics	0.2
Others[b]	5.3
Total	40.1

[a]H.G. Hauthal, *SÖFW-Journal*, 10, 130, 2004
[b]Including lignin, petroleum sulfonates, and tall oil derivatives

that will react with alcohols, amines, carboxylic, and water. EO and PO will also self-polymerize to form high-molecular-weight polymers.

Ethylene oxide has been produced commercially by two basic routes: the ethylene chlorohydrin and direct oxidation processes. The chlorohydrin process was first introduced during World War I in Germany by BASF. The process involves the reaction of ethylene with hypochlorous acid followed by dehydrochlorination of the resulting chlorohydrin with lime to produce EO and calcium chloride. This process has now been replaced by direct oxidation which utilizes the catalytic oxidation of ethylene with oxygen over a silver-based catalyst.

PO is produced by one of the two commercial processes: the chlorohydrin process or the hydroperoxide process. The chlorohydrin process involves reaction of propylene and chlorine in the presence of water to produce the two isomers of propylene chlorohydrin. This is followed by dehydrochlorination using caustic or lime to produce PO and salt. The hydroperoxide process involves oxidation of propylene to PO by an organic hydroperoxide.

Alkylbenzene. Alkylbenzene is an intermediate for the production of alkylbenzene sulfonate (ABS). Alkylbenzene consists of a mixture of phenyl substituted *n*-alkanes of 9–14 carbon atoms. Prior to 1965, alkylbenzene was synthesized from propylene tetramer, obtained by oligomerization of propylene. The resulting “hard” alkylate was a highly branched chain compound. However, the slow biodegradability of propylene tetramer-based materials soon became apparent and by 1965, most of the detergent industry had switched over to linear alkylbenzene. Extensive research has demonstrated biodegradation effectiveness in sewage treatment plants in excess of 95% [32–35].

Today, there are two main routes for producing linear alkylbenzene as shown in Fig. 36.24. The most prevalent route to alkylbenzene uses partial dehydrogenation of paraffins, followed by alkylation of benzene with a mixed olefin/paraffin feedstock using liquid hydrogen fluoride (HF), aluminum chloride ($AlCl_3$), or solid alkylation (DETAL) catalyst. A second older process, which has largely been discontinued, uses partial chlorination of paraffins, followed by alkylation of the chloroparaffin feedstock in the presence of an aluminum chloride catalyst. Distillation is used in all the processes to separate alkylation reaction products into the targeted detergent alkylate homologue distribution. A coproduct stream of high molecular weight linear alkylbenzene and dialkylbenzene is also produced for use in lubricant detergents and industrial applications.

Products with low 2-phenyl isomer content (13–22%) are produced by HF alkylation of internal olefins obtained from paraffin dehydrogenation. High 2-phenyl products (25–35%) are produced using $AlCl_3$ or solid alkylation catalyst (DETAL). Super high 2-phenyl (>80%) products have been reported in the literature using shape-selective catalysts, but are not commercially produced at this time. Generally, high 2-phenyl products are preferred for use in liquid detergents due to increased solubility. Low 2-phenyl products are typically used in powder detergents but can be used in liquid formulations with the appropriate choice of hydrotrope.

Alkylphenol. Alkylphenol is a common surfactant intermediate used to produce alkylphenol ethoxylates. Phenol reacts with an olefin thermally without a catalyst but with relatively poor yields. Catalysts for the reaction include sulfuric acid *p*-toluene sulfonic acid (PTSA), strong acid resins, and boron trifluoride (BF_3). Of these, strong acid resins and BF_3 are mostly widely used for the production of surfactant-grade alkylphenols. The most common alkylphenols are octylphenol, nonylphenol, and dodecylphenol. Mono nonylphenol (MNP) is by far the most common hydrophobe. It is produced by the alkylation of phenol with nonene under acid conditions. All commercially produced MNP is made with nonene based on propylene trimer. Because of the skeletal rearrangements that occur during propylene oligomerization, MNP is a complex mixture of branched isomers.

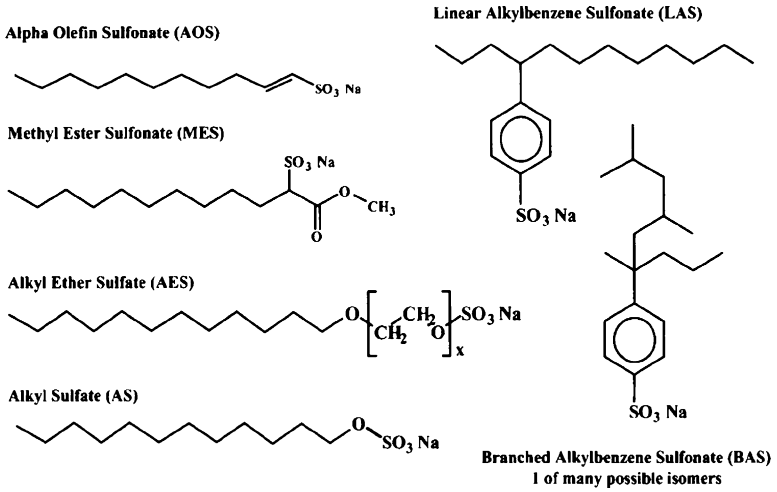


Fig. 36.25 Common anionic surfactants

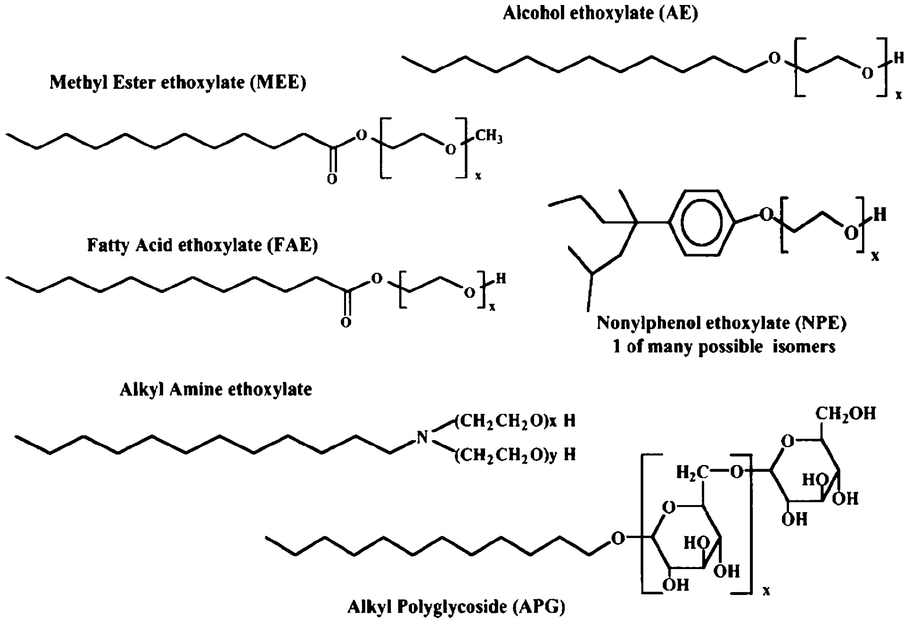


Fig. 36.26 Common nonionic surfactants

Fatty Amines. A variety of different fatty amines is used as intermediates for surfactant production. Primary amines are used to prepare alkoxylated amines. The primary amine is produced by reacting fatty acid with ammonia to form the nitrile followed by hydrogenation. Tertiary amines are used as an intermediate for producing betaine, amine oxide, and HEQ. The tertiary amine is produced by reacting olefin with HBr followed by substitution with dimethylamine.

World Surfactant Consumption

The worldwide consumption of surfactants in 2003 is shown in Table 36.5. In many parts of the world, soap is still the primary surfactant for textile washing and personal care. Synthetic detergents are widely used in the devolved regions of the world including North America, Western Europe, and Japan. Anionic surfactants dominate the world surfactant market. Common anionic surfactants include LAS, AS, and AES. The second largest class of surfactants is the nonionics of which AE and APE are most common. The use of APE is in decline because of concerns with biodegradability and endocrine disruption. Quats are commonly used in antibacterial formulations because of their ability to lyse the cell membrane of harmful bacteria and pathogens. Amphoterics are typically used in personal care products because they are relatively mild to the skin compared to typical anionics. The molecular structure of

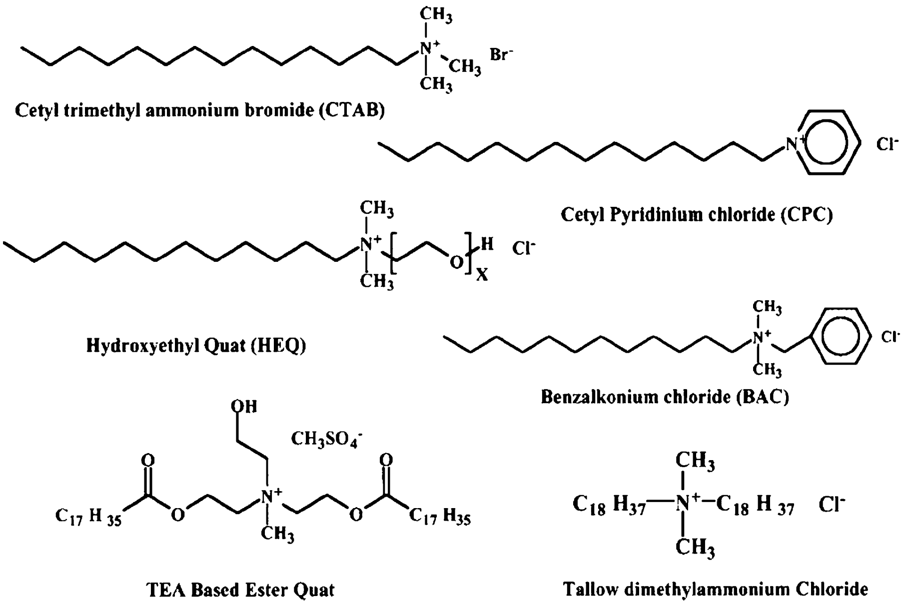


Fig. 36.27 Common cationic surfactants

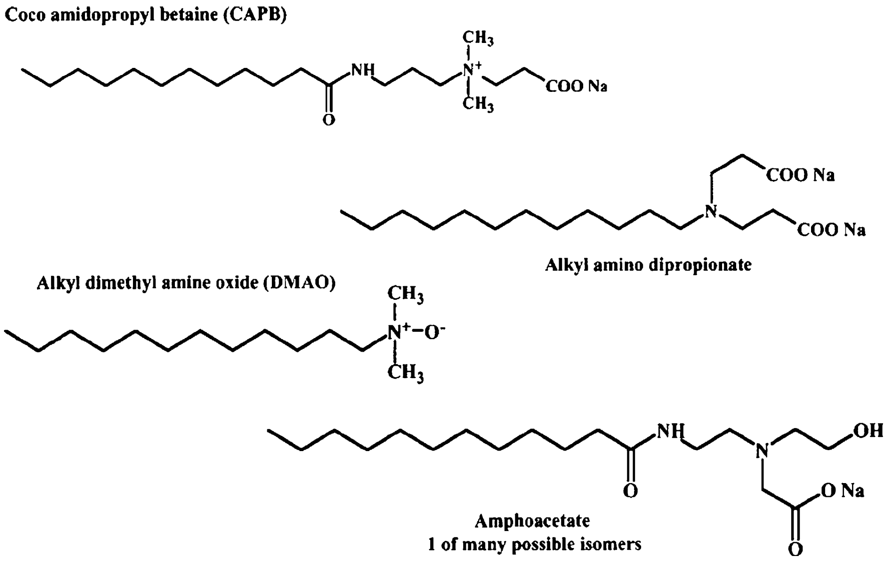


Fig. 36.28 Common amphoteric/zwitterionic surfactants

common surfactant types is shown in Figs. 36.25, 36.26, 36.27, and 36.28.

Anionic Surfactants

Alkylaryl Sulfonates. ABS has been the workhorse of the detergent industry for more than 40 years and accounts for approximately 50% of the total synthetic anionic surfactant volume. LAS has largely replaced branched alkylbenzene sulfonate (BAB) in most parts of the world due to its superior biodegradability. Products are typically sold as the free acid or neutralized with a strong base such as sodium hydroxide to give concentrated slurry, typically referred to as "paste." The major use of paste is in the production of spray-dried powder detergents. Paste can also be drum-dried to a powder or flake, or spray dried to light-density granules. Dried forms of LAS are used primarily for industrial and institutional cleaning applications.

To be useful as a surfactant, alkylbenzene must first be sulfonated. Oleum and sulfur trioxide (SO_3) are commonly used for sulfonation, with the latter growing in popularity in recent years. Oleum sufonation requires relatively inexpensive equipment and can be run as batch or continuous processes. However, the oleum process has major disadvantages in terms of cost compared to SO_3; the need for spent acid stream disposal and the potential corrosion problems due to sulfuric acid generation. The chemical reactions involved in oleum sulfonation are shown in Fig. 36.29. The oleum

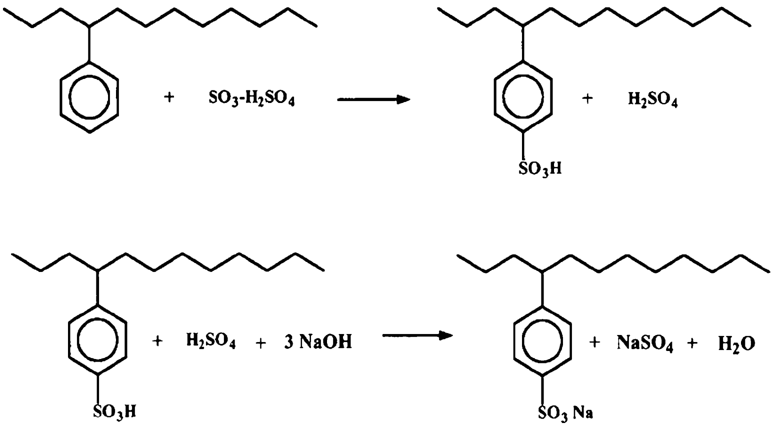


Fig. 36.29 Oleum sulfonation process

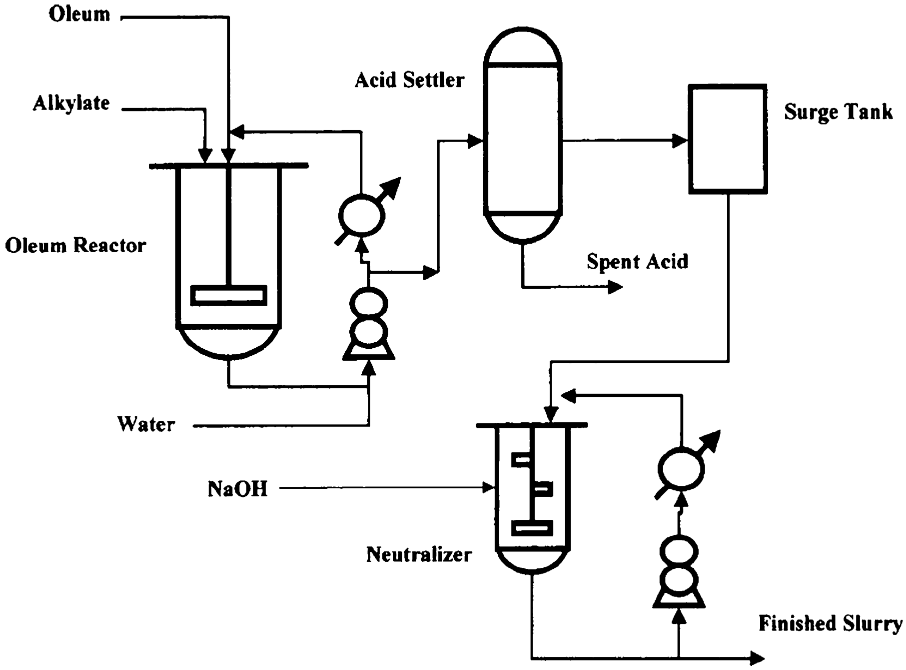


Fig. 36.30 Batch oleum sulfonation process

process typically yields around 90% ABS, 6–10% sulfuric acid, and 0.5–1% unsulfonated oils.

A typical batch oleum sufonation process is shown in Fig. 36.30. There are four basic unit operations prior to neutralization: sulfonation, digestion, dilution, and phase separation. In the sulfonation stage, alkylbenzene and oleum are mixed under an inert atmosphere. The sulfonation reaction is highly exothermic and heat removal is achieved by reactor jacketing and/or recirculating heat exchanger loops. Key variables in controlling the extent of reaction and product color are the temperature, acid strength, reaction time, and the ratio of oleum to alkylate. The product leaving the sulfonation zone is digested 15–30 min to allow for reaction completion. Following digestion, the mixture is diluted with water to quench the reaction. The reaction product is then fed to a separator to allow for gravity settling of the spent sulfuric acid layer from the lighter sulfonic acid. The separation time is dependent upon the separator vessel configuration, sulfonic acid viscosity, temperature, and the degree of aeration in the feed streams.

The gaseous air/SO_3 sulfonation process normally generates a product composed of 95–98% ABS, 1–2% sulfuric acid, and 1–2.5% unsulfonated oils. Although the initial capital costs are much higher than those for an oleum process, relative sulfonic acid yields and spent acid disposal costs are substantially lower. The chemical reactions involved in air/SO_3 sulfonation are shown in Fig. 36.31. The SO_3 sulfonation reaction differs from oleum in that the crude reaction product consists of a mixture of ABS and anhydride. At the end of the reaction, the anhydride is reacted with water to produce two moles of sulfonic acid.

A typical continuous air/SO_3 sulfonation process is shown schematically in Fig. 36.30. Sulfur trioxide is produced from vaporization of liquid SO_3 or generated on-site

Fig. 36.31 Sulfur trioxide sulfonation process

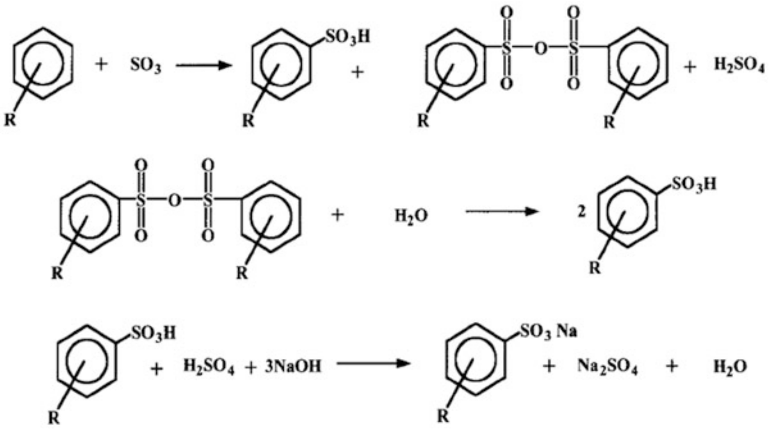
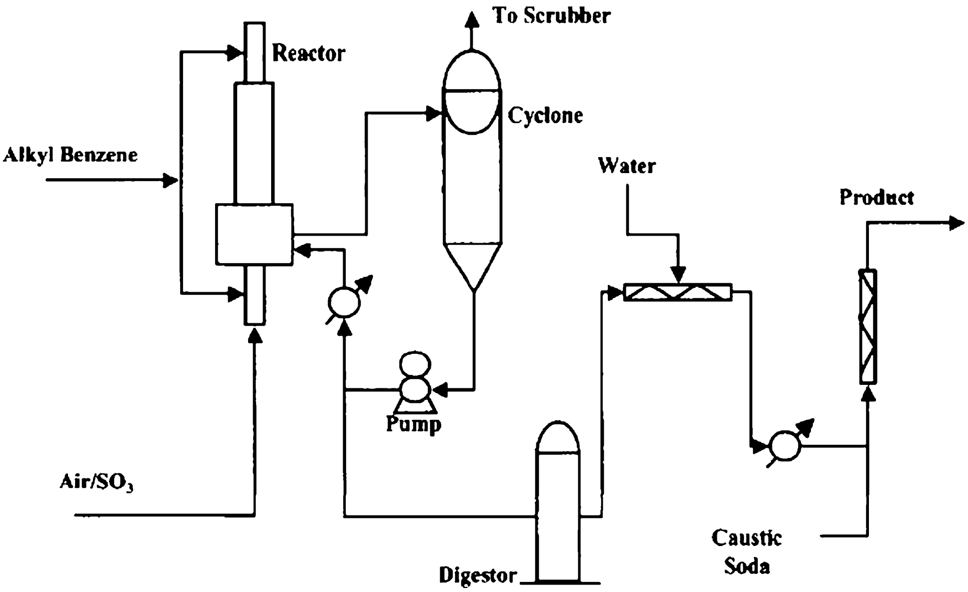


Fig. 36.32 Alkylbenzene sulfonation process



by burning sulfur and then catalytically converting SO_2 to SO_3 over a vanadium catalyst. Increased shipping regulations and sensitivity to storage of the highly hazardous liquid SO_3 has increased the trend toward on-site sulfur trioxide generation. The vaporized SO_3 is diluted to 4–8 weight percent in dry air prior to entering the sulfonation reactor. Air serves as both a diluent and heat sink for the highly exothermic sulfonation process. In the sulfonation reactor, the molar ratio of SO_3/alkylate is controlled between 1.03 and 1.0 to achieve optimum conversion to sulfonic acid without increased side reaction or color degradation.

The Chemithon [12, 13], Ballestra [15], and Mazzoni [36, 37] reactor designs each have unique falling film designs to provide intimate mixing of the alkylate and the air/SO_3 streams. The reactor temperature is controlled between 110 and 150 °F, depending on the desired production rates and product quality requirements. The digestion and hydrolysis temperatures are controlled between 110 and 125 °F. During hydrolysis, approximately 2 parts of water per 100 parts of reaction product are added to convert acid anhydrides to sulfonic acid. Both the oleum and SO_3 sulfonation processes are quite exothermic and almost instantaneous. In order to prevent decomposition and maintain optimum product color, an efficient reactor heat removal system is necessary (Fig. 36.32).

The Chemithon Corporation has patented a novel venture reactor [38, 39]. The organic to be sulfonated is injected into a stream of gas containing SO_3 at a venture. The reaction mixture is quenched and recycled downstream until sulfonation is complete. The reactor design can be used for highly viscous, high molecular weight specialty surfactants for industrial applications such as lubrication oil additives.

Cold sulfonation is also practiced using SO_2 as a solvent and refrigerant [40, 41]. The product color is extremely good due to the low reaction temperature and rapid removal of the reaction exotherm via SO_2 vaporization. Equipment consists of fairly simple glass-lined reactors equipped with gaseous SO_2 recovery. Purging of SO_2 from the reaction product and control of SO_2 recovery are key to achieving a high-quality

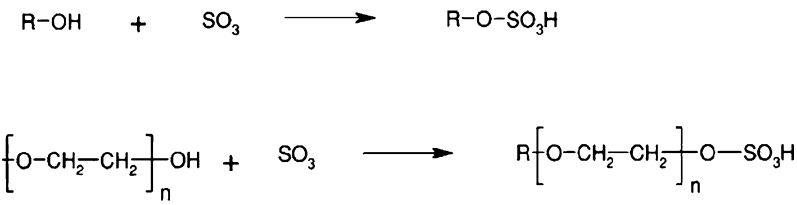
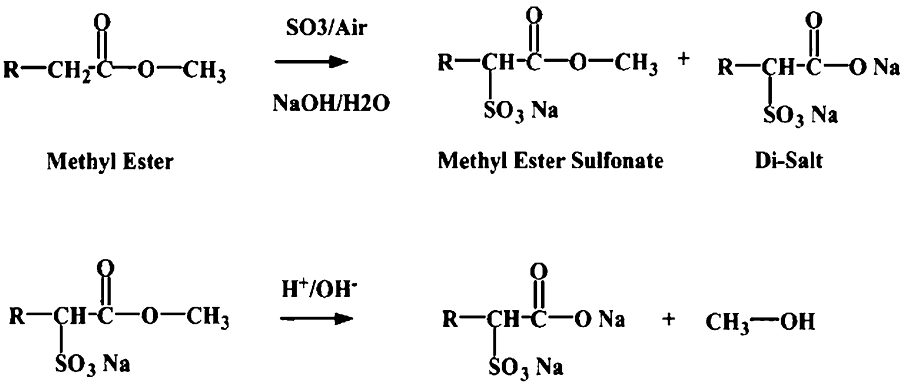


Fig. 36.33 Sulfation of alcohols and alcohol ethoxylates



Under acid or alkaline conditions, MES will hydrolyze to di-salt

Fig. 36.34 Synthesis of methyl ester sulfonates

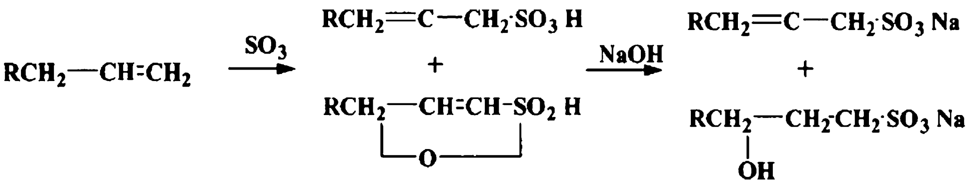


Fig. 36.35 Synthesis of α-olefin sulfonates

product. As in the case of air/SO_3 sulfonation, this process can be used to produce all detergent-grade surfactants including alkylbenzene, alcohol, and alcohol ethoxylates.

Neutralization of sulfonic acid is similar regardless of sulfonation process. ABS acid can be neutralized with an aqueous solution of NaOH, KOH, NH_4OH, or amines. Sodium salts are commonly used in spray-dried powders for household laundry detergents. Ammonium and amine salts are commonly used in light-duty liquid (LDL) detergents and in high foaming personal care products such as hair shampoos and body wash. ABS acid can also be processed to a dry powder form by in situ neutralization with a base salt, such as sodium carbonate. The neutralized sodium salt can be dry blended with other detergent additives to produce a free flowing powder detergent (Fig. 36.33).

Alkyl and Ether Sulfates. The second largest class of anionic surfactants is sulfated alcohols and ethoxylates. Falling film reactors are also used to sulfate these materials, however, cooling water and sulfation temperature are adjusted to lower values for these materials. The digestion and hydration steps are not required in sulfation and these systems are bypassed. The acids are neutralized immediately to minimize degradation and side reactions in storage. Typical SO_3 sulfation reactions for alcohols and alcohol ethoxylates are shown in Fig. 36.31.

Methyl Ester Sulfonates. MESs are an old product whose time may be coming. MES is produced by reacting detergent-range methyl esters with SO_3 to produce a mixture of α-sulfonated methyl ester and di-salt as shown in Fig. 36.34. The crude product is highly colored and must be bleached to produce acceptable products.

Alpha Olefin Sulfonates. The direct reaction of alpha-olefins with a strong sulfonating agent yields a mixture of alkane sulfonates, hydroxyllalkane sulfonates, and disulfonated products. These are known commercially as AOS. AOS manufacture is generally more complex than ABS because

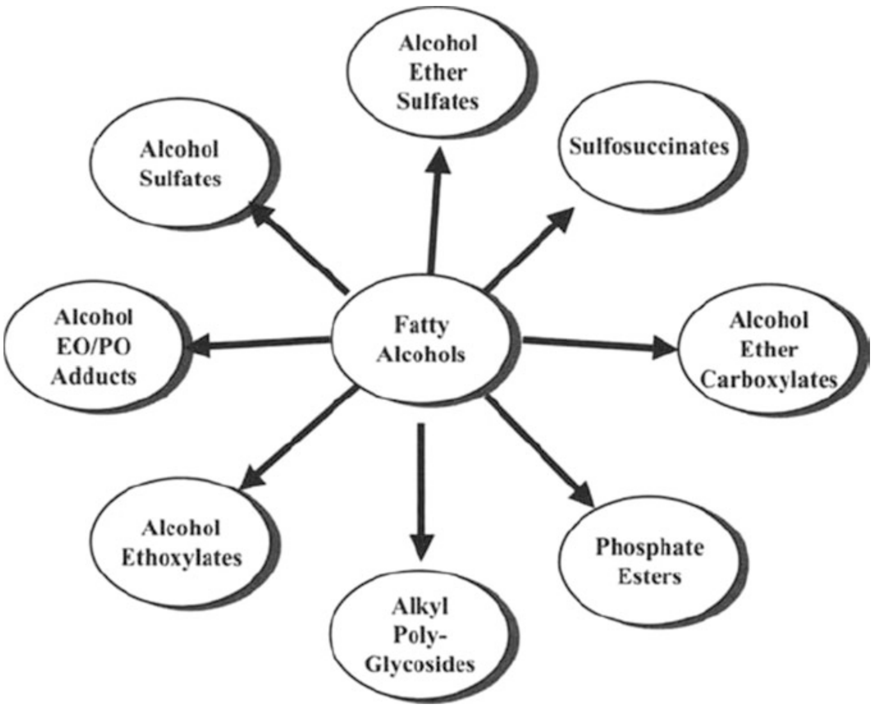


Fig. 36.36 Surfactants obtained from fatty alcohols

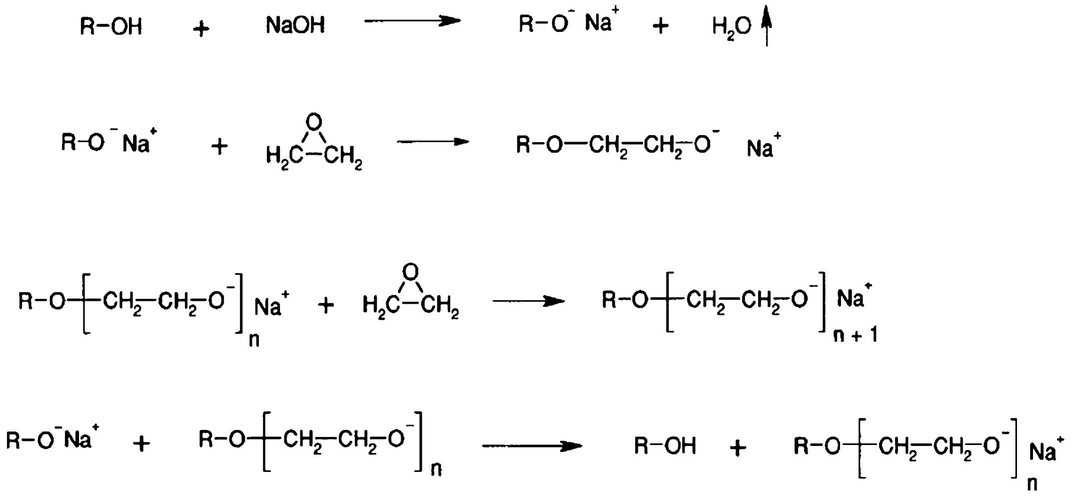


Fig. 36.37 Fatty alcohol ethoxylation chemistry

the initial reaction product is a mixture of sultones as shown in Fig. 36.35.

Nonionic Surfactants

Nonionic surfactants are increasingly popular active ingredients. The majority of nonionic surfactants are ethylene or propylene oxide derivatives of alkylphenols or fatty alcohols, although fatty acids, fatty amines, and alkanolamides are also used. In general, one ethylene oxide unit is required to solubilize each methylene unit in the hydrocarbon tails. By changing the ratio of ethylene oxide to fatty groups, the nature of the surfactant can be modified from oil-soluble to water-soluble products. This versatility makes applications possible in numerous areas, such as detergents, agricultural products, corrosion resistance, emulsion polymerization, textile applications, and de-inking to name a few.

Fatty alcohols are one of the most useful intermediates for the production of nonionic surfactants, some of which are listed in Fig. 36.36. A detailed discussion on all the

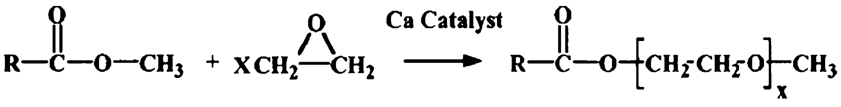


Fig. 36.38 Synthesis of methyl ester ethoxylates

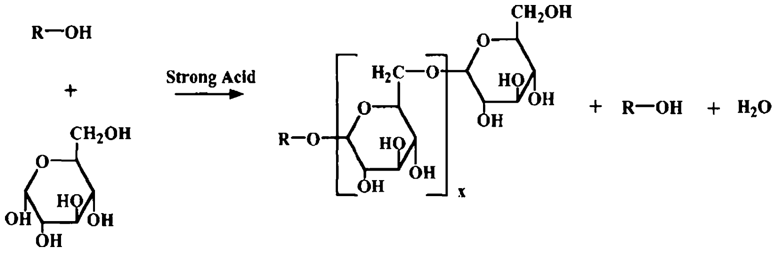


Fig. 36.39 Synthesis of alkyl polyglycosides

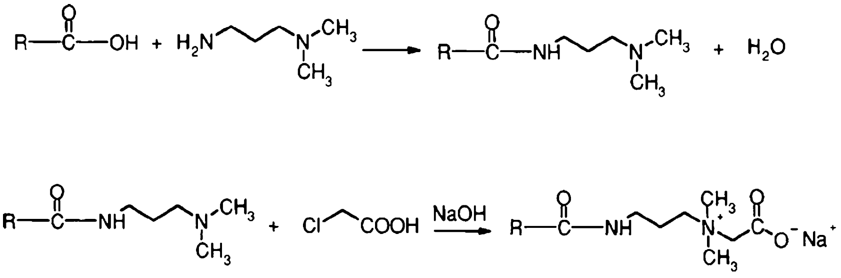


Fig. 36.40 Synthesis of coco amido propyl betaine

commercially available nonionic surfactants is beyond the scope of this work and only the major surfactant types are covered. A more complete discussion of different surfactant classes and their properties is available elsewhere [42].

Ethylene and propylene oxide are highly reactive species that will react with any molecule containing an active hydrogen. The alkoxylation reaction is best described as a two-step process. In the first step, the catalyst forms a charged complex with the initiator. In practice, the reaction is usually run using a base catalyst such as NaOH or KOH to avoid creating a large amount of polyethylene glycol (PEG) and other by-products. Alcohol and nonylphenol ethoxylates (NPE) are two of the most common types of nonionic surfactant. The reaction chemistry for the alcohol ethoxylate is shown in Fig. 36.37. The degree of ethoxylation follows a normal Poisson distribution.

To produce ethoxylated nonionics, the initiator and base catalyst are charged to a pressure reactor and purged with an inert gas to remove moisture. The reaction mixture is heated to 115–200°C and the oxide is added until the desired pressure is reached. The reaction is highly exothermic and heat removal rates control the rate of reaction. The reaction temperature and pressure are controlled until the desired quantity of oxide has been added. The batch is then cooled and the catalyst neutralized with a Lewis acid or removed by filtration.

Ethoxylated Methyl Esters. Even though methyl esters do not have active hydrogen, it is possible to insert EO into the ester group using calcium- or magnesium-based catalysts. The oligomer distribution is typically narrower than that obtained using a conventional base catalyst. The reaction is thought to take place by a complex series of transesterification reactions as shown in Fig. 36.38.

Alkyl Polyglycosides. A new type of nonionic surfactant has recently been developed that is not based on ethylene oxide. APG are prepared by reacting an excess of fatty alcohol with dextrose using a strong acid catalyst as shown in Fig. 36.39. The unreacted fatty alcohol is removed using a wiped film evaporator and the crude reaction product bleached to give light colored products. APG is readily biodegradable and nonirritating to human skin. APG is widely used in personal care and hard surface cleaning products but has not found wide acceptance in laundry detergents due to its poor cost/performance vs. other nonionics.

Amphoteric Surfactants

These surfactants contain both cationic and/or anionic groups on the same molecule. Amphoteric surfactants are relatively expensive due to the raw material and processing costs.

Typical applications include liquid dishwashing detergents, shampoos, body wash, and other personal care applications. In general, amphoteric surfactants detoxify and mitigate the irritation of anionic surfactants in the formulation. A typical reaction scheme for coco amido propyl betaine (CAPB) is shown in Fig. 36.40. Other examples of amphoteric surfactants include amine oxides, amphoacetates, and amino propionates. A detailed discussion of amphoteric surfactants and their properties is available elsewhere [43, 44].

Cationic Surfactants

Cationic surfactants tend to be specialty items that are tailored for diverse uses such as germicides, textile applications, corrosion inhibitors, and fabric softeners. Their primary growth has been in fabric softeners for home laundry detergents. Because cationics are antistatic agents, they perform well in removing static charges associated with synthetic fabrics.

Cationic surfactants have at least one hydrophobic group attached directly to a positively charged nitrogen atom. Cationic surfactants may be formed from nitriles, amines, amidoamines, or quaternary nitrogen bases. Significant literature reviews cover this subject in great detail [29, 30].

Detergent Additives

In addition to surfactants, modern detergent formulations contain many different additives to increase performance or assist in the detergent manufacture.

Detergent Builders. Builders are typically added to a detergent formulation to extend or improve the cleaning performance of a formulation across a wide range of use conditions. The combination of builders and surfactants exhibits a synergistic effect to boost total detergency and cleaning efficacy, as compared with an equal amount of either compound alone. The major properties and characteristics that are desirable for a compound to be considered a detergent builder include:

1. The ability to control water hardness and other metal ions
2. A contribution to final product alkalinity
3. Buffer capacity in the desired pH range
4. Deflocculation capability
5. Compatibility with other formulation ingredients and detergent additives
6. Consumer safety
7. Environmental acceptability
8. Processibility
9. Adequate cost/performance

In today's detergent products, a variety of builder materials is used singly or in combination with other builders to contribute unique properties for final product performance enhancement.

Phosphates. Tetrasodium pyrophosphate ($Na_4P_2O_7$, TSPP) became commercially available in the mid-1930s and soon gained acceptance as the preferred builder for soap-formulated products. The combination of TSPP and soda ash was more effective than either ingredient used separately. By the mid-1940s, sodium tripolyphosphate ($Na_5P_3O_{10}$, STPP) was commercially available and had supplanted TSPP as the preferred builder because of superior detergent processing, solubility, and hard-water sequestering properties. Sequestration is defined as the reaction of a cation with anion to form a soluble complex. The sequestration of Ca^{++} and Mg^{++} ions leads to softened water which results in better surfactant solubility and improved cleaning performance. STPP was the major builder ingredient in heavy-duty laundry detergents until the 1970s when phosphates were blamed for eutrophication of lakes and streams and thus have been almost completely removed from commercial laundry detergents. STPP is still widely used in automatic dishwashing detergents and in I&I cleaning applications.

Other possible builder salts include tetrapotassium pyrophosphate, sodium trimetaphosphate, trisodium and tripotassium phosphates, and sodium acid pyrophosphates. Significant literature reviews cover this subject in great detail [31, 45].

Silicates. Both sodium and potassium silicate solids or solutions have valued functionality including emulsification, buffering, deflocculation, and antiredeposition ability. Silicates also provide corrosion protection to metal parts in washing machines, as well as to the surfaces of china patterns and metal utensils in automatic dishwashers. Silicates are manufactured in liquid, crystalline, or powdered forms and with different degrees of alkalinity. The alkalinity of the silicate provides buffering capacity in the presence of acidic soils and enhances the sequestration ability of the builder system in the formulation. The silicate/alkali ratios of the silicates are selected by the formulator to meet specific product requirements. Silicate ratios of 1/1 are commonly used in dry blending applications with silicate ratios of 2/1 and higher commonly used in laundry and autodish applications.

Carbonates. In most areas of the United States, phosphates have been legislatively banned from use in home laundry detergents. Sodium carbonate, commonly called soda ash (Na_2CO_3), or a combination of soda ash and zeolite has replaced STPP as the builder in granular laundry products. Sodium carbonate softens water by precipitating hardness ions as insoluble carbonate salts. It is one of the most economical sources of alkalinity for detergents and provides some degree of soil dispersion and suspending action. Sodium carbonate is commonly used in powder laundry detergents, autodish powders, and hard surface cleaners.

Sodium Citrate. Sodium citrate ($Na_3C_6H_{25}O_7$) has builder characteristics due to its ability to sequester hard-water ions and deflocculate soils. Sodium citrate's principal use is in liquid detergents and personal care products.

Zeolites. Zeolites are an important alternative builder for powdered laundry detergents and structured surfactant systems. Zeolites are crystalline hydrated aluminosilicates of Group I and Group II elements. They are a complex mixture of sodium, potassium, calcium, and magnesium salts, and appear in natural and synthesized forms. Zeolite 4A ($Na_2{\cdot}Al_2O_3{\cdot}2SiO_2{\cdot}4.5H_2O$, commonly called Zeolite A) is the most widely used form of laundry detergents.

Zeolites are not water-soluble but they reduce calcium water hardness by ion exchange. The calcium ion passes through the zeolite pore openings and exchanges with the sodium ion. Magnesium ions are not significantly removed by zeolites because of the large atomic radium of the Mg^{++} ion. Zeolites are commonly used in combination with other builders such as sodium carbonate to control magnesium and other ions present in the wash water.

Sodium Nitrilotriacetate. One of the first compounds developed as a replacement for STPP was sodium nitrilotriacetate ($N(CH_2CO\ ONa)_3{\cdot}H_2O$—NTA), which has excellent sequestration and/or chelating properties. This product was quickly accepted by the detergent industry as an alternative builder. However, its use in detergents was suspended in the United States in 1970, when three major detergent manufacturers voluntarily agreed to discontinue its usage due to preliminary data suggesting that the compound could be teratogenic. Subsequent assessment by the EPA indicated that the risk to persons exposed to NTA from detergents is not significant. However, the EPA stated that NTA should not be used in products such as shampoos, hand-dishwashing detergents, foods, or other consumer products designed for direct dermal or oral exposure. Although NTA has not been used in the US household laundry products since the voluntary ban, it has been widely used as a detergent builder in Canada since 1970, and in the United States in I&I cleaning products. NTA is normally used in liquids as the primary builder or in combination with phosphates in spray-dried detergents.

Sodium Sulfate. Although considered a weak builder by some manufacturers, sodium sulfate (Na_2SO_4) contributes little to detergent performance. It is commercially available from natural sources and as a by-product from rayon processing. In dry mix and agglomerated type products, sodium sulfate is used to improve finished product flow characteristics. In spray-dried products, sodium sulfate acts as an inert filler and aids in density control.

Sodium Chloride. Sodium chloride (NaCl) is used by some detergent manufacturers, its main function being as an inert filler or diluent. In spray-drying operations, NaCl is used to control Crutcher slurry viscosity and the density of the spray-dried bead or granules. In liquid formulations, NaCl is used to control product viscosity through the salt effect. Although salt is used to some extent in autodish formulations, it is not recommended because of potential machine and dishware corrosion.

Antiredeposition Polymers. The surfactants employed in synthetic detergents are effective in removing soil and stains from the surface of the washed fabric. Under heavy soil loading, there is a tendency for these soils to redeposit on the fabric. Various charged polymers have been found to prevent soil redeposition. Sodium carboxymethylcellulose (NaCMC) is effective in reducing soil redeposition at a formulation level of 0.1–0.5%. The sodium salt of polyacrylic acid is also an effective antiredeposition polymer. Molecular weights of 5,000 g/mole give a good combination of performance-handling characteristics. Copolymer of styrene and maleic anhydride are also commonly used, although the optimum molecular weight is approximately 45,000. The polyacrylates are typically used at 0.5–1% in the formulation.

Optical Brighteners. Although the surfactant system in formulated detergents is effective at cleaning clothes, white fabrics tend to gray or lose their white appearance after multiple wash cycles. Optical brighteners are added to detergent formulations to visually improve the whiteness or brightness of the fabric.

All major liquid or powdered detergent products contain one or more fluorescent whitening agents [32]. Substantive to most fabrics, these materials absorb ultraviolet radiation and convert it to visible blue-white light that masks the natural yellowing or graying of the fabric. Most brightener systems are more effective on cotton than on synthetic fabrics. The usage range varies from 0.05 to 0.2% in the detergent formulation.

Hydrotropes. Hydrotropes are used primarily in liquid detergents to increase the solubility of sparingly soluble ingredients and improve storage stability. They are also used occasionally as viscosity modifiers in the slurry phase of manufacturing spray-dried detergent powders. Reduction of the detergent slurry viscosity improves its processing characteristics, reduces product bulk density, and insures

Table 36.6 Spray-dried detergent formulation

Ingredients	Phosphate built detergent wt%	Nonphosphate built detergent wt%
Sodium tripolyphosphate	20–35	0
Sodium carbonate	5–10	5–10
Zeolite A	0	0–20
NaLAS	0–15	0–20
Alcohol ethoxylate	0–15	0–20
NaAES	0–5	0–5
Sodium silicate	3–6	3–6
Sodium sulfate	10–30	10–30
NaCMC	0	0.1–1
Enzymes	0.1–1	0.1–1
Optical brightener	0.1–0.3	0.1–0.3
Bleach	0–5	0–5
Water	3–8	3–8

crisp, free-flowing granules. The most common hydrotropes are the ammonium, potassium, or sodium salts of toluene, xylene, or cumene sulfonates.

Enzymes. A number of different enzymes have been developed for use in detergent systems to remove common soils and stains. Protease enzymes are used to cleave the protein-based soils into smaller pieces more easily removed by the detergent system. Lipase enzymes are used to cleave lipid bonds resulting in fatty acid and glycerol. Lipase enzymes work better at elevated temperature and tend to become active in the drying operation. Noticeable removal of stains occurs after multiple wash cycles. Amylase enzymes tend to break carbohydrate bonds in sugar-based soils. Cellulase enzymes are designed to remove microfibulas in cotton fabric to help retain color intensity and appearance. Enzymes are available in liquid and powder form. Care must be taken when handling solid enzymes that the powder does not become airborne where it can be breathed in by production workers. Enzymes are used in both liquid and powder laundry detergents.

Pearlescent, *Opacifying*, *Thickening Agents*. These compounds are used in liquid detergent systems to produce a specific aesthetic effect. Pearlescent compounds, such as mica or ethylene glycol distearate (EGDS), are added to give a luxurious appearance to liquid formulations. Opacifying agents include water-soluble salts of styrene and maleic anhydride and are used to reduce translucence, modify the viscosity characteristics of liquids, and provide a creamlike texture to the product. Polymeric thickeners and clays are used to modify the rheological properties of liquids for various applications where cling to vertical surfaces is beneficial.

Anticaking Agents. The use of nonphosphate builders in laundry powders has changed the flow characteristics and caking tendencies. Products formulated with increased anionic surfactant levels and higher sodium carbonate content tend to exhibit compaction in storage. Anticaking agents have been added to the formulation to reduce this tendency, especially in warm humid climates. A wide variety of products is available, including sodium benzoate, tricalcium phosphate, colloidal aluminum oxide, silicon dioxide, magnesium silicate, calcium stearate, and microcrystalline cellulose.

Production of Synthetic Detergents

Modern detergent products come in a variety of different formats including liquids, powders, tablets, sachets, and wipes. Product format preferences change with different regions of the globe. Liquid laundry detergents account for 70% of the market in North America whereas powders are preferred in Europe and Japan. Tablets and sachets are a relatively new format which is very popular in the United Kingdom and other countries in the European Union. Recently, wipes have become popular especially for hard-surface cleaners. The following is a brief description of the common manufacturing processes.

Spray Drying. A major portion of powdered laundry detergents is spray-dried. In this process, the detergent raw materials are mixed with water to form a thick paste or slurry, atomized into spherical droplets, and dried to produce a free-flowing granular product. The art of spray drying was first patented in 1883, but was not practiced to any extent until the 1930s. With the advent of synthetic detergents in the 1940s, spray drying became the dominant method of detergent processing. The advantages of spray drying are numerous: the product density can be varied from 0.20 to 0.80 g/cm^3, the detergent granule solubility is significantly improved, very high production rates are easily achieved, and multiple formulations can be prepared using the same equipment. The major disadvantages of spray drying are the initial capital investment required and the energy-intensive nature of the process. A modern spray-drying operation consists of several unit operations: slurry preparation, atomizing, drying and conditioning of the product, post-additions, packaging, waste reclamation, and pollution control.

Spray-dried powders are formulated to give the best overall cleaning performance and flow characteristics at the lowest possible formulation costs. Typical spray-dried detergent formulations are shown in Table 36.6. Slurry preparation is considered to be the single most important factor in spray drying. Control of the raw material order of addition, slurry solids level, slurry temperature, viscosity,

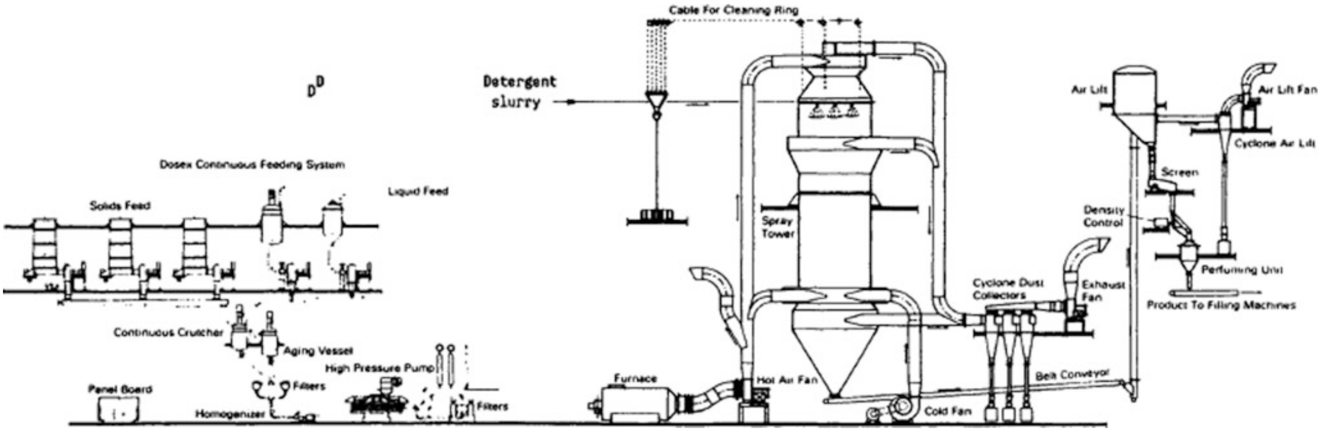


Fig. 36.41 Spray tower and slurry preparation

and aeration enable manufacturers to deliver the desired final product characteristics. The spray tower itself is simply a contained heat source that removes moisture at a uniform rate dependent upon the quality and the quantity of feed to the dryer. If feed variables are not controlled, variability in the density and moisture levels of the powder will vary. For example, if the slurry solids level is not constant (±2%), the evaporative load, production rate, particle sizing, and moisture levels of the spray-dried granules will vary. Compensating for nonuniform feed to the spray tower is difficult because of the interaction of control variables such as tower temperature, pump pressure, and gas velocity. Although some minor tower adjustments are possible, the major control for routine production is consistent and uniform preparation of the detergent slurry.

Slurries are prepared via batch, semi-continuous and continuous crutching units. Solids levels are normally kept as high as possible (64–72%) to reduce the heat requirements and maximize the tower throughput. After the detergent raw materials are mixed, the slurry is charged to a homogenizer and any large lumps or gritty particles are desized and screened to prevent plugging of the spray nozzle. The slurry is then charged to a booster pump, deaerated, and sent to a high pressure pump (300–1,200 psig) for atomization.

Most detergent slurries are atomized into spherical droplets via high pressure nozzles. The nozzles produce a hollow conical pattern with a spray angle of 45–90°. The slurry emerges from the nozzle as a film, which is sheared into droplets by the centrifugal force developed by specially grooved cores inserted into the nozzle before the discharge point. A series of spray nozzles (2–25) is placed near the top of the spray tower and angled so that one edge of the desired spray pattern is vertical to the spray tower side walls. Some spray towers use two or three rings of nozzles, depending upon the formulation and drying capacity of the unit. Production rates of detergent spray towers range from 5,000 to 80,000 lbs/h. In most cases, the rates are controlled by the slurry feed facilities, solids levels, and the drying capacity. A schematic of a typical spray drying facility is shown in Fig. 36.41. Spray towers vary in size from approximately 10 to 35 ft in diameter and range from 40 to 200 ft in height. The majority of detergent spray towers use countercurrent airflow patterns although some use a concurrent in design. In general, a countercurrent tower will produce a more high-density particle than the concurrent type. In countercurrent systems, the heated gas stream (250–270°C) is uniformly dispersed into the lower section of the tower through a series of plenum chambers or baffles equally spaced around the unit's periphery. The hot gases are cooled as they exhaust the top of the tower by the drying of atomized detergent droplets upon contact with the hot air stream. Because of the relatively low temperature (110–140°C) at this point in the tower, the particles do not expand to their full capacity as water evaporates. This lower expansion rate results in a bead with increased bulk density and thicker shell. If a similar formulation were dried in a concurrent spray tower, the higher temperature (250–270 °F) and increased drying rate would increase the evaporative rate and result in a thin-walled bead of lighter density.

In most spray towers, the air flow rate and the temperature are adjusted at the inlet ports to maintain balanced air/temperature patterns. When the atomized slurry is sprayed into the drying chamber, the air flow pattern is disturbed and becomes very turbulent. This effect can cause localized temperature gradients (hot or cold spots) and affect the particle drying characteristics. It is possible to have both over- and under-dried particles discharged at the same time from the tower because of poor adjustment of the air flow in the tower. Adjustment of the airflow patterns, as well as sending a balanced spray into the tower, can help counteract this situation. Increased impingement of the product on the tower walls is another negative effect of excessive turbulence and of excessive spray nozzle velocity. A ring of product may form approximately 10 to 20 ft below the

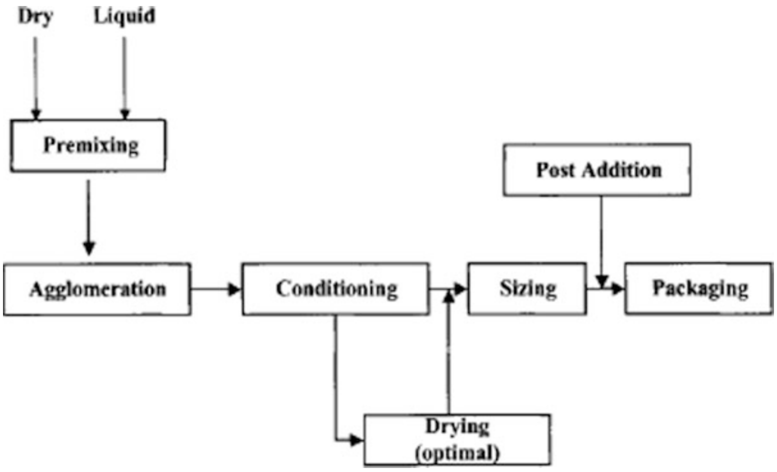


Fig. 36.42 Unit operations for batch agglomeration

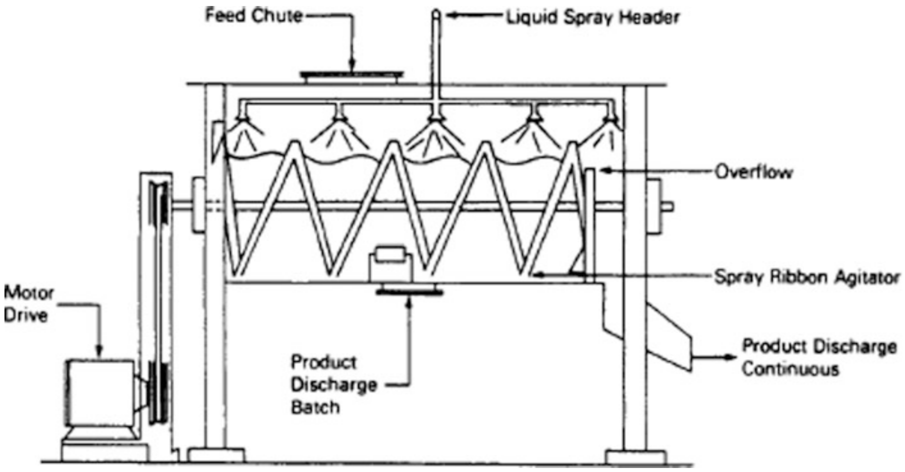


Fig. 36.43 Ribbon blender

spray nozzles and build in size until it cracks or scales off the walls. Ring removal methods, such as periodic cleaning using air or water lances, vibration, sonic shocks, or mechanical cleaning rings, sometimes are used to prevent a buildup that could result in blockage of the tower discharge.

After drying in the spray tower, the detergent granules are conveyed to storage silos prior to packaging. In some processes, product storage is preceded by fluid-bed conditioning. The product for the tower is very warm (75–120°C) and typically contains moisture levels of 1–3%. Conditioning is employed to prevent caking or product flow problems due to either the high temperature or excessive free moisture content. Following conditioning, perfume and other heat-sensitive ingredients (enzymes, surfactants, bleaches) may be added to the product. Post-addition is accomplished by metering the additives into the spray-dried base in a rotary drum or other mixing device. The post-additions normally comprise 2–15% of the total detergent formulation. An excellent review of spray drying and key patents is available. [32, 33]

Agglomeration

Agglomeration is a technique of binding a mixture of granular and/or dry powdered raw materials into an attractive granular form with enhanced aesthetic and flow characteristics. The majority of granular automatic-dishwashing compounds and some laundry powders are manufactured using the agglomeration process. The advantages of agglomeration are low capital costs relative to spray towers and reduced energy consumption. Agglomeration also yields a higher-density product, which results in reduced packaging costs. The major disadvantage of agglomeration is the limited surfactant loading due to raw

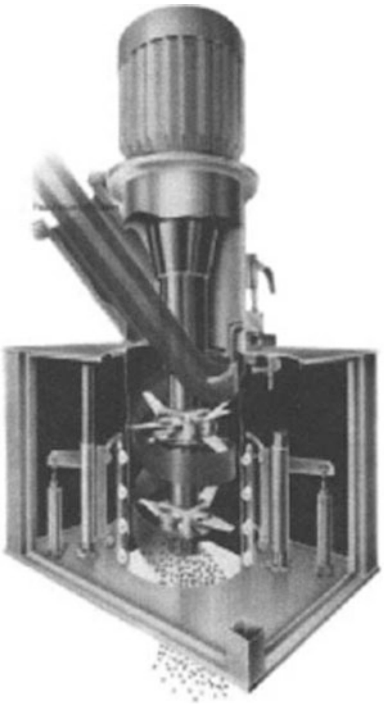


Fig. 36.44 Shugi Flex-o-Matic agglomerator

material absorption characteristics, particularly in laundry applications. As energy costs have assumed increasing importance, and higher-density products have gained in popularity, agglomerated or combinations of agglomerated and spray-dried powders have found wider use.

The typical process steps for formulating a laundry or autodish detergent usually are defined in the laboratory prior to continuous plant processing; however, they generally follow the scheme shown in Fig. 36.42. Frequently the dry raw materials are premixed and then metered into the agglomeration equipment. In the agglomerator, the surface of the dry raw materials is exposed to atomized sprays of surfactant or silicates to begin building the agglomerated particle. Following addition of liquids to the dry raw materials, the product has a wet or pasty texture and requires conditioning to remove excess water prior to storage or packaging. Conditioned agglomerates are screened to remove oversize particles and then blended with other additives (bleaches, enzymes, dyes, and perfumes) that would not be stable under agglomeration conditions.

Numerous equipment types can be used for agglomeration, including horizontal ribbon and vertical mixers, rotary drums, zig-zag mixers, and pan agglomerators. Horizontal ribbon mixers as shown in Fig. 36.43 are used primarily for dry mix formulating, where only limited agglomeration occurs. They comprise a U-shaped trough and a ribbon or paddle-type agitator. Liquids are sprayed onto the surface of the dry material and slowly dispersed into the mix. Mixer operations can be either batch or continuous.

Vertical mixers have become more commonly used due to low space requirements and high efficiency of operations. The vertical mixer shown in Fig. 36.44 utilizes high speed agitation to intimately mix and uniformly blend liquids and dry raw materials prior to discharge. The contact time is reduced to less than 5 s. The rotary drum units generally contain baffles that roll or lift the product for uniform dispersion or liquid feeds. Pan and zig-zag agglomerators generally are used for limited, specialty product manufacture.

The vertical and the rotary drum agglomeration equipment find most widespread use in the detergent industry because of improved agglomeration, increased liquid feed dispersion levels, and optimized product uniformity. In addition to agglomeration of detergent and autodish detergents, these units can be used for hydration of STPP, dry neutralization of sulfonic acids, post-addition of raw materials and fines to spray-dried products and particle encapsulation. An excellent review of detergent agglomeration is available [46, 47].

Liquid Detergent Processing

Liquid heavy-duty products account for approximately 60% of the 2001 US laundry detergent market [48]. The growth of heavy-duty liquids began in 1974, as the result of the phosphate ban in home laundry. Unbuilt liquids based on nonionic and anionic/nonionic surfactant combinations gradually increased in volume for various reasons including convenience, solubility, and improved performance vs. nonphosphate containing powdered detergents. In the mid-1980s the introduction of higher-performance liquids accelerated growth in this segment, with the market share essentially doubling from 1984 to 1990 [49].

Liquid products are approximately 50–60% water, with the remainder being a combination of surfactants, builders, foam regulators, enzymes and enzyme stabilizers, hydrotropes, antiredeposition polymers, optical brighteners, corrosion inhibitors, dye, and perfume. Two-in-one formulations also contain antistatic and fabric softening ingredients. Liquid systems require careful selection and blending of raw materials to achieve a stable product. Special attention is necessary for the following items.

- *Viscosity*. The product must be pourable and must retain the same viscosity characteristics from batch to batch. This variable can be controlled by proper solvent or hydrotrope selection.
- *Clear-cloud point*. The composition should have sufficient solubility to prevent hazing or separation of the product when subjected to storage in cool temperatures.

- *Freeze–thaw stability*. The formulation must be compounded to prevent phase separation or solidification at freezing temperatures.

Most LDL detergents are formulated using a combination of LAS and AES with various amphoteric foam stabilizers (CAPB, amine oxides, etc.). Heavy-duty liquid detergents often use a combination of actives such as LAS, AES, AE, and AOS. The manufacture of liquid products requires relatively basic equipment for mixing the various ingredients. The raw materials are added batchwise to a large mixing vessel or proportionated to in-line mixers before being fed to the filling lines. In 1986, liquid autodish detergents were introduced. These products were based on thixotropic slurries of STPP and soda ash. The rheology helped to solve issues with product dispensing rate and product physical stability. The slurries were generally viscous enough to prevent premature dispensing into the dishwasher prior to the wash cycle. The thixotropic nature of the product was achieved using natural clays of high molecular weight polymers. High-shear mixing equipment such as homogenizers is used to manufacture these products in batch or continuous stirred tank systems.

Detergent Trends

In the United States, liquid detergent sales continue to grow at the expense of powders. The trend is largely due to convenience as cost per load for liquids is significantly higher than for powders. In the United States, liquid products are clear and isotropic with vibrant colors. In Europe, liquid products have a lower market share and are typically based on structured surfactant systems with suspended STPP and/or zeolites. These systems tend to be opaque products and it remains to be seen if the US market will accept these products. The performance of liquid laundry detergent performance has been increased through the use of multiple enzyme and ancillary ingredients.

Recently, unit dose laundry tablets and liquid sachets have been introduced in the United States, which take customer convenience to the next step. Presently, the market share for these products is small and it also remains to be seen how these products will fare in the US market.

The machines used to wash clothes are also changing. The DOE has mandated energy efficiency standards on new washing machines. In order to meet these efficiency standards, most manufacturers are offering horizontal-axis washing machines similar to those that have been available in Europe for many years. The H-axis machines use less water and energy than conventional top-loader machines; however, they are currently more expensive than conventional washing machines. The new machines generate more mechanical actions, which in turn generates more foam. Detergent manufacturers now must formulate lower foam detergents for use in these machines.

The environmental movement, which began in Europe, is now beginning to appear in the United States. The use of NPE in household products is declining. To a large extent, NPEs have been replaced with alcohol ethoxylates. The use of environmentally acceptable surfactants continues to grow. Surfactants based on natural, regrowable raw materials are commercially available. APG based on sugar and fatty alcohol, MESs, and MEEs are examples of naturally derived surfactants. This trend can be expected to continue into the next decade.

References

1. (1981) Soaps and detergents. The Soap and Detergent Association, New York
2. Spitz L (1996) Soaps and detergents: a theoretical and practical approach. AOCS Press, Champaign, IL
3. Gupta S (1990) Chemistry, chemical and physical properties and raw materials. In: Spitz L (ed) Soap technology for the 1990's. American Oil Chemists Society, AOCS, Champaign, IL, pp 48–93
4. Jungermann E (1979) Soap. In: Swern D (ed) Bailey's industrial oil and fat products, vol 1, 4th edn. Wiley, New York, pp 511–585
5. Ghaim JB, Volz ED (2001) Skin cleansing bars. In: Barel AO, Payne M, Maibach HI (eds) Handbook of cosmetic science and technology. Marcel Dekker, Inc., New York, pp 485–497
6. By courtesy of G. Mazzoni S.p.A., Busto Arsizio, Italy. See also Spitz, L (1967) J Am Oil Chem Soc 45:423
7. Dieckelmann G, Heinz HJ (1988) The basics of industrial oleochemistry. Peter Pomp GmbH, Essen, W. Germany
8. Woollatt E (1985) The manufacture of soap, other detergents and glycerin. Ellis Horwood, Sussex, England
9. By courtesy of Alfa-Laval, Tumba, Sweden
10. Sonntag NOV (1979) J Am Oil Chem Soc 56:729A, 861A
11. Markley KS (1967) Fatty acids, 2nd edn. Interscience, New York
12. By courtesy of Lurgi GmbH, Frankfurt am Main, Germany
13. Gosewinkel L, Knuth M (1988) Fett Wiss Technol 90:155
14. Stoiculescu P, Svet V (1971) Ind Aliment (Bucharest) 22:320
15. Leshchenko PS (1973) Maslo-Zhir Prom 11:21
16. Brady CL, Metcalfe L, Slaboszewski S, Frank D (1988) J Am Oil Chem Soc 65:917
17. Kwon DY, Rhee JS (1984) Korean J Chern Eng I:153
18. Park YK, Pastore GM, de Almeida MM (1988) J Am Oil Chem Soc 65:252
19. Holmberg K, Osterb E (1988) J Am Oil Chem Soc 65:1544
20. Mange HK, Vaidya SD (1979) J Oil Technol Assoc India 11:73
21. Vaidya SD, Subrahmanyan VVR, Kane JG (1975) Indian J Technol 13:528
22. Berger R, McPherson W (1979) J Am Oil Chem Soc 56:743A
23. Jach KW, Stage H (1988) Fett Wiss Technol 90:501
24. Stage H (1984) J Am Chem Soc 61:204
25. Jach KW, Stage H (1988) Fett Wiss Technol 90:501
26. Haradsson G (1984) J Am Oil Chem Soc 61:219
27. Zilch KT (1979) J Am Oil Chem Soc 56:739A
28. Luddy FE (1979) J Am Oil Chem Soc 56:759A

29. Haupt DE (1984) Soap Cosmet Chem Spec 60:42
30. Leikham JW (1988) J Am Oil Chem Soc 89
31. Henkel KGA (1982) Fatty alcohols: raw materials, process, applications. Schneider & Hense GmbH, Dusseldorf
32. Swisher RD (1970) "Surfactant biodegradation", surfactant science series, vol 3. Marcel Dekker, New York
33. Huddleston RL, Nielsen AM (1979) Household & personal products industry. p 72
34. Gledhill WE, Huddleston RL, Kravetz L, Nielson AM, Sedlak RL, Vashon RD (1989) Tenside 26:276
35. Schoberl P (1989) Tenside 26:86
36. Linfield WM, Barauskas RA, Silvieri L, Serota S, Stevenson RW Sr (1984) J Am Oil Chem Soc 61:191
37. Linfield WM, O'Brien DJ, Serota S, Barauskas RA (1984) J Am Oil Chem Soc 61:1067
38. Brooks RJ et al (1978) U.S. Patent 4,113,428, September 12, 1978
39. Brooks RJ (1980) U.S. Patent 4,185,030, January 22, 1980
40. Hennig H et al (1957) U.S. Patent 2,802,026, August 6, 1957
41. Logan RS (1965) U.S. Patent 3,188,183, June 8, 1965
42. Schick MJ (1987) "Nonionic surfactants", surfactant science series, vol 23. Marcel Dekker, New York
43. Bluestein BR, Hilton CL (1982) Amphoteric surfactants, surfactant science series, vol 12. Marcel Dekker, New York
44. Lomax EG (1996) Amphoteric surfactants, surfactant science series, vol 59. Marcel Dekker, New York
45. Griffin WC (1954) J Soc Cosmet Soc 5:249
46. De Groot WH et al (1995) The manufacture of modern detergent powders. Herman de Groot Academic, The Netherlands
47. Dolan MJ (1988) Detergent agglomeration technology. Soap Chem Spec
48. Marchie M (2002) The laundry detergent market. Happi 76
49. Cannon DR (1987) Chemical week. p 22
50. Griffin WC (1949) J Soc Cosmet Soc 1:311
51. Brooks RJ et al (1966) U.S. Patent 3,257,715, June 21, 1966
52. Brooks RJ et al (1996) U.S. Patent 3,259,645, July 5, 1966
53. Ballestra M (1965) U.S. Patent 3,180,699, April 27, 1965
54. Lantini A (1976) U.S. Patent 3,931,272, January 6, 1976
55. Swartz AM, Perry JW (1949) Surface active agents, vol 1. Wiley, New York
56. Jungermann E (1979) "Cationic surfactants", surfactant science series, vol 4. Marcel Dekker, New York
57. Richmond JM (1990) "Cationic surfactants", surfactant science series, vol 34. Marcel Dekker, New York
58. Wazer V (1961) Phosphorus and its compounds, vol II. Wiley, New York
59. Sorgenfrei ME (1980) Increased use of STPP. Soap Cosmet Chem Spec 3(2):60
60. Findley WR (1988) JAOCS 65(4)
61. Sittig M (1979) Detergent Mfg. Including Zeolite Builders. Chem Tech Review No. 128. Noyes Data Corporation, New Jersey

Chemical Explosives

37

Don H. Cranney and Walter B. Sudweeks

Introduction

The average citizen in today's world gives little thought to the important role that commercial explosives play in their lives and how their use is linked to our standard of living and our way of life. Explosives provide the energy required to give us access to the vast resources of the earth for the advancement of civilization. In 2010, the Mineral Information Institute estimated that the average baby born in America will need the following quantities of minerals, metals, and fuels in their lifetime: copper—932 lb; salt—31,779 lb; clays—12,121 lb; zinc—544 lb; stone, sand, and gravel—1,100,000 lb; petroleum—72,499 gal; lead—777 lb; other minerals and metals—43,822 lb: natural gas—5.93 million ft^3; cement—41,181 lb; iron ore—14,530 lb; bauxite (aluminum)—4,040 lb; coal—542,968 lb; phosphate rock—15,152 lb; and gold—1.383 troy oz [1]. Availability of all of these materials, which total 2.9 million lb/individual, depends on the use of explosives.

Without explosives, the steel industry and our entire transportation system would not be possible. The generation of electricity has been largely dependent on coal, and coal mining today is still the largest consumer of industrial explosives. Rock quarrying for road building and cement production, excavations for skyscrapers, tunnels, roads, pipelines, and utilities are direct beneficiaries of the labor-saving use of explosives. Exploration for oil and gas involves the use of explosives in seismic studies to map underground strata to find structures that may have these resources present. Our food supply depends on the use of phosphate based fertilizers. Significantly, the industrialization cycle being experienced in several areas of the world, such as in China and other parts of Southeast Asia, is greatly ramping up the use of explosives worldwide to supply the needed materials for the massive building and infrastructure improvements.

Commercial Explosives Market

The use of commercial explosives in the United States over the last 10 years shows some cyclical behavior that basically follows the economic cycles. This is illustrated in Fig. 37.1, which indicates explosives usage by year as reported by the U.S. Geological Survey (USGS) [2] from 2000 to 2009. The data compiled in this figure show fairly level explosives sales for the years 2000 through 2004, hovering around 2.5 million metric tons/year. The years 2005 through 2008 show a distinct step up to volumes over three million metric tons consistent with the world wide commodities boom period and the booming U.S. economy during that period. This was followed by a sharp drop to below 2.5 million tons as the world wide recession of 2008/2009 sharply dropped explosives demand. Figure 37.1 also separates the volumes by industry use. The open-pit coal mining industry continues to be the largest user, as it has been for many years.

Table 37.1 shows the commercial explosive usage by the ten leading states for 2008 and 2009 [3]. These ten states represent 75% of the total explosive use in the U.S. In the year 2009 (the latest statics available at the time of this update), the four states consuming the most explosives (in decreasing order) were: Wyoming, West Virginia, Kentucky, and Indiana; all coal mining states. These states alone accounted for over 50% of the total U.S. market. The coal market continues a shift from the eastern states to the western states with lower BTU, but more importantly, lower sulfur coal. This trend is congruent with the shift to large open-pit mining operations, both in the west and in other regions, in order to increase production and competitiveness. The dramatic effect of the economic downturn was indicated

D.H. Cranney (✉)
Senior R& D Mgr./Global Explosives Team, Dyno Nobel Inc., 10535 South Featherwood Drive, South Jordan, UT 84095, USA
e-mail: don.cranney@am.dynonobel.com; don.patc@comcast.net

W.B. Sudweeks
Retired 680 East 3230, Utah 84043, North Lehi

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_37, © Springer Science+Business Media New York 2012

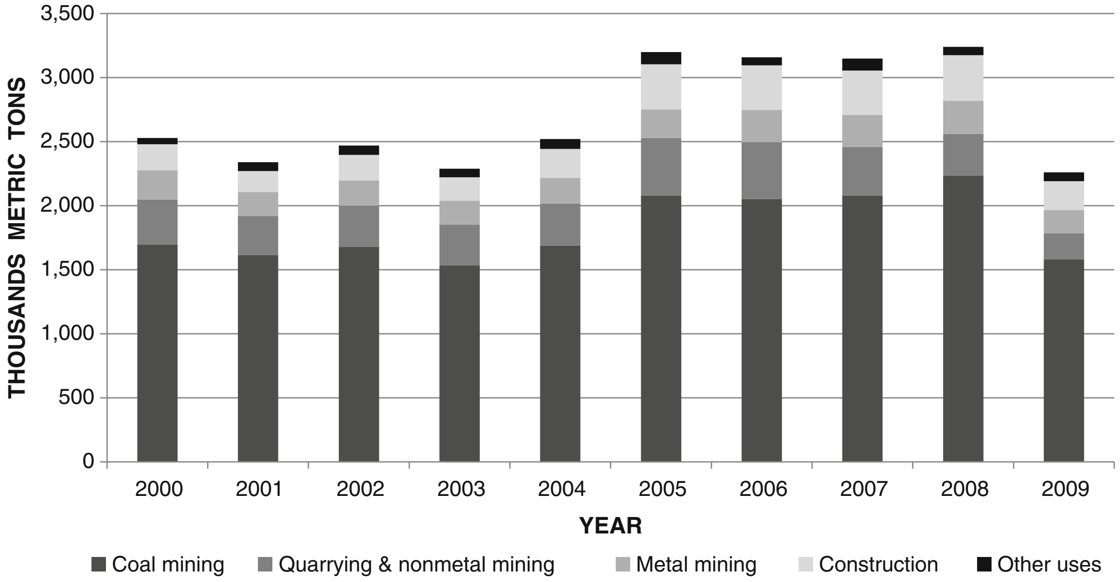


Fig. 37.1 Total sales of U.S. Industrial Explosives from 2000 to 2009. Relative volumes within major market segments indicated (Compiled from United States Geological Survey [2])

Table 37.1 Industrial explosives and blasting agents sold for consumption in the United States, by state and class[a]

Top ten consuming states only, in descending order (2009 compared to 2008) (Metric tons)

	2008				2009			
	Fixed high explosives		Blasting agents and oxidizers		Fixed high explosives		Blasting agents and oxidizers	
State	Permissibles	Other		Total	Permissibles	Other		Total
Wyoming	–	517	675,000	377,000	29	240	377,000	378,000
West Virginia	290	1,240	423,000	348,000	154	758	348,000	349,000
Kentucky	308	2,010	386,000	292,000	206	1,330	292,000	293,000
Indiana	–	975	187,000	189,000	70	980	189,000	190,000
Virginia	348	945	166,000	114,000	299	1,560	114,000	116,000
Alabama	25	301	87,300	95,900	14	257	95,900	96,200
Pennsylvania	64	1,290	102,000	90,400	36	1,390	90,400	91,800
Utah	19	321	94,800	64,100	43	219	64,100	64,400
Texas	(3)	1,470	118,000	62,700	18	699	62,700	63,500
Montana	–	1,880	73,500	55,200	–	2,010	55,200	57,200
Total	*1,053*	*10,950*	*2,312,600*	*1,688,300*	*869*	*9,444*	*1,688,300*	*1,699,100*
Total, all states	*1,200*	*35,800*	*3,380,000*	*2,240,000*	*1,610*	*23,700*	*2,240,000*	*2,270,000*

Source: Apodaca [3]
Data Source: Institute of Makers of Explosives
[a] Data are rounded to no more than three significant digits
(3) Less than 1/2 unit

by the drop of Nevada from the 6th position in the leading explosives consumers in 2008 to the 20th position in 2009 (not shown in Table 37.1). This was due to the slump in gold mining, which at the time of this update is recovering strongly. Other metal and commodity mines are also in a positive upswing, while construction and quarry (aggregate) operations remain in a slump.

Chemistry of Combustion and Explosion

For a simple understanding of explosives, it is helpful to compare an explosive reaction with the more familiar combustion or burning reaction. Three components are needed to have a common fire: fuel, oxygen from the air, and a source of ignition. The process of combustion is basically an oxidation–reduction (redox) reaction between the fuel and oxygen. Once initiated, this reaction can become self-sustaining, producing large volumes of gases and heat. The heat given off further expands the gases and provides the stimulus for the reaction to continue by heating and igniting surrounding fuels.

The basic burning reaction is a relatively slow, diffusion-controlled process that occurs within the flames or near the surface of glowing embers. The size of the fire depends upon how much fuel is involved and on the rate of the combustion reaction. The rate of the combustion reaction depends on how finely divided the fuel is and how rapidly the oxygen reaches the flame, that is, the intimacy of contact between the fuel and the oxygen in the air. Burning rate is greatly increased when convection of the air, natural (wind) or man-made (fanning the flame), joins diffusion in supplying oxygen to the flame.

Another result of an intimate mixture of the fuel and air is the completeness or efficiency of the reaction. In a complete combustion all the fuel elements are oxidized to their highest oxidation state. Thus, burning of wood, being mainly cellulose and gasoline being generally a hydrocarbon (e.g., octane), produces primarily carbon dioxide and water vapor upon complete combustion. Once initiated, these burning reactions give off heat energy, which sustains the reactions. Heat is released because the oxidized products of the reaction are in a lower energy state (more stable) than the reactants. The maximum potential energy release can be calculated from the respective heats of formation of the products and reactants. Actual heats of combustion can be measured experimentally by causing the reaction to occur in a bomb calorimeter. The calculated energy values for the above reactions are 3,857 cal/g for cellulose and 10,704 cal/g for octane.

In the case of an inefficient burn, some less stable or higher-energy products are formed so that the resultant heat energy given off is lower than that for complete combustion. In the above examples, inefficient combustion could result from lack of accessible oxygen, producing carbon monoxide or even carbon particles instead of carbon dioxide. A smoky flame is evidence of unburned carbon particles and results from inefficient combustion where fuel particles are so large or so dense that oxygen cannot diffuse to the burning surface fast enough. If this inefficiency is great enough, insufficient heat is given off to keep the reaction going, and the fire will die out.

All chemical explosive reactions involve similar redox reactions; so the above principles of combustion can help illustrate, in a very basic way, the chemistry involved in explosions. As in a fire, three components (fuel, oxidizer, ignition source) are needed for an explosion. Figure 37.2 shows an explosion triangle, which is similar to a fire triangle. In general, the products of an explosion are gases and heat, although some solid oxidation products may be produced, depending upon the chemical explosive composition. As in normal combustion, the gases produced usually include carbon dioxide and water vapor plus other gases such as nitrogen, again depending upon the composition of the chemical explosive.

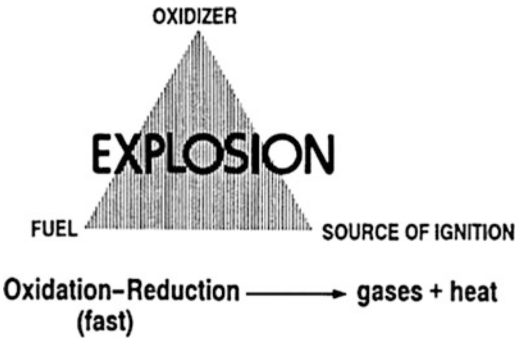


Fig. 37.2 An explosion triangle

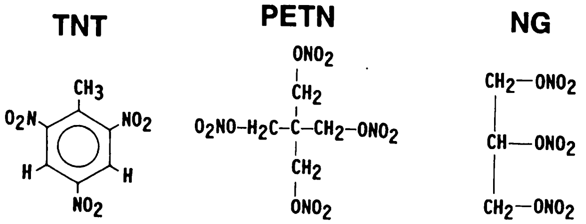


Fig. 37.3 Chemical structure of three molecular explosives

It should be noted that an explosion differs from ordinary combustion in two very significant ways. First, oxygen from the air is not a major reactant in the redox reactions of most explosives. The source of oxygen (or other reducible species) needed for reaction with the fuel—the oxidizer—may be part of the same molecule as the fuel or a separate intermixed material. Thus, an explosive may be thought of as merely an intimate mixture of oxidizer and fuel. This high degree of intimacy contributes to the second significant difference between an explosion and normal combustion—the speed with which the reaction occurs.

Explosives in which the oxidizer and fuel portions are part of the same molecule are called molecular explosives. Classical examples of molecular explosives are 2,4,6-trinitrotoluene (TNT), pentaerythritol tetranitrate (PETN), and nitroglycerin (NG) or, more precisely, glycerol trinitrate. The chemical structures of these explosives are shown in Fig. 37.3.

As can be seen in the structures, the oxidizer portions of the explosives are the nitro (–NO_2) groups in TNT and the nitrate (–ONO_2) groups in PETN and NG. The fuel portions of all three explosives are the carbon and hydrogen (C and H) atoms. Comparison of the ratios of carbon to oxygen in these explosives (i.e., approximately 1:1 for

TNT, approximately 1:2 for PETN, and 1:3 for NG) shows that TNT and PETN are deficient in oxygen; that is, there is insufficient oxygen present in the molecule to fully oxidize the carbon and hydrogen. Consequently, products such as carbon monoxide, solid carbon (soot), and hydrogen are produced, as well as carbon dioxide and water vapor. Prediction of the exact products of explosion is complex, especially for the oxygen-deficient explosives, because the amounts of CO_2, CO, H_2O, and H_2 will vary as will a host of trace products such as residual hydrocarbons, depending upon reaction conditions (explosive density, degree of confinement of the explosive, etc.) [4, 5]. The following equations show typical ideal reaction products along with calculated heats of reaction for these molecular explosives:

$$\begin{aligned}\text{TNT:}\quad & C_7H_5N_3O_6 \rightarrow 1.5CO_2 \\ & + 0.5CO + 2.5H_2O + 1.5N_2 + 5C + 1{,}290\,\text{cal/g}\end{aligned}$$

$$\begin{aligned}\text{PETN:}\quad & C_5H_8N_4O_{12} \rightarrow 4CO_2 \\ & + 4H_2O + 2N_2 + C + 1{,}510\,\text{cal/g}\end{aligned}$$

$$\begin{aligned}\text{NG:}\quad & C_3H_5N_3O_9 \rightarrow 3CO_2 \\ & + 2.5H_2O + 1.5N_2 + 0.25O_2 + 1{,}480\,\text{cal/g}\end{aligned}$$

Explosives in which the oxidizer and fuel portions come from different molecules are called composite explosives because they are a mixture of two or more chemicals. A classic industrial example is a mixture of solid ammonium nitrate (AN) and liquid fuel oil (FO). The common designation for this explosive is the acronym, ANFO. The oil used (typically #2 diesel fuel) is added in sufficient quantity to react with the available oxygen from the nitrate portion of AN. The redox reaction of ANFO is as follows:

$$\begin{aligned}& 3\underset{\text{AN}}{NH_4NO_3} + \underset{\text{FO}}{-CH_2-} \\ & \rightarrow CO_2 + 7H_2O + 3N_2 + 880\,\text{cal/g}\end{aligned}$$

"Oxygen balance" (O.B.) is the term applied to quantify either the excess oxygen in an explosive compound or mixture (beyond what is needed for complete combustion of the fuel elements) or oxygen deficiency (compared to the amount required for complete combustion). It is expressed as either a percentage or a decimal fraction of the molecular weight of the oxygen in excess (+) or deficiency (−) divided by the molecular weight of the explosive or the ingredient being considered. Individual components of an explosive mixture have O.B. values that may be summed for the mixture. Shown below are the O.B. calculations for AN and FO:

$$\underset{\substack{\text{AN}\\ \text{Mol. wt.=80}}}{NH_4NO_3} \rightarrow 2H_2O + N_2 + \underset{\substack{O_2\\ \text{Mol. wt.=32}}}{\tfrac{1}{2}O_2}$$

$$\text{O.B.} = \frac{(1/2)(32)}{80} = 0.20$$

$$\underset{\substack{\text{FO}\\ \text{Mol. wt.=}\sim 14n}}{-(CH_2)_n-} + \tfrac{3n}{2}O \rightarrow nCO_2 + nH_2O$$

$$\text{O.B.} = \frac{-(3n/2)(32)}{14n} = 3.43$$

From the O.B. values, one can readily determine the ratio of ingredients to give a zero O.B. mixture for optimum efficiency and energy. Thus, the weight ratio for ANFO is 94.5 parts of AN and 5.5 parts of FO ($94.5 \times 0.20 = 5.5 \times 3.43$). For the molecular explosives shown previously, the respective oxygen balances are: TNT, −0.74; PETN, −0.10; and NG, ±0.04. Thus, NG is nearly perfectly oxygen-balanced; PETN is only slightly negative; but TNT is very negative, meaning significantly deficient in oxygen. Therefore, combinations of TNT and AN have been employed to provide additional oxygen for the excess fuel, as, for example, in the Amatols developed by the British in World War I [6].

Modern commercial explosives react in a very rapid and characteristic manner referred to as a detonation. Detonation has been defined as a process in which a combustion-induced, supersonic shock wave propagates through a reactive mixture or compound. This high pressure shock wave compresses and interacts with the reactive material it contacts, resulting in very rapid heating of the material, initiation of chemical reaction, and liberation of energy. This energy, in turn, continues to drive the shock wave, thereby sustaining the detonation. Pressure in a detonation shock wave may reach millions of pounds per square inch. Once initiated, molecular explosives tend to reach a steady-state reaction with a characteristic detonation velocity. Composite or mixture explosives also have steady-state detonation velocities, but these velocities are more variable than those of molecular explosives and are influenced by such factors as diameter of the charge, temperature, density, and confinement.

Historical Development

The first known explosive material was black powder, a mixture of potassium nitrate (saltpeter), charcoal, and sulfur. As such it is a composite explosive whose properties are dependent upon how finely divided each of the ingredients is, and how intimately they are mixed. The exact origins of black powder are lost in antiquity. Publications referring to it seem about equally divided between those that attribute its origin to third- or fourth-century China [7, 8] and those that place it

Table 37.2 Nineteenth century explosive discoveries

1800	Mercury fulminate
1831	Safety Fuse
1846	Nitrocellulose
1846	Nitroglycerin
1847	Hexanitromannite
1862	Commercial production of nitroglycerin
1866	Dynamite
1867	Blasting cap, fuse initiated
1867	Ammonium nitrate explosive patented
1875	Blasting gelatin and gelatin dynamite
1884	Smokeless powder
1886	Picric acid
1891	2,4,6-Trinitrotoluene (TNT)
1894	Pentaerythritol tetranitrate (PETN)

closer to the thirteenth century, at about the time of Roger Bacon's written description in 1242 [9–13]. Nevertheless, its use did not become very popular until the invention of the gun by Berthol Schwartz in the early 1300s; and its first recorded use in mining did not occur for over 300 years after that. First used for blasting in 1627, the production and application of black powder played a critical role in the rapid expansion of the United States in the early nineteenth century as canals were dug and railroads built to span the continent.

For over 200 years black powder was the only blasting agent known, but the 1800s brought a number of rapid developments that led to its demise, replacing it with safer and more powerful explosives. Table 37.2 presents a chronological summary of some of the significant discoveries of the 1800s. Credit for the first preparation of NG is generally ascribed to Ascanio Sobrero in Italy in 1846.

Swedish inventors Emmanuel Nobel and his son Alfred took an interest in this powerful liquid explosive and produced it commercially in 1862. However, its transportation and its handling were very hazardous, and eventually Alfred Nobel discovered that NG absorbed into a granular type of material (kieselguhr) was still explosive, but was much safer to handle and use than the straight liquid. This new invention, called "dynamite," was difficult to ignite by the usual methods used for pure NG. Therefore, also in 1867, Alfred Nobel devised the fuse initiated blasting cap using mercury fulminate. With this development dynamite became the foundation of the commercial explosives industry.

For military and gun applications, black powder continued to be the only explosive of choice as a propellant or bursting charge until the inventions of the late 1800s, when smokeless powder, based on nitrocellulose, proved to be a cleaner, safer, and more effective propellant than black powder. The synthesis of picric acid (2,4,6-trinitrophenol) followed by TNT and PETN gave solid, powerful, molecular explosives of more uniform performance for use in bombs and artillery shells. The main explosives used in World War I were TNT, Tetryl (2,4,6-trinitrophenylmethylnitramine), and Hexyl (hexanitrodiphenylamine), and in World War II they were TNT, PETN, and RDX (1,3,5-trinitro-1,3,5-triazacyclohexane) [14].

In the industrial arena the production of black powder in the United States dropped precipitously after reaching a peak of 277 million lb in 1917 [15]. By the mid-1960s, it had ceased to be of commercial significance in the explosives industry, but during the same time period dynamite production rose from 300 to 600 million lb.

In 1947, a spectacular accident of catastrophic proportions ushered in the next revolution in explosives. Fertilizer-grade AN, in the form of prills (small spherical particles coated with paraffin to prevent caking), was being loaded into ships in Galveston Bay, TX. Along with other cargo, one of these ships, the partially loaded SS Grandcamp, contained 2,300 tons of this material. On the morning of April 16, soon after loading was resumed aboard the Grandcamp, a fire was discovered in one of the holds containing AN. Efforts to extinguish the fire were unsuccessful, and an hour later the bulk of the coated fertilizer detonated, killing 600 people and injuring 3,000 [16].

This tragedy, along with several other large-scale accidents involving AN explosions, finally led researchers to the conclusion that inexpensive, readily available, fertilizer-grade AN could be used as the basis for modern industrial explosives.

Soon after the advent of porous AN prills, introduced in the early 1950s, investigators realized that these prills could readily absorb just the right amount of FO to produce an oxygen-balanced mixture that was both an inexpensive and effective blasting agent, in addition to being safe and simple to manufacture. This technology was widely adopted and soon constituted 85% of the industrial explosives produced in the United States [17]. With ANFO's cost and safety characteristics, it became practical for surface miners to drill larger boreholes and to utilize bulk ANFO delivery systems. Nevertheless, ANFO had two significant limitations: AN is very water soluble, so wet boreholes readily deactivated the explosive; and ANFO's low density of 0.85 g/cc limited its bulk explosive strength. Cook [18] hit upon the idea of dissolving the AN in a small amount of hot water, mixing in fuels such as aluminum powder, sulfur, or charcoal, and adding a thickening agent to gel the mixture and hold the slurried ingredients in place. As this mixture cooled down, the AN salt crystals would precipitate, but the gel would preserve the close contact between the oxidizer and the fuels, resulting in a detonable explosive. Other oxidizers also could be added, and the density could be adjusted with chemical foaming agents to vary the bulk explosive strength of the product. With the addition of a cross-linking agent, the slurry could be converted to a semisolid material, generally called a water gel, having some water resistance. The latest significant development in industrial explosives actually was invented only a few years after slurries [19, 20]. Water-in-oil emulsion explosives involve essentially the same ingredients that slurry composite explosives do, but in a different physical form. Emulsion explosives are discussed fully under the section titled "Explosives Manufacturing and Use."

The main developments in military types of explosives since World War II have been trends toward the use of plastic bonded explosives (PBXs) and the development of insensitive high explosives. Driving these trends are desires for increased safety and improved economics in the process of replacing aging TNT-based munitions and bomb fills. PBXs involve the coating of fine particles of molecular explosives such as RDX and HMX (1,3,5,7-tetranitro-1,3,5,7-tetrazacyclooctane) with polymeric binders and then pressing the resultant powder under vacuum to give a solid mass with the desired density. The final form or shape usually is obtained by machining. Explosives such as triaminotrinitrobenzene (TATB), nitroguanidine [21], and hexanitrostilbene (HNS) [22] are of interest because of their high levels of shock insensitivity and thermal stability. The development of new, explosive compounds and compositions is an ongoing area of research [23, 24], including interest in composite explosives similar to those used by industry. Examples are EAK, a eutectic mixture of ethylenediamine dinitrate, AN, and potassium nitrate [25], and nonaqueous hardened or cast emulsion-based mixtures [26].

Classification of Explosives

The original classification of explosives separated them into two very general types: low and high, referring to the relative speeds of their chemical reactions and the relative pressures produced by these reactions. This classification still is used, but is of limited utility because the only low explosives of any significance are black powder and smokeless powder. All other commercial and military explosives are high explosives.

High explosives are classified further according to their sensitivity level or ease of initiation. Actually sensitivity is more of a continuum than a series of discrete levels, but it is convenient to speak of primary, secondary, and tertiary high explosives. Primary explosives are the most sensitive, being readily initiated by heat, friction, impact, or spark. They are used only in very small quantities and usually in an initiator as part of an explosive train involving less sensitive materials, such as in a blasting cap. They are very dangerous materials to handle and must be manufactured with the utmost care, generally involving only remotely controlled operations. Mercury fulminate, used in Nobel's first blasting cap, is in this category, as is the more commonly used lead azide. On the other end of the spectrum are the tertiary explosives that are so insensitive that they generally are not considered explosive.

By far the largest grouping is secondary explosives, which includes all of the major military and industrial explosives. They are much less easily brought to detonation than primary explosives and are less hazardous to manufacture. Beyond that, however, generalizations are difficult because their sensitivity to initiation covers a very wide range. Generally, the military products tend to be more sensitive and the industrial products less sensitive, but all are potentially hazardous and should be handled and stored as prescribed by law. Table 37.3 lists some of the more prominent explosives of each type, along with a few of their properties.

For industrial applications, secondary explosives are subdivided according to their initiation sensitivity into two classes: Class 1.1 and Class 1.5. Class 1.1 explosives are sensitive to initiation by a blasting cap and usually are used in relatively small-diameter applications of 1–3-in. boreholes. Class 1.5 (historically known as blasting agents) are high explosives that are not initiated by a Standard #8 electric blasting cap under test conditions defined by the U.S. Department of Transportation (DOT) and that pass other defined tests designed to show that the explosive is "so insensitive that there is very little probability of accidental initiation to explosion or of the transition from deflagration to detonation" [27]. Being less sensitive, blasting agents are generally used in medium- and large-diameter boreholes and in bulk applications. Dynamites are always Class 1.1, but other composite explosives made from mixtures of oxidizers and fuels can be made either Class 1.1 or 1.5, depending upon the formulation and the density. Density plays a significant role in the performance of most explosives, and this is especially true for commercial, composite explosives such as slurries and emulsions where the density may be adjusted by air incorporation, foaming agents or physical bulking agents, irrespective of the formulation. The Class 1.5 explosives (blasting agents) are of interest because regulations governing transportation, use, and storage are less stringent than for Class 1.1 explosives. (Propellants and fireworks are classified by the DOT as Class 1.2 or 1.3 explosives, and blasting caps and detonating cord as Class 1.4.)

Structural Characteristics of Explosives

The number of potentially explosive compounds is virtually unlimited. A listing by the U.S. Bureau of Alcohol, Tobacco and Firearms of explosive materials under federal regulation [28] numbered 225, and many of the items listed were broad, general categories. The ten-volume *Encyclopedia of Explosives and Related Items* compiled by the U.S. Army

Table 37.3 Some properties of common explosives

Common name	Symbol	Composition	Molecular weight	Density (g/cc)	Detonation velocity (km/s)	Detonation pressure (kbar)	Explosive energy (cal/g)
Primary explosives							
Mercury fulminate		$Hg(CNO)_2$	284.7	3.6	4.7	220	428
Lead azide		$Pb(N_3)_2$	291.3	4.0	5.1	250	366
Silver azide		AgN_3	149.9	5.1	6.8	–	452
Lead styphnate		$C_6H(NO_2)_3O_2Pb$	468.3	2.5	4.8	150	368
Mannitol hexanitrate (nitromannite)	MHN	$C_6H_8(ONO_2)_6$	452.2	1.7	8.3	300	1,420
Diazodinitrophenol	DDNP	$C_6H_2N_4O_5$	210.1	1.5	6.6	160	820
Tetrazene		$C_2H_8N_{10}O$	188.2	1.5	–	–	658
Secondary explosives							
Nitroglycerin	NG	$C_3H_5(ONO_2)_3$	227.1	1.6	7.6	253	1,480
Pentaerythritol tetranitrate	PETN	$C(CH_2ONO_2)_4$	316.2	1.6	7.9	300	1,510
Trinitrotoluene	TNT	$CH_3C_6H_2(NO_2)_3$	227.0	1.6	6.9	190	900
Ethyleneglycol dinitrate	EGDN	$C_2H_4(ONO_2)_2$	152.1	1.5	7.4		1,430
Cyclotrimethylenetrinitramine (hexogen or cyclonite)	RDX	$C_3H_6N_3(NO_2)_3$	222.1	1.6	8.0	347	1,320
Cyclotetramethylenetetranitramine (octogen)	HMX	$C_4H_8N_4(NO_2)_4$	296.2	1.9	9.1	393	1,350
Trinitrophenylmethylnitraminc (tetryl)		$(NO_2)_3C_6H_2N(CH_3)NO_2$	287.2	1.4	7.6	251	950
Nitroguanidine	NQ	$CH_4N_3NO_2$	104.1	1.6	7.6	256	721
Nitromethane	NM	CH_3NO_2	61.0	1.1	6.2	125	1,188
Nitrocellulose	NC	Variable	–	1.4	6.4	210	950
Triaminotrinitrobenzene	TATB	$C_6H_6N_3(NO_2)_3$	258.2	1.8	7.9	315	829
Diaminotrinitrobenzene	DATB	$C_6H_5N_2(NO_2)_3$	243.2	1.6	7.5	259	993
Ethylenediamine dinitratc	EDDN	$C_2H_{10}N_4O_6$	186.1	1.5	6.8	–	948
Ethylenedinitramine (haleite)	EDNA	$C_2H_6N_2(NO_2)_2$	150.1	1.5	7.6	266	1,080
Picric acid		$C_6H_3O(NO_2)_3$	229.1	1.7	7.4	265	1,000
Ammonium picrate (explosive D)		$C_6H_6NO(NO_2)_3$	246.1	1.6	6.9	–	800
Picramide		$C_6H_6N(NO_2)_3$	228.1	1.7	7.3	–	1,070
Hexanitrostilbene	HNS	$[C_6H_3C(NO_2)_3]_2$	450.2	1.7	7.1	200	1,005
TACOT-Z		$C_{12}H_4N_8O_8$	388.2	1.6	7.2	181	980
Azobishexanitrobiphenyl	ABH	$C_{24}H_6N_{14}O_{24}$	874.4	1.8	7.6	–	
Dinitrotoluene	DNT	$CH_3C_6H_3(NO_2)_2$	182.1	1.5	5.0	–	700
Composition B		49/50/1 TNT/PDX/wax	–	1.7	8.0	294	1,100
Pentolite		50/50 TNT/PETN	–	1.6	7.7	245	1,100
Amatol		50/50 TNT/AN	–	1.6	6.5	–	950
Dynamite		Variable NG and various oxidizers and fuels	–	0.8–1.6	1.8–7.6	30–160	675–1,090
Prilled AN-fuel oil	ANFO	94/6 AN/FO	–	0.8–0.9	1.5–4.0 (depends on diameter)	–	880
Slurries or water gels		Variable mixtures of oxidizers, fuels, and water	–	0.9–1.4	3.5–5.0	–	600–1,200
Emulsions		Variable solutions of oxidizers in water and fuels	–	0.9–1.4	4.5–6.0	–	700–1,100
Heavy ANFO		50–75 % AN with 50–25 % emulsion	–	1.1–1.3	4.0–4.5	4.0–4.5	755–815
Tertiary explosives							
Mononitrotoluene	MNT	$CH_3C_6H_4NO_2$	137.1	1.2			
Ammonium perchlorate	AP	NH_4ClO_4	117.5	1.9	3.4	187	488
Ammonium nitrate	AN	NH_4NO_3	80.1	1.4	3.2	–	346

Picatinny Arsenal over a 25-year period contains several thousand entries. New organic molecular explosives are still being synthesized; composite explosives, such as current commercial products that are mixtures of oxidizers and fuels, present an infinite number of possible combinations. The complexity of trying to comprehensively list the chemical structures of explosives is shown by a 1977 reference that listed 13 separate categories just for primary explosives [29]. However; the majority of the most important explosives can be grouped into a few classes sharing common structural features that are of value to researchers in understanding and predicting explosive properties.

The following seven categories [30, 31], updated to include the relatively recent fluoroderivatives [23], appear to be the most encompassing. Many explosives may contain more than one category, but not every compound that contains one of these chemical groups is necessarily an explosive.

–NO
–N–N–, –NN–, and –N≡N
–C≡N– and –C≡N
–C≡C–
–ClO
–N–X, where X = Cl, F, I
–O–O–

Category 1 is by far the largest. It includes nitro groups, both aliphatic and aromatic; nitrate esters; nitrate salts; nitramines; and nitrosamines. Nearly all of the explosives listed in Table 37.3 fall into this category. Prominent examples are: nitromethane, an aliphatic nitro compound; TNT, an aromatic nitro compound; NG and PETN, nitrate esters; EDDN and ammonium nitrate, nitrate salts; and RDX and HMX, nitramines. Category 2 represents the hydrazine, azo, diazo, and azide compounds, both organic and inorganic. Hydrazine, tetrazene, and lead azide are examples of this group. Category 3 is represented by the explosives mercury fulminate and cyanogen, respectively. Acetylene and metallic acetylide salts constitute category 4. Category 5 consists mainly of inorganic and organic ammonium salts of chloric and perchloric acid, but would also include various chlorine oxides. Category 6 is generalized to include most of the amine halogens, nitrogen triiodide being a classic example. Also, considerable synthetic work has focused on inserting the energetic difluoroamine groups into various organic molecules to form explosives that fall into this category. Category 7 includes organic peroxides and ozonides as well as hydrogen peroxide itself.

Commercial industrial explosives such as dynamites, slurries, and emulsions are included in these categories because their major components, nitrate esters and nitrate and perchlorate salts, are listed. However, mixtures of fuels and oxygen or other gases that may be explosive at certain ratios are not covered, including the liquid oxygen explosives that saw limited application earlier in the twentieth century.

Explosives Manufacturing and Use

Details of the synthesis and larger-scale production of a number of molecular explosives including dynamites are given in the four-volume series by Urbanski (*Chemistry and Technology of Explosives*, Pergamon Press, 1964–1984) and in various military books such as *Engineering Design Handbook: Explosive Series*. Formulations of commercial slurries and emulsions generally are considered proprietary and are described mainly in the patent literature. Some specific examples of prominent explosives with general preparation methods are given below.

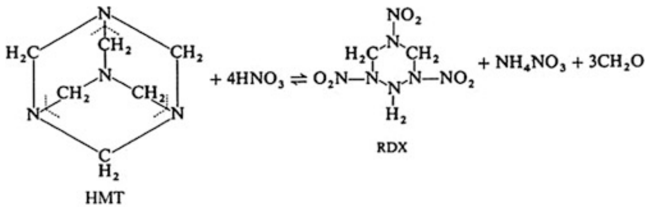
TNT

TNT is no longer manufactured in the United States for either commercial or military use. It is produced commercially in other countries and is imported into the United States for use in cast boosters to initiate industrial blasting agents. TNT use in military applications is still significant, but starting to decline as other higher technology munition and bomb fills come into use. Current TNT military needs are supplied from off-shore sources as well as through recycling from demilitarized weapons. A lot of the demilitarized TNT is also used commercially in the U.S. in cast boosters. In a relatively straightforward process, TNT is made by the direct tri-nitration of toluene with nitric acid. Most modern processes are set up for continuous production in a series of nitrators and separators with the nitrating acid flowing counter-currently. This procedure avoids having to isolate the intermediate mono- and dinitration products and may also employ continuous purification and crystallization, being carried out simultaneously with production.

Mixed nitric and sulfuric acids sometimes are used with the addition of SO_3 or oleum. The sulfuric acid or oleum helps drive the reaction to completion by removing the water produced by nitration and by dehydrating nitric acid to form the more reactive nitronium ion (NO_2^+). Because toluene is not very soluble in the acid, powerful agitation is required. The spent acid is removed in successive separation steps, and the sulfuric acid is reused after the addition of more nitric acid. The molten TNT product is purified with multiple water and sodium sulfate washes, which produce significant quantities of “yellow water” and “red water” waste streams, respectively, that must be properly handled to avoid environmental problems. The low melting point of TNT (80–82°C) is ideal for melt casting, and TNT usually is employed as a mixture with other higher-melting explosives such as PETN, RDX, HMX, and Tetryl. This feature and the excellent chemical stability of TNT have made it, historically, the most popular and widely used military explosive in the world.

RDX and HMX

Both RDX and HMX are cyclic nitramines made by nitrolysis of hexamethylene tetramine (HMT). Their good thermal stabilities, high melting points (7200°C), and high energy properties make these crystalline compounds popular as projectile and bomb fills and for use in cast boosters and flexible, sheet explosives. HMX has superior detonation properties and a higher melting point than RDX, but it is more difficult and more expensive to manufacture. Reaction 1 shows the formation of RDX by the action of nitric acid on HMT. Schematically, RDX formation can be pictured as nitration of the three "outside" nitrogen atoms of HMT (in more accurate, three-dimensional representations all four nitrogens are equivalent) with removal of the "inside" nitrogen and methylene (–CH_2–) groups. AN (NH_4NO_3) and formaldehyde (CH_2O) are produced as by-products, but can be used to form more RDX with the addition of acetic anhydride, as shown in Reaction 2. In actual practice these two reactions are run simultaneously, as shown in the combined reaction to produce approximately 2 mol of RDX for each mole of HMT.



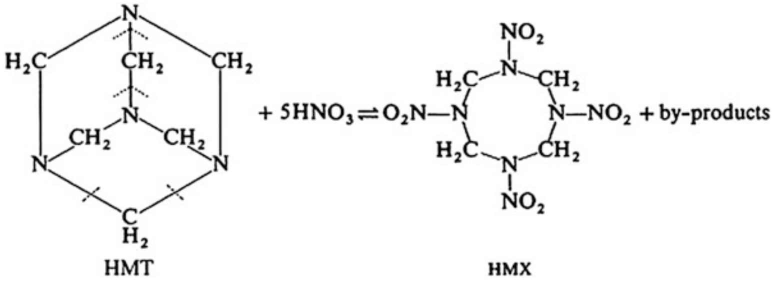
Reaction 2

$$\underset{\text{Formaldehyde}}{3CH_2O} + \underset{\text{AN}}{3NH_4NO_3} + \underset{\text{Acetic anhydride}}{6(CH_3CO)_2O}$$

$$\rightarrow \underset{\text{RDX}}{(CH_2 \cdot N \cdot NO_2)_3} + \underset{\text{Acetic acid}}{12CH_3COOH}$$

Combined Reaction

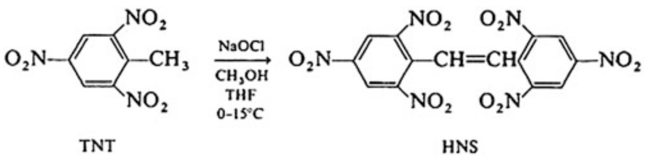
$$(CH_2)_6N_4 + 4HNO_3 + 2NH_4NO_3 + 6(CH_3CO)_2 \rightarrow 2(CH_2 \cdot N \cdot NO_2)_3 + 12CH_3COOH$$



HMX was discovered as an impurity produced in the RDX reaction. It is composed of an eight-membered ring rather than the six-membered ring of RDX. The latter is more readily formed than the eight-membered ring, but with adjustment of reaction conditions (lower temperature and different ingredient ratios), HMX formation can be favored. Schematically, its formation can be pictured by nitration of all four nitrogens in HMT and removal of two methylene groups as indicated in Reaction 3. To obtain pure HMX, the RDX "impurity" must be removed by alkaline hydrolysis or by differential solubility in acetone.

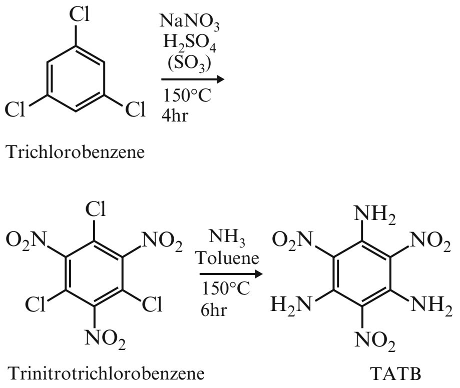
HNS (2,2,4,4,6,6-Hexanitrostilbene)

This explosive was prepared unequivocally for the first time in the early 1960s [32, 33]. It is of interest primarily for two reasons: (1) its high melting point (316°C) and excellent thermal stability, and (2) its unique crystal-habit-modifying effects on cast TNT. The former makes HNS useful in certain military and space applications as well as in hot, very deep wells, and the second property is used to improve TNT castings. It can be manufactured continuously by oxidative coupling of TNT as shown below. This relatively simple process from readily available TNT and household bleach (5% NaOCl solution) has been shown to involve a series of intermediate steps that give HNS in only low to moderate yields (30–45%) with many by-products. Although it also involves the use of expensive organic solvents that must be recovered, this synthesis is used commercially [34, 35].



TATB (1,3,5-Triamino-2,4,6-Trinitrobenzene)

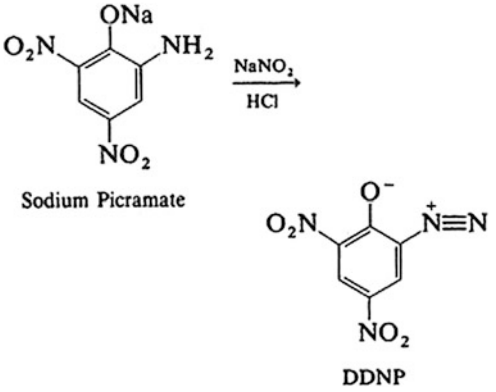
This highly symmetrical explosive molecule has even higher thermal stability than HNS (greater than 400°C) and has ongoing interest because of its extreme insensitivity [36–38]. Because accidental initiation is highly unlikely, TATB has been used in nuclear warheads and has been explored for use in plastic bonded systems for a number of military and space applications [21]. It is manufactured in large-scale batch processes that are little changed from its original synthesis over 100 years ago. The two-step process involves tri-nitration of trichlorobenzene followed by amination to displace the chlorine groups as shown below.



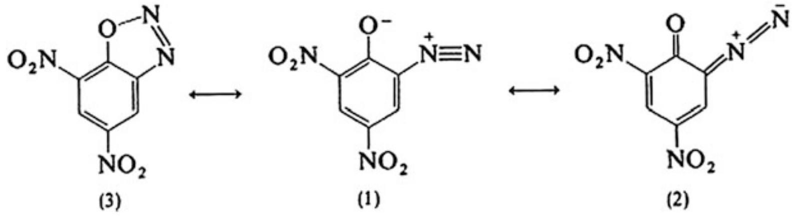
Both steps require high temperature and considerable reaction time but give 80–90% yields. The major problem areas are chloride impurities in the final product and the excessively fine particle size of the final product. Because TATB is highly insoluble in most solvents, it is difficult to purify the product or to change its particle size by recrystallization. Also the starting material is expensive and not very readily available. More recently, a similar synthetic procedure starting with 3,5-dichloranisole was reported [39].

DDNP

This yellow-to-brown crystalline material (melting point 188°C) is a primary explosive used as the initiator charge in electric blasting caps as an alternative to lead azide. It is less stable than lead azide but much more stable than lead styphnate and is a stronger explosive than either of them because it does not contain any metal atoms. 2-Diazo-4, 6-dinitrophenol (DDNP) is also characterized as not being subject to dead pressing (tested at pressures as high as 130,000 psi). It was the first diazo compound discovered (1858) and was commercially prepared in 1928. It is manufactured in a single-step, batch process by diazotizing a slurry of sodium picramate in water.

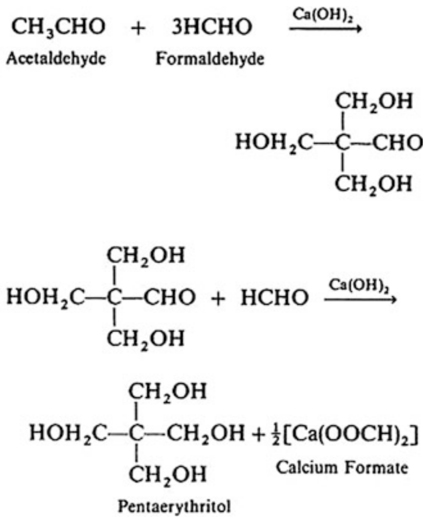


The structure shown in this reaction is convenient for visualization purposes, but DDNP actually exists in several tautomeric forms as shown below with form (2) apparently predominating. The sodium picramate starting material is itself explosive, but is commercially available as a chemical intermediate. It can be made by the reduction of picric acid with reducing agents such as sodium sulfide. The key to making useful DDNP is to control the rate of diazotization so that relatively large, rounded crystals are formed instead of needles or platelets that do not flow or pack well.



PETN

Although known as an explosive since 1894, PETN was used very little until after World War I when the ingredients to make the starting material became commercially available. The symmetrical, solid alcohol starting material, pentaerythritol, is made from acetaldehyde and formaldehyde, which react by aldol condensation under basic catalysis followed by a crossed Cannizzaro disproportionation to produce the alcohol and formate salt. Although the reaction takes place in a single mixture, it is shown below in two steps for clarity.



For PETN manufacture the pentaerythritol starting material can be readily purchased as a commodity chemical from commercial suppliers. The nitration is relatively simple, involving only nitric acid (96–98%) and the solid alcohol added slowly with mixing and cooling. PETN is not very soluble in nitric acid or water and is readily filtered directly from the acid or after dilution of the acid with water. Water washing and recrystallization from acetone–water mixtures give the desired particle size ranges and the desired purity. PETN can be made either batchwise or continuously for large-scale production.

Pure PETN is a white, crystalline solid with a melting point of 141.3°C. Because of its symmetry, it is said to have higher chemical stability than all other nitrate esters [40]. Relatively insensitive to friction or spark initiation, PETN is easily initiated by an explosive shock and has been described overall as one of the most sensitive, non-initiating, military explosives [41]. As with most explosives, the detonation velocity of PETN varies with the bulk density of the explosive. Most military applications of PETN have been converted to RDX because of its greater thermal stability. However, in industry PETN is widely used as a major component in cast boosters for initiating blasting agents, as the explosive core in detonating cord, and as the base load in detonators and blasting caps. For safety in handling, PETN is shipped in cloth bags immersed in water–alcohol mixtures and dried just before use.

NG (Nitroglycerin or Glyercol Trinitrate)

This nitrate ester is one of only a very few liquid molecular explosives that are manufactured commercially. It is a clear, oily liquid that freezes when pure at 13°C. As seen in the historical section, the first practical use of NG was in dynamites, where it is still used today more than 140 years later. It also is used as a component in multi-based propellants and as a medicine to treat certain coronary ailments. This latter usage is attributed to NG's ability to be rapidly absorbed by skin contact or inhalation into the blood, where it acts as a vasodilator. (At high exposure levels such as in dynamite manufacture and handling, this property is responsible for the infamous powder headache.) NG is undoubtedly the most sensitive explosive manufactured in relatively large quantities. Its sensitivity to initiation by shock, friction, and impact is very close to that of primary explosives, and extreme safety precautions are taken during manufacture. Pure glycerin is nitrated in very concentrated nitric and sulfuric acid mixtures (typically a 40/60 ratio), separated from excess acid, and washed with water, sodium carbonate solution, and water again until free from traces of acid or base. Pure NG is stable below 50°C, but storage is not recommended. It is transported over short distances only as an emulsion in water or dissolved in an organic solvent such as acetone. Traditionally, it has been made in large batch processes, but safety improvements have led to the use of several types of continuous nitrators that minimize the reaction times and quantities of explosives involved. Because of its sensitivity, NG is utilized only when desensitized with other liquids, combined with absorbent solids or compounded with nitrocellulose.

Packaged Explosives

Packaged explosives dominated the explosives market from the time dynamite was invented in 1867 until the middle of the twentieth century. At that time, other composite type packaged technologies began appearing, particularly in the 1960s and 1970s, concurrent with increasing market penetration of bulk explosives. During those years of rapid technology development, two packaged product types emerged, first water-gel explosives, followed in the 1970s and early 1980s by packaged emulsion explosives. All of these packaged product technologies remain active in parts of the world as of the date of this edition, but in the U.S., emulsions are pre-eminent followed by dynamites. Both are discussed below.

Dynamite

Dynamite is not a single molecular compound, but a mixture of explosive and non-explosive materials formulated in cylindrical paper or cardboard cartridges for a number of different blasting applications. Originally Nobel simply absorbed NG into kieselguhr, an inert diatomaceous material, but later he replaced that with active ingredients—mixtures of finely divided fuels (including absorbent combustibles) and oxidizers called dopes. Thus, energy is derived not only from the NG, but also from the reaction of oxidizers such as sodium nitrate with the combustibles.

The manufacture of dynamite involves mixing carefully weighed proportions of NG and various dopes to the desired consistency and then loading preformed paper shells through automatic equipment. Because dynamites represent the most sensitive commercial products produced today, stringent safety precautions such as the use of non-sparking and very-little-metal equipment, good housekeeping practices, limited personnel exposure, and barricaded separations between processing stations are necessary for manufacturing. Today, the "NG" used in dynamite is actually a mixture of EGDN (ethylene glycol dinitrate) and NG formed by nitrating mixtures of the two alcohols (ethylene glycol and glycerin), in which NG is usually the minor component. Table 37.4 lists the common general types of dynamites with their distinguishing features. The straight dynamites and gelatins largely have been replaced by the ammonia dynamites and ammonia gelatins for better economy and safety characteristics.

The use of NG-based dynamite continues to decline throughout the world. For example, by 1995 there was only one dynamite manufacturing plant left in North America, and in 2010 the dynamite production at this plant had dropped to about one-third the amount produced in 1990. The reasons for the declining use of dynamite are its unpopular properties of sensitivity to accidental initiation and the headache-causing fumes. However (despite its higher cost), for some difficult blasting applications it remains the product of choice due to its high density and high energy and sensitivity.

Packaged Emulsions

The first major non-dynamite package product technology was water gels. These were composite mixtures of hot, concentrated, dissolved oxidizer solution slurried with other ingredients including soluble and insoluble liquid fuels, particulate fuels, solid oxidizer salts, and thickeners consisting of natural and synthetic water soluble polymers. The thickened solution could be gelled by addition of cross-linking agents, typically inorganic salts of certain metals. The gelled structure gave significant, but not excellent water resistance.

Packaged water-in-oil emulsions are made from essentially the same oxidizer solutions as used for water-gel products, but are fueled with an organic external liquid hydrocarbon phase that provides much better water resistance. These products are basically formed with the same manufacturing equipment as the bulk emulsions (see next section). The fuel component usually contains waxes and other thickeners to give the emulsions a thick, putty-like consistency, and the oxidizer solution often contains both AN and a second oxidizer salt to produce optimum energy, physical stability, and improved after-blast fumes. After manufacture, the hot, thick emulsion is extruded into packaging material, normally a plastic film. The final product is then clipped together with metal clips forming firm, sausage-like chubs and is cooled which further thickens the product. Some packaged emulsions are also available in paper cartridges, designed to simulate dynamite packaging. To obtain reliable detonability in small diameters, the density of packaged emulsions must be maintained at a relatively low value, typically 1.10–1.20 g/cc. On the other hand, some dynamites are available with densities in excess of 1.40 g/cc. Figure 37.4 shows some commercial packaged emulsion cartridges in both plastic and paper wrappings.

Table 37.4 General types of dynamite

Straight dynamite	Granular texture with NG the major source of energy
Ammonia dynamite ("extra" dynamite)	AN replacing part of the NG and sodium nitrate of the straight dynamite
Straight gelatin dynamite	Small amount of nitrocellulose added to produce soft to tough rubbery gel
Ammonia gelatin dynamite ("extra" gelatin)	AN replacing part of the NG and sodium nitrate of the straight gelatin
Semigelatin dynamite	Combination of types 2 and 4 within-between properties
Permissible dynamite	Ammonia dynamite or gelatin with added flame retardant

Fig. 37.4 Commercial packaged emulsion cartridges (Courtesy Dyno Nobel)

Bulk Explosives

Ammonium Nitrate and ANFO

Ammonium nitrate (AN) continues to be the most widely used component of commercial explosives. It is used in nearly all of the packaged and bulk explosives on the market. Ammonia is the main raw material needed to manufacture AN. Some of the AN manufacturers make their own ammonia and some purchase it on the open market. It is obvious that the cost of manufacturing AN will depend on the price of ammonia and, even more basically, natural gas from which ammonia is made. Figure 37.5 shows the volatility of natural gas over the last decade. This volatility translates into volatility in the cost of explosives, although significant natural gas development in the U.S. over the past few years seems to be suppressing some of the price fluctuation.

There are many producers of AN in North America making both AN solution and explosive-grade AN prills. The AN solution is used in the manufacture of packaged and bulk emulsion and water-gel explosives, and explosive-grade AN prills are used to make ANFO, a composite explosive described in an earlier section. These low density AN prills are made by a specialized process, in which internal

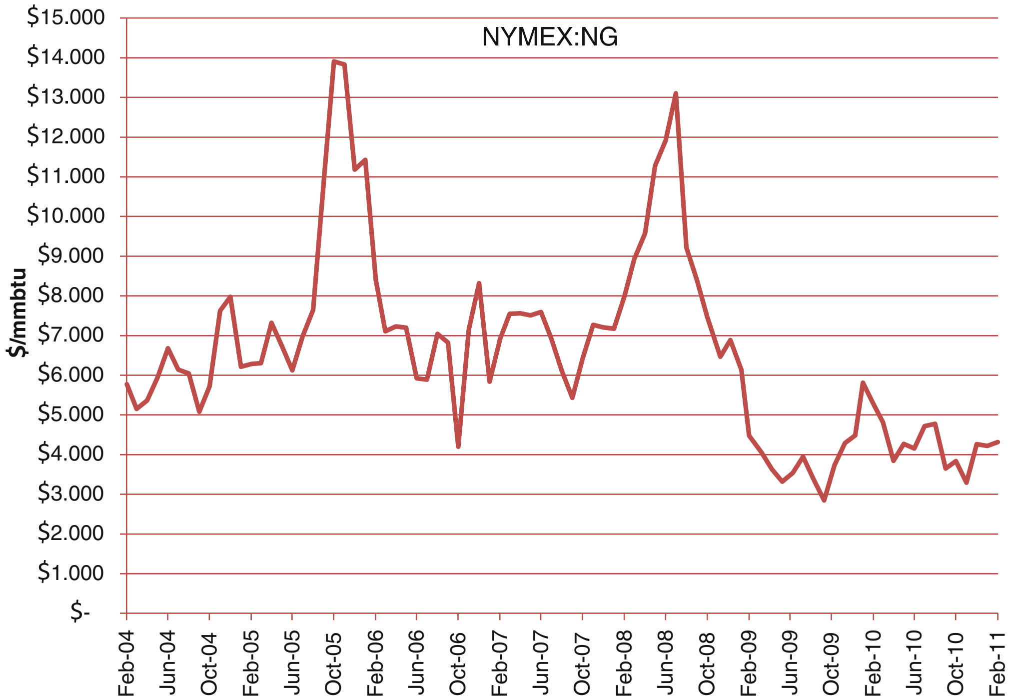


Fig. 37.5 Volatility of natural gas prices from 2004 to 2010. New York Mercantile Exchange, Natural Gas, Compiled Monthly

voids are created, making the prills porous and able to absorb the required 5.5–6% FO. ANFO remains the largest volume commercial explosive in use today around the world. Because of this, ANFO is commonly used as a reference when defining and comparing explosive properties. Some of these important explosive properties include density, detonation velocity, and energy release.

The crystal density of AN is about 1.72 g/cc, and the particle density of explosive-grade AN prills ranges from about 1.40 to 1.45 g/cc depending upon the manufacturing process. This difference in crystal and particle density reveals the volume of pores or voids created by the specialized prilling process. The porosity of AN prills is the property desired in the manufacture of ANFO, since this determines how much FO can be absorbed. This intimate mixture of AN with FO is critical to the efficient detonability of ANFO. The AN prill particle density and inherent void-space value also become important when predicting and calculating the densities of ANFO blends with water-gel and emulsion explosives.

The bulk density of ANFO ranges from about 0.80 to 0.87 g/cc. So, clearly about half of the ANFO is air or void space. Most commercial explosives require a certain amount of entrained void space in order to detonate optimally. These void spaces play a major role in the detonation reaction by creating reaction sites, "hot spots," under adiabatic compression and interaction with the shock wave in the detonation front [42]. The amount of void space in any given explosive and the resultant change in density have a significant impact on the detonation properties like detonation velocity, sensitivity, and even energy release.

Generally speaking, the detonation velocity of a composite explosive will increase with density until a failure point is reached. This failure point is commonly referred to as the critical density of that particular explosive in that particular diameter or other configuration. The density at that point is so high and the void space so low that the detonation cannot be sustained and failure occurs.

Other important parameters that affect the detonation velocity and performance of ANFO are charge diameter and confinement. The detonation velocity of commercial composite explosives such as ANFO will drop dramatically as the diameter of the charge nears the failure diameter. At a given density, the diameter below which a charge fails to propagate a detonation is called the critical diameter. As a charge diameter is increased from this point, the detonation velocity increases asymptotically to a maximum value for that particular explosive formulation. The detonation velocity of ANFO will also increase noticeably when the charge confinement is changed, such as from a PVC tube to a Schedule 40 steel pipe. A summary of test data on ANFO velocity vs. confinement and diameter is shown in Fig. 37.6. The temperature of composite explosives also affects sensitivity and critical properties.

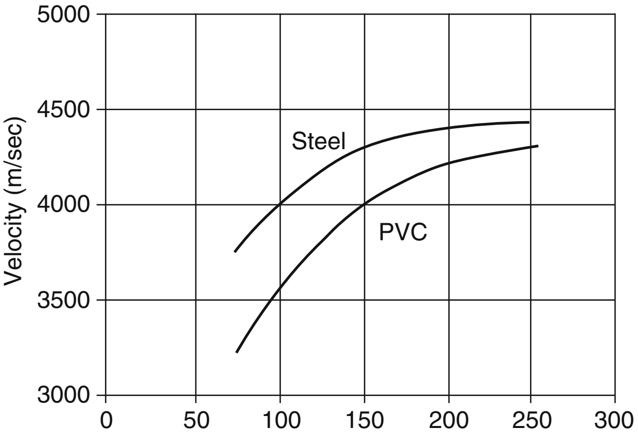


Fig. 37.6 Detonation velocity of ANFO vs. diameter and confinement

The basic chemical reaction of ANFO can be described with the following equation:

$$3NH_4NO_3 + CH_2 \rightarrow CO_2 + 7H_2O + 3N_2 + 880\,\mathrm{cal/g}$$

Using CH_2 to represent FO is generally accepted, but it really is an over-simplification, since it is a mixture of hydrocarbons. The heat energy release of 880 cal/g is the theoretical maximum value based upon the heats of formation of the reactants and products. Of course, all of the products of detonation are gases at the detonation temperature of about 2,700 °K.

The theoretical work energy that is released from an explosive reaction can be calculated using a variety of equations of state and computer programs [43]. Explosive energy can also be measured by a variety of techniques including underwater detonation of limited size charges with concurrent measurements of the shock and bubble energies [44]. Each explosive manufacturer has an energy measurement and equation of state that is used to calculate and report their product properties. This often leads to confusion and controversy when explosive consumers try to compare product lines when given only technical information sheets. Since theoretical calculations must of necessity be based on a number of assumptions, the most valid comparisons are done in the field with product testing and detailed evaluation of results.

Bulk Emulsions

During the past 50 years, the commercial use of bulk delivered explosives, including ANFO and other composite explosives, has continued to increase while the use of packaged explosives continued to fall on a percentage basis. Due to low cost and superior sensitivity compared to earlier water-resistant technologies such as water gels, bulk emulsions have been leading this trend over the last 25–30 years. Bulk products initially became very popular in large volume open-pit mining operations and this accelerated with the advent of bulk emulsion products and blends with ANFO. More recently, bulk emulsions have increasingly replaced packaged products in underground mining and particularly in quarry operations.

Emulsions are made by combining an oxidizer solution and a fuel solution using a high-shear mixing process. The oxidizer solution is normally 90–95% by weight of the emulsion. It contains AN, water, and sometimes a second oxidizer salt such as sodium or calcium nitrate. The solution must be kept quite hot, since the water is minimized for increased energy, and the crystallization temperature is typically 50–70°C. The fuel solution contains liquid organic fuels, such as FO and/or mineral oils, and one or more emulsifiers. An emulsifier is a surface active chemical that has both polar and nonpolar ends of the molecule. In the high-shear manufacturing process, the oxidizer solution is broken up into small droplets, each of which is coated with a layer of fuel solution. The droplets in this meta-stable, water-in-oil emulsion are basically held together with the emulsifier molecules, which migrate to the surface of the dispersed droplets and form a link between the hydrophobic external (continuous) oil phase and the hydrophilic phase within the droplets. In today's explosives industry, much of the research work is directed towards developing better and more-efficient, emulsifier molecules that will improve the storage life and handling characteristics of the bulk and packaged emulsions. The emulsifiers currently used in commercial explosives range from relatively simple fatty acid esters with molecular weights of 300–400 to the more complex polymeric emulsifiers having molecular weights in excess of 2,000.

Figure 37.7 shows a photomicrograph of an emulsion explosive at 400 power with the typical distribution of the fuel-coated oxidizer solution droplets (nominally 1–5 μm in diameter). Figure 37.8 shows a bulk emulsion exiting a loading hose and displaying the soft ice cream-like texture typical of bulk emulsions. The viscosities of bulk emulsions can range from nearly as thin as 90 weight oil to thicker than mayonnaise, depending upon the application requirements. Emulsion viscosity increases with product cooling, but most emulsions continue to remain stable at temperatures well below 0°C, dramatically below the crystallization temperature of the oxidizer solution. The oxidizer solution droplets in the emulsion are therefore held in a supersaturated state. Over time, the surface layers created by the emulsifier molecules can become less stable due to chemical degradation (oxidation) of the emulsifier or other spontaneous physical and chemical processes. Movement of the emulsion is also stressful to these systems, which in reality are thermodynamically unstable, tending towards crystallization and coalescence. Crystals can begin to form that break though the emulsifier layer and the emulsion begins to "break down" and further crystallize, thereby diminishing some of the desirable physical and detonation properties. For this reason emulsion compositions must be optimized for a particular application in terms of product stability and usable storage life.

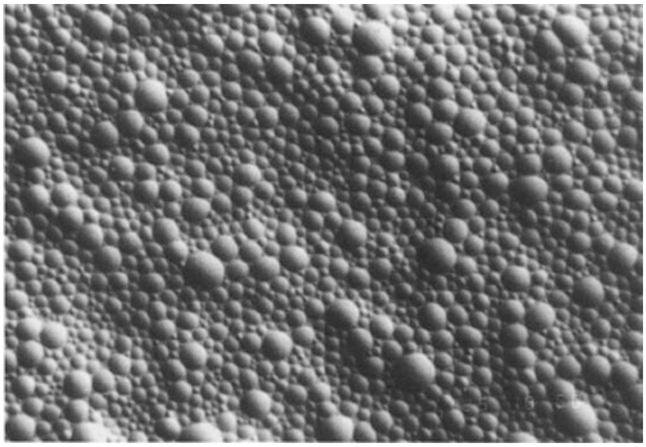


Fig. 37.7 Photomicrograph of a bulk emulsion

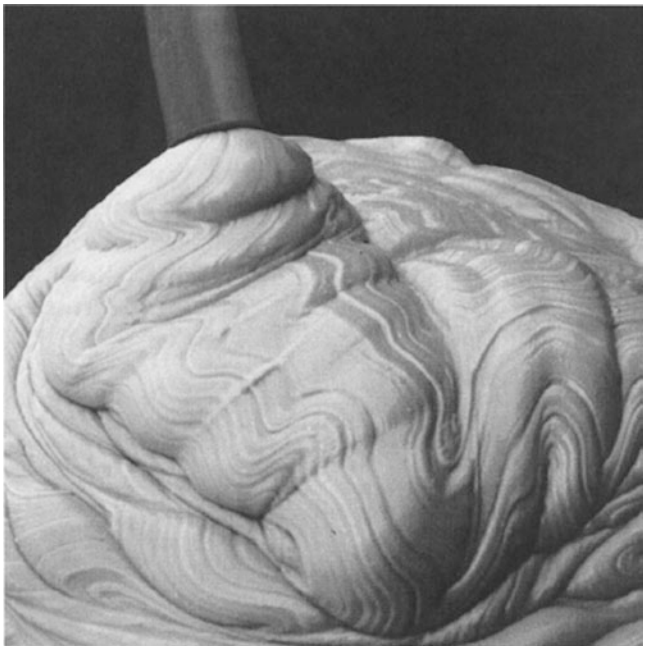


Fig. 37.8 Bulk emulsion exiting a loading hose

The intimate mixing of oxidizer and fuel in emulsions gives these explosives much higher detonation velocities when compared to ANFO. For example, in 150 mm diameter PVC, ANFO at a density of ~0.82 g/cc has a velocity of about 4,000 m/s, and a sensitized emulsion would have a velocity closer to 6,000 m/s at a density of 1.20–1.25 g/cc.

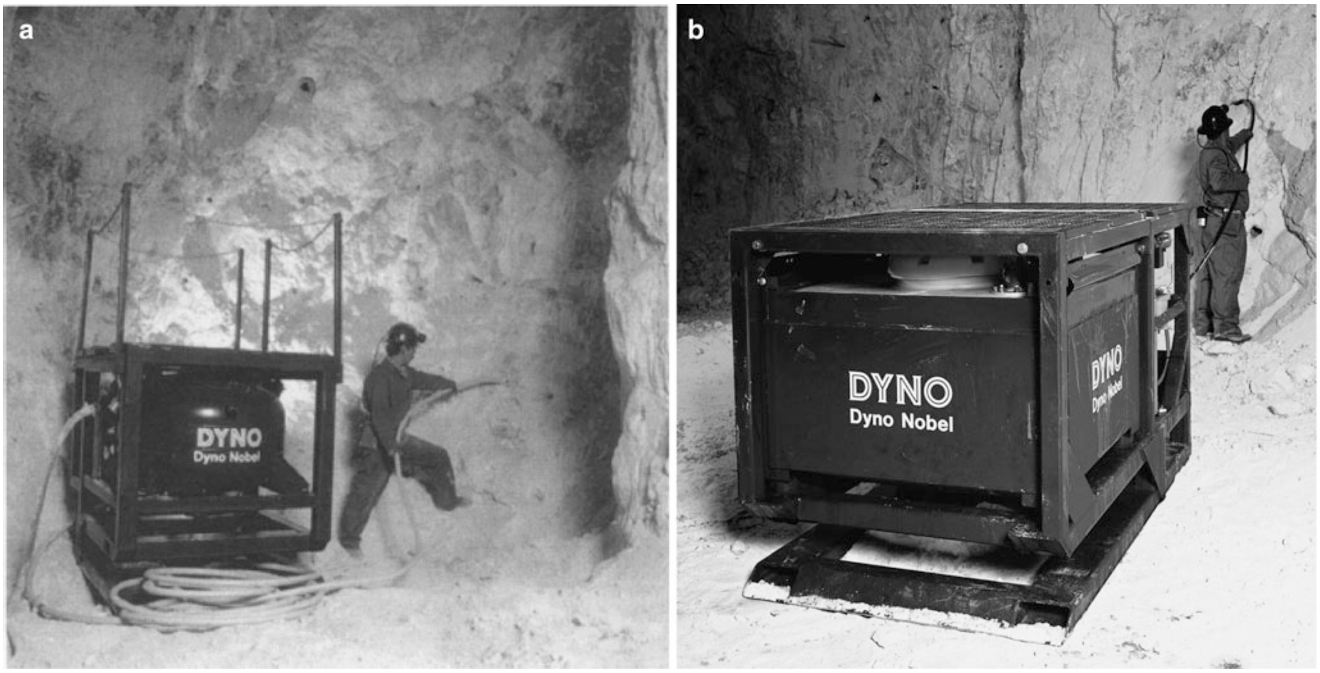


Fig. 37.9 (**a**) Underground pressure vessel loader. (**b**) Compressed-air, powered pumper for underground bulk emulsion (Courtesy Dyno Nobel)

Also, the layer of oil surrounding each oxidizer solution droplet protects the emulsion from extraneous water intrusion and subsequent deterioration of the explosive. Many studies have shown that when mining operations use emulsion explosives rather than ANFO, which has basically no water resistance, the amount of nitrate salts in mine ground water is reduced considerably. This can be a very important factor in today's environmentally conscious mining and explosives industry.

Bulk emulsions are generally non-detonable per se and are typically sensitized with some type of density control medium to become usable blasting agents. These voids interact with the shock wave, creating "hot spots," which are required to sustain the detonation front. The two most commonly used density control methods are hollow solid microspheres and gas bubbles created by an in situ chemical reaction. Both glass and plastic hollow microspheres are commercially available and used by explosives manufacturers. The in situ chemical gassing techniques require considerably more expertise and generally utilize proprietary technology. Emulsion blended with ANFO, with its internal and interstitial air voids, is also widely used as a less efficient but effective sensitizer in larger diameter boreholes.

Underground Bulk Emulsions

In the past decade the use of sensitized bulk emulsions has continued to increase in underground mining. In addition to the necessary small diameter sensitivity of emulsion based explosives in these applications, much of this trend has been due to the development of innovative loading equipment and techniques. One example of this is shown in Fig. 37.9a, which shows a small-volume pressure vessel that can be used for development and tunnel rounds utilizing horizontally drilled boreholes. The bulk emulsion blasting agent is pressurized inside the vessel and literally squeezed through the loading hose into the boreholes. A continuous column of explosives is assured by inserting the loading hose to the back of the hole and extracting it as the product is loaded. A more recently designed air-powered pumper unit for underground delivery of bulk emulsion is shown in Fig. 37.9b. This unit typically uses compressed air, available in most underground mines, to power a pump that delivers the bulk explosive. Much more complex underground loading units are available for loading bulk emulsion into boreholes drilled at any angle to the horizontal from straight up to straight down.

The emulsion explosives used in these specialized loading units were specifically designed for underground use nearly 30 years ago and remain essentially the same today. They are manufactured in central plants, either sensitized with microballoons to form as a 1.5 explosive or as a gassable matrix that can be transported and stored as a non-explosive, thence sensitized with gas bubbles at the point of use. They have been successfully used in underground mining operations around the world. The fuel and oxidizer contents are carefully balanced, and this, combined with the excellent water resistance and detonation efficiency,

Fig. 37.10 Bulk emulsion loading truck in a Florida quarry (Courtesy Dyno Nobel)

results in the near elimination of after-blast toxic fumes, such as CO, NO, and NO_2. The fume characteristics of this product have been shown to be considerably superior to either dynamite or ANFO. For example, a series of tests in an underground chamber in Sweden compared the after-blast fumes of this emulsion to ANFO. The CO was reduced from 11 to less than 6 L/kg of explosive, and the NO plus NO_2 was reduced from about 7 to less than 1 L/kg of explosive [45].

Surface Bulk Emulsions

Many open-pit quarries also use bulk emulsions for their blasting operations. As the size of the quarry increases, the size of the explosive loading trucks also must increase. Truck payloads can range from 5,000 to 30,000 lb of product. Figure 37.10 shows an emulsion pumper truck in a quarry in south Florida. These particular trucks, with a payload of about 20,000 lb, are specially designed for a site-mixed system, in which each truck is an emulsion manufacturing unit. Combining non-explosive raw materials directly on the truck maximizes safety and minimizes requirement for explosive storage. This particular bulk emulsion is manufactured at a rate between 300 and 500 lb/min and sensitized to the desired density with a chemical gassing system as it is loaded into the boreholes.

Figure 37.11 shows a Florida blast in progress. Note the ejection of cardboard tubes from some holes. These tubes must be used in most areas to keep the boreholes from collapsing in the layered, coral limestone formation.

Other surface, bulk delivery trucks use emulsions that are manufactured in central plants, then shipped to a surface depot, off-loaded into a bin or overhead silo, and loaded as needed into delivery trucks by pumping or gravity flow. The products can be shipped pre-sensitized with microballoons or as a non-sensitized, gassable matrix. The application of the product can be as a pumped straight emulsion or as a pumped or augured emulsion/ANFO blend.

Many of the large volume metal and coal mining operations around the world have both bulk emulsion and AN prills stored either on site or nearby so that any combination of these two products can be used. Figure 37.12 shows a typical explosive staging area in a large open-pit coal mine in Wyoming. The explosive truck in the foreground has compartments on board for emulsion, AN prills, and FO, so any combination of products ranging from straight emulsion to straight ANFO can be loaded. The truck has a capacity of about 50,000 lb and can deliver product to the boreholes at up to a ton per minute. Each borehole can contain as much as 5 tons of explosive, and some of the blast patterns can contain as much as ten million total pounds.

The emulsion/ANFO explosive blend selection to be used in any given mining application depends upon many factors. Typically, ANFO is the least expensive product, but it also has the lowest density and no water resistance. As emulsion is added to ANFO, it begins to coat the AN prills and fill the interstitial voids between the solid particles. This increases the density, detonation velocity, and water resistance compared to ANFO. The density increases nearly linearly with percent emulsion from about 0.85 g/cc with ANFO to about

Fig. 37.11 Florida quarry blast in progress (Courtesy Dyno Nobel)

1.32 g/cc with a 50/50 blend. This range of emulsion/ANFO blends is commonly referred to as Heavy ANFO. As the density increases, the amount of explosive that can be loaded into each borehole increases, and either drill patterns based on ANFO can be spread out or better blasting results can be obtained.

It is commonly accepted in the explosive and mining industry that at least 45–50% emulsion is required to protect the Heavy ANFO blend from borehole water intrusion. Pumped explosive blends with 60–80% emulsion can be used for even better water resistance when severe water conditions are encountered. These products can be pumped through a loading hose, which can be lowered to the bottom of the borehole and displace the water during loading. Trucks similar to that shown in Fig. 37.10 can be used for these products. Most Heavy ANFO products are more simply mixed through an auger and discharged into the top of boreholes by trucks similar to the one shown in Fig. 37.12. For Heavy ANFO products, especially those with less than ~45% emulsion, the holes should be either dry or dewatered using pumps.

The basic chemical composition of a typical all-AN oxidizer emulsion explosive would be: AN plus about 16% water plus about 5% fuels. The fuels may contain fuel oil, mineral oil, and emulsifiers; the majority of which can generally be described as CH_2 hydrocarbon chains. Therefore, a very simplified chemical reaction for a basic emulsion is similar to that for ANFO shown earlier.

By adding 16% water to the ANFO reaction described earlier, the theoretical heat energy release is reduced from 880 to 680 cal/g. The difference is the energy price paid for using water due to converting it to steam in the detonation reaction plus the energy loss in diluting the ANFO with a non-energy producing additive (water). The advantages and disadvantages of using ANFO or emulsions begin to become clear. ANFO is easily mixed and is probably the least expensive form of explosive energy, but it has no water resistance and has a relatively low loading density. Emulsions are considerably more complicated to formulate and manufacture, but they have excellent water resistance and more flexibility in terms of density, velocity, and higher bulk energy to match rock types and blasting applications.

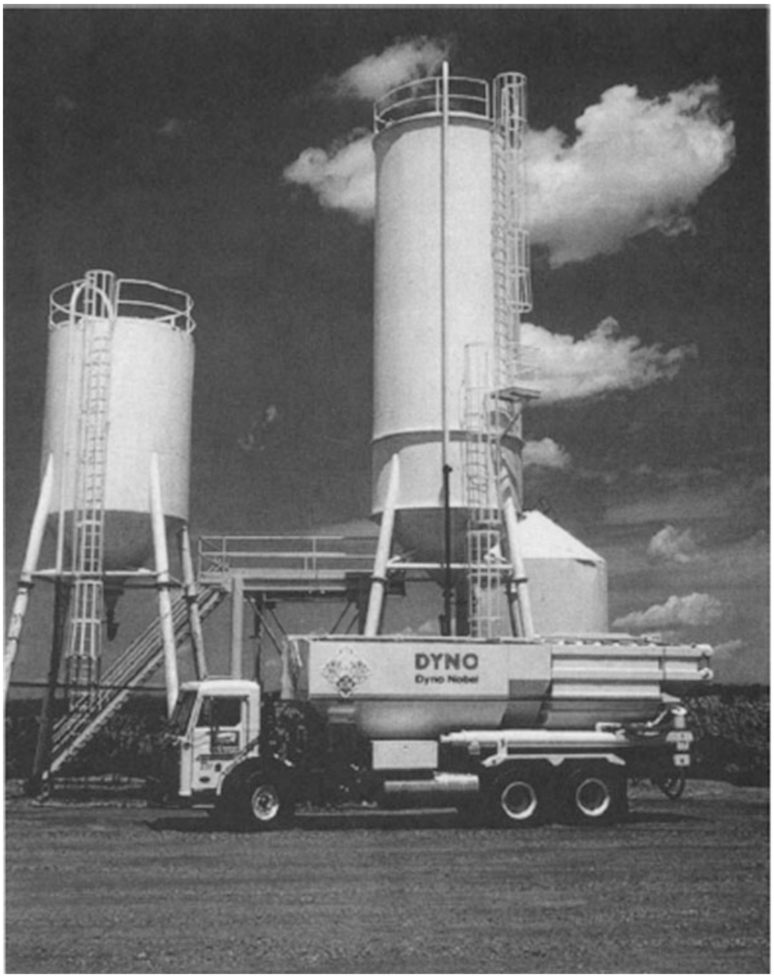


Fig. 37.12 Typical bulk explosives staging area in a large opt-pit mine

Initiation Systems

History of Initiation Systems

The first reliable initiation system for commercial explosives could probably be traced back to Alfred Nobel's invention of the blasting cap in 1864. These early caps had to be initiated with a strong shock, limiting their safety and convenience. In 1867, Nobel developed a cap that could be initiated with a fuse, a more convenient system that allowed an element of timing (length of the fuse) to the initiation of a blast, and hence, some increase in safety and predictability. This, combined with his invention of dynamite in 1866 [46], basically ushered in the modern era of blasting. The fuse cap and dynamite dominated the emerging blasting industry for several decades. In the century that followed, the initiation systems became more and more sophisticated and safe. The electric blasting cap which could be initiated by electrical current was developed around the turn of the century, adding further safety to blasting. Detonating cord, a flexible cord made of cloth or plastic with a core load of high explosives—usually PETN, was developed in Europe around the same time. Strings or circuits of detonating cord could be utilized to initiate several explosive charges with only one blasting cap. Electric caps were introduced in 1946 equipped with built-in, variable-delay elements (pyrotechnics that burn reproducibly) [46]. This allowed long rows of boreholes to be initiated in a sequential manner, optimizing blasting effectiveness as well as better controlling ground vibration and air blast damage to the surroundings. With caps having delays ranging from a few milliseconds to many hundreds and eventually thousands of milliseconds, the number of sequentially fired boreholes in a row depended only on the desired timing and the amperage capability of the electric blasting machine or other firing circuit. Sequential blasting machines were developed that could be used to control firing times between a number of rows of blastholes, greatly expanding the timing possibilities and size of blast patterns.

Fig. 37.13 Variety of commercial cast boosters (*Courtesy Dyno Nobel*)

Boosters

Prior to about 1950, most of the commercial explosives in the market were reliably detonable with just a blasting cap or detonating cord as the initiator. However, the use of non-cap-sensitive explosives (blasting agents) began emerging into the explosives market in the 1930s and 1940s. These products really took off with the advent of ANFO in the mid 1950s and water gels, invented by Melvin A. Cook in 1957 [47]. Emulsion based blasting agent products, which began significant commercialization in the mid 1970s, became the dominant non-ANFO product by the late 1980s and remains so to the present day. All of these blasting agents were adaptable to larger diameter packaged products and particularly to bulk loading.

These blasting agent explosive products were considerably less sensitive than dynamites and required larger "booster" charges for reliable detonation. At first, a high density and high velocity dynamite was used as the booster charge. Later, TNT-based cast boosters came into the market. Cast TNT by itself is not reliably detonable with a blasting cap or detonating cord, and so 40–60% PETN is normally added to the TNT melt and subsequent cast. The combination of TNT and PETN is called Pentolite. TNT has a melting point of about 80°C, which makes it an excellent base explosive for casting into forms. The military has used this concept for decades for filling bomb casings. Once the TNT has melted, other material can be added to give the final cast explosive composition the desired properties. Additives used for military purposes have included such as aluminum, ammonium nitrate, RDX, and also PETN, but Pentolite been dominant in commercial cast boosters. Cast boosters are available in a variety of sizes from about 10 to 800 g, as shown in Fig. 37.13, and continue to be predominant today in nearly all large mining operations and other blasting applications. Cap-sensitive composite explosives such as emulsion cartridges, dynamite, etc. are also used in less demanding situations. Coupled with the flexibility of blasting caps with variable delays to initiate the booster charge which then initiates the main charge, large patterns containing many thousands of pounds of explosive can be blasted.

Non-electric Initiation

In 1967, Per-Anders Persson of Nitro Nobel AB in Sweden invented a non-electric initiation system, designated Nonel®, that eventually revolutionized the explosives industry [46]. The Nonel system consists of an extruded hollow plastic tube (shock tube) that contains an internal coating of a mixture of powdered molecular explosive and aluminum. The plastic tube is inserted into and attached to a specially designed detonator or blasting cap. The Nonel tubing can be initiated by a number of starter devices, one of which uses a simple shotgun shell primer. The explosive/aluminum mixture explodes down the inside of the tube at about 2,000 m/s and will run at this velocity until all the interconnected tubing reacts, including initiating all the blasting caps. The tubing is about 3 mm outside diameter and 1 mm inside diameter, and the explosive core load is only about 18 mg/m, not even enough to rupture the tubing. The Nonel product is not susceptible to the hazard associated with electric blasting caps wherein premature initiation by extraneous electric sources can occur. Figure 37.14 is a photograph of both an electric blasting cap with the two electrical wires and a typical Nonel unit with the plastic tubing.

During the 1980s and 1990s, Nonel products continued to replace both electric blasting caps and detonating cord down lines around the world. It has long been known that detonating cord down lines disrupt and partially react with blasting agents causing some degree of energy and sensitivity loss. Also, the use of surface detonating cords to initiate blasts can lead to noise complaints. Using a range of long-lead delay caps down the borehole, a surface trunk-line of shock tubing, surface connectors to tie the tubing together, and surface delay elements that are available with this technology, huge and complex blasting patterns can be laid out and blasted with a multitude

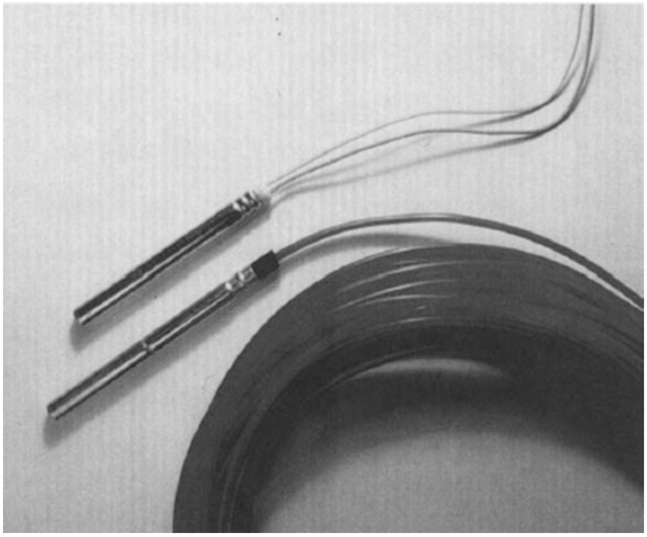


Fig. 37.14 An electric blasting cap showing electrical wires, typical Nonel unit with plastic tubing

of possibilities as to timing and sequence. As delay elements were perfected for the Nonel blasting caps, their application and use grew even further, especially in underground mining where a large percentage of blasting caps is used.

Electronic Detonators

The development of detonators with delay periods that were based on pyrotechnic elements (timing controlled by the burn rate and length of the element) had a dramatic effect on the explosives industry, enabling much of the size and complexity of modern blasting, as well as reducing off-site blast effects caused by ground vibration and air blast. When a blasting pattern is properly timed, the detonation of individual holes is such that damage from air blast and ground vibration is greatly reduced. This is a result of destructive wave interference, i.e., the ground vibration waves from detonation of each hole combine with each other and the air blast waves from the detonation of each hole combine with each other in such a way that the resulting wave patterns propagating from the blast have lower amplitudes and higher frequencies. This is key to controlling off-site structural damage. Over many years of development, the accuracy and repeatability of pyrotechnic delays was improved dramatically. However, delay times can still vary from detonator to detonator and from the target delay time by a few milliseconds to tens of milliseconds depending on the delay period of the detonator. There are also occasional fliers that occur well outside the normal scatter for a given delay detonator that can upset the blasting pattern. It was surmised that these variations could result in less than optimum blasts, considering not just air blast and ground vibration but also ground movement, fragmentation, ore dilution, fly-rock, etc.

In the late 1980s and early 1990s, several companies began developing and testing detonators wherein the delay timing was controlled with an integrated electronic circuit. These were initially very expensive and, by today's standards, were crude devices. However, they did demonstrate the promise of very accurate caps. As electronics technology improved into the 1990s, testing and field evaluation of such caps increased. However, cost and complexity in assembling the caps, the development of user friendly controllers to program and initiate them, and the unfamiliar equipment and technology on the blasting patterns slowed widespread implementation. It wasn't until mounting commercial experience began to demonstrate improved blasting results that the tide began to turn. This in turn resulted in accelerated improvements in manufacturing technology to make the caps less expensive, more reliable, improve ease of use, and increase expertise available to blasters in the field. By the 2000s, the technical case for electronic detonators had been well established. Despite the significant cost differences that remain between electric or non-electric detonators and electronic detonators, amounting to factors of perhaps five to ten times more costly per unit, there are dramatic increases in the use of electronic detonators in the blasting industry today. As of the date of this edition, they are still a minor part of the total global detonator market, but the market share is increasing. Several manufacturing companies with independent detonator technologies are involved and use is occurring in all segments of mining, construction, and other commercial blasting applications.

The construction of an electronic detonator has some features in common across all manufacturers. The timing and programmability features depend on an integrated circuit. One or more capacitors are included to store the electrical energy necessary to run the circuitry and to initiate the detonator via bridge or igniter elements. The detonator can be built to have a preset delay time or can be programmable as to delay period at the point of use. Electrical and electronic communication with the detonators is established via two or four lead wire connections, depending on the system chosen and the blast design, flexibility, control, and programmability features desired or needed by the blasters. The lead-wires are connected to the blasting machine which enables all of the features in the particular technology chosen, including programming (if needed), arming, and eventually initiating each line of detonators. The timing accuracy is on the order of only 1 to a few milliseconds depending on delay times, and variability is typically reduced by an order of magnitude compared to pyrotechnic detonators [48, 49]. Figure 37.15 shows the components of a four wire system

Fig. 37.15 Components of an electronic cap initiating system. *L–R*, logger, cap, base station (blasting machine), connectors (Courtesy Dyno Nobel)

and includes the cap, connectors, tagger (identifies and tags each detonator as it is placed in the borehole), and blasting machine (herein called the base station).

References

1. Mineral Information Institute. Soc Metallurg Explor Foundation. http://www.mii.org/MiiBabyMain
2. United States Geological Survey. Minerals Yearbook, Explosive Statistics and Information. Compiled Information from years 2000–2009
3. Apodaca LE (2009) United States Geological Survey. Minerals Yearbook, Explosive Statistics and Information (advance release)
4. Ornellas J (1968) Phys Chem 72:2390
5. Kaye SM (1978) Encyclopedia of explosives and related items, PATR 2700, vol 8. U.S. Army Armament Research and Development Command, Dover, pp 99–100
6. Engineering design handbook: explosives series. Army Material Command Pamphlet 706–177, AD 764, 340, p 12, distributed by NTIS, Jan 1971
7. Fedoroff BT, Sheffield OE (1962) Encyclopedia of explosives and related items, PATR 2700, vol 2. Picatinny Arsenal, Dover, p B165
8. Butler AR (1990) ChemTech:202
9. Van Gelder AP, Schlatter H (1972) History of the explosives industry in America. Arno Press, New York (reprint of 1927 edition)
10. Marshall A (1917) Explosives, vol I, History and manufacture. P. Blakiston's Son, Philadelphia
11. Gregory CE (1973) Explosives for North American engineers. Trans Tech Publications, Cleveland
12. "Black Powder" (1983) Explos and Pyrotech 16(7)
13. Atlas Powder (1987) Explosives and rock blasting. Atlas Powder, Dallas
14. Johansson CH, Persson PA (1970) Detonics of high explosives. Academic, Reinhold Book Corporation, New York
15. Du Pont (1980) Blasters handbook, 16th edn. pp 14, 109
16. Kintz GM et al (1948) Report of investigations 4245. U.S. Department of Interior, Bureau of Mines, Feb 1948
17. Chem Mark Rep:6 (1983)
18. Cook MA (1974) The science of industrial explosives. IRECO Chemicals, Salt Lake City, p 2
19. Egly RS, Neckar AE (1964) US Patent 3,161,551, 15 Dec 1964
20. Bluhm HE (1969) US Patent 3,447,978, 3 June 1969
21. Bower JK et al (1980) 1&EC Prod Res Dev:326
22. Mohan VK, Field JE (1984) Combust Flame 56:269
23. Spear RJ, Wilson WS (1984) J Energy Mater 2:61
24. Coburn MD et al (1986) Ind Eng Chem Prod Res Dev 25:68
25. Dobratz BM (1983) LA-9732-H, VC-45. Los Alamos National Laboratory, Los Alamos, May 1983
26. Cranney DH, Hales RH (1990) Proceedings of the fourteenth symposium on explosives and pyrotechnics, Franklin Research Center, Feb 1990
27. Code of Federal Regulations, Title 49 Transportation, Part 173.114a, p 272, 1 Oct 1979 Revision
28. Fed Regist 55(9):1306 (1990)
29. Tarver CM et al (1977) Structure/property correlations in primary explosives (Final report, 76-2). Stanford Research Institute, Menlo Park, 4 Feb 1977
30. Brunswig H (1909) Explosivstoffe. Barth, Braunschweig, p 17
31. Plets V (1953) Zh Obshch Khim 5:173
32. Shipp KG (1964) J Org Chem 29:2620
33. US Patent 3,505,413, 7 Apr 1970
34. Kayser EG (1983) J Energy Mater 1:325
35. Gallo AE, Tench N (1984) J Hazard Mater 9:5
36. Delistraty I, Brandt H (1982) Propell Explos Pyrotech 7:113
37. Rizzo HE et al (1981) Propell Explos 6:27
38. Kolb IR, Rizzo HE (1979) Propell Explos 4:10
39. Ott DG, Benziger TM (1987) J Energy Mater 5:343
40. Urbanski T (1965) Chemistry and technology of explosives, vol II. Pergamon, New York, p 181
41. Military explosives, Department of the Army technical manual TM 9-1300-214, chapter 7, p 32, Washington, DC, 1967

42. Cook MA (1958) The science of high explosives. American Chemical Society monograph series no. 139, Reinhold Book Corporation, pp 178–183
43. Ibid., Appendix II
44. Cole RH (1948) Underwater explosions. Princeton University Press, Princeton
45. Hanto K (1998) Internal report on fumes from emulsion and ANFO. Dyno Nobel, Europe
46. History of explosives and blasting. Intern Soc Explos Eng. http://www.explosives.org/HistoryofExplosives
47. Cook MA (1974) The science of industrial explosives. IRECO Chemicals, Salt Lake City, pp 5–7
48. Watson J (2002) Electronic Blast Initiation, A Practical Users Guide, J Explos Eng 19(3):6
49. Lusk BT, Hoffman J, Wedding WC (2011) Electronic Detonator and Modern Non-Electric Shocktube Detonator Accuracy, Proceedings thirty-seventh annual conference on explosives and blasting technique, International Society of Explosives Engineers, p 563, 2011

Electrochemical Energy Storage: Applications, Processes, and Trends

38

Gerardine G. Botte and Madhivanan Muthuvel

Introduction

Energy consumption in the world has increased significantly over the past 20 years. In 2008, worldwide energy consumption was reported as 142,270 TWh [1], in contrast to 54,282 TWh in 1973; [2] this represents an increase of 262%. The surge in demand could be attributed to the growth of population and industrialization over the years. In 2009, energy consumption was reported as 140,700 TWh, a slight decrease (1.1%) when compared to 2008 due to the world financial crisis [1], while in 2010 there was a rise in the consumption to 149,469 TWh, due to the recovery of the economy at that time [3]. Conversely, the total supply of energy in the world had caught up with the consumption as shown in Table 38.1 [2, 4]. Approximately 10–14% of the total energy supply in the world is delivered as electric energy. In addition, the amount of power supplied by renewables had increased over the years, from 37 TWh in 1973 to 612 TWh in 2008 (as shown in Table 38.1), which represents a growth of 94%. However, the total amount of energy available from renewables based on current technology could reach up to 834,280 TWh (distributed as: 53.2% solar, 20.0% wind, 16.7% geothermal, 8.4% biomass, and 1.7% hydropower); [5] that is, 5.7 times the world energy supply in 2008. Nevertheless, renewable sources of energy such as solar and wind are intermittent and only abundant in certain regions, which causes a limitation on the use and distribution of such sources of energy. An undersized world energy surplus (based on a total energy balance including supply, consumption, and losses) is usually reported annually; a comprehensive analysis is presented in the literature [2].

Given the increase in energy consumption as the world's population grows, the scarcity of traditional energy supplies (i.e., petroleum, oil, and gas), and the environmental impact caused by conventional power generation systems, it has become imperative to utilize unconventional energy sources and renewables, and to redesign traditional processes to make them more energy efficient. Within this context, energy storage and conversion play a significant role. Energy storage refers to devices, or physical media, that collect different types of energy to be used at a later time. Perhaps the use of devices to accumulate energy is the most popular way, as it brings to mind the term "batteries," which has become extremely important with the spiraling growth of modern electronic applications. Indeed, batteries are a way to store energy. Batteries are unique devices, as the energy that is stored is chemical, but delivered as electrical. That is, batteries play a dual role: they store chemical energy, which is then delivered or converted directly into electrical energy when needed (we will return to this definition later). The definition of energy storage also includes physical media, which can be easily related to fuels (e.g., gasoline, diesel, hydrogen). Fuels are ways to store chemical energy that needs to be conventionally converted into electrical energy quite differently than the way batteries do it. Traditional power generation involves heat engines (e.g., turbines). In this process, a fuel is combusted to release thermal energy, the thermal energy is then transformed into kinetic energy through a rotor (i.e., using a turbine), and finally the kinetic energy is transformed into electrical energy by an electrical generator. The efficiency of the conventional process to convert chemical energy into electrical energy is limited by the Carnot cycle and is much lower than the efficiency of batteries.

There are different ways to store energy: chemical, biological, electrochemical, electrical, mechanical, thermal, and fuel conversion storage [6]. This chapter focuses on electrochemical energy storage and conversion. Traditionally, batteries, flow batteries, and fuel cells are considered as electrochemical energy storage devices. However, the

---

G.G. Botte (✉) • M. Muthuvel
Department of Chemical and Biomolecular Engineering,
Center for Electrochemical Engineering Research, Ohio University,
Athens, OH, USA
e-mail: botte@ohio.edu

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2_38,

authors believe that with the growth of renewable energy and intermittent energy sources, the concept of electrochemical energy storage can be extended to the electrochemical synthesis and production of fuels, chemicals, petrochemicals, etc. The vision of the approach is shown in Fig. 38.1 [7]. For example, the inspiration for the hydrogen economy is to use hydrogen as an energy carrier. The intention is to produce hydrogen through water electrolysis (electrochemical synthesis of hydrogen from water) by using electricity (electrical power) from intermittent (i.e., not available all the time) renewable power generation devices, such as solar panels and wind turbines. This way, hydrogen could be used to produce electricity when the intermittent source of electricity is not available, and to distribute the energy to remote sites. Conversely, the intention is to produce electricity by using hydrogen as the fuel in devices called fuel cells (we will return to this definition later). With this in mind, the vision of Fig. 38.1 is broader than the concept of the hydrogen economy. The final goal is to use intermittent and renewable sources of electricity (which represents 5.7 times the supply of energy in the world) to transform different feedstock into fuels, chemicals, petrochemicals, etc. That is, intermittent energy sources are used to store energy as products, commodities, or fuels via electrochemical synthesis.

Table 38.1 Energy and power supply in the world

Year	Energy (TWh)	Power (TWh)	Fraction of energy supplied as power (%)	Power from renewables[a] (TWh)	Fraction of power from renewables (%)
1973	71,117	6,116	9	37	0.6
1990	102,569	11,821	12	277	2.3
2000	117,687	15,395	13	303	2.0
2001	118,147	15,502	13	301	1.9
2002	120,451	16,104	13	340	2.1
2003	124,675	16,688	13	361	2.2
2004	129,568	17,476	13	404	2.3
2005	133,602	18,258	14	450	2.5
2006	136,893	18,971	14	481	2.5
2007	139,390	19,804	14	561	2.8
2008	143,851	20,181	14	612	3.0

Sources
For 1973: IEA [2]
For 1990–2008: Energimyndigheten [4]
[a] Includes geothermal, solar, wind, combustible renewables and waste, and heat

In this chapter, the authors outline the basic concepts and theories associated with electrochemical energy storage, describe applications and devices used for electrochemical energy storage, summarize different industrial electrochemical processes, and introduce novel electrochemical processes for the synthesis of fuels as depicted in Fig. 38.1.

Basic Concepts

There are some important terminologies and fundamental theories associated with electrochemistry and electrochemical engineering. Electrochemical engineering is the branch of engineering dealing with the technological application of electrochemical phenomena. Electrochemical engineering evolved from industrial electrochemistry, which was defined in the early years as [8, 9] "the application of electric current

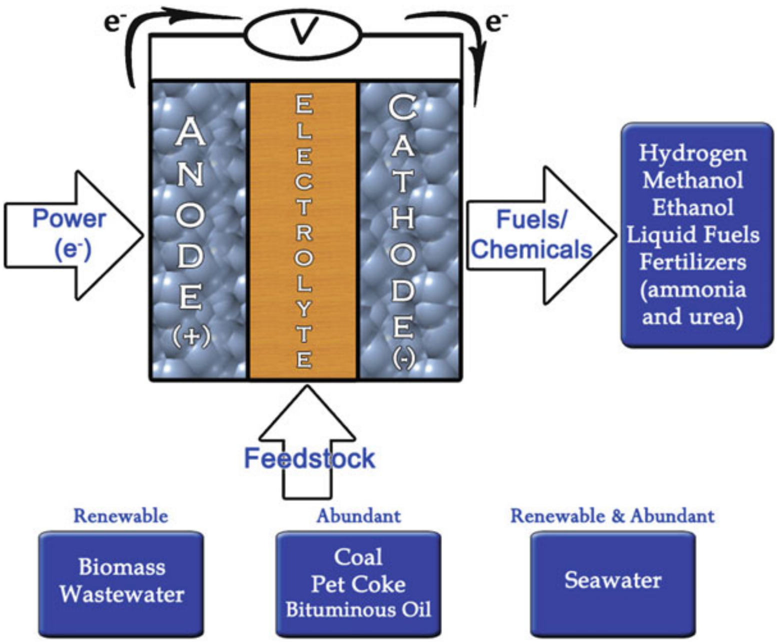


Fig. 38.1 Electrochemical synthesis of fuels, chemicals, and petrochemicals from different feedstocks [7]. Within the vision electrical power could be supplied from different sources, however, the concept creates the ideal platform to use renewal electrical energy from solar panels and wind turbines

to the chemical and metallurgical arts.” In this section, terminologies that are pertinent to the electrochemical energy storage devices are described.

The basis for a traditional electrochemical energy storage system (batteries, fuel cells, and flow batteries) and the extended electrochemical energy storage concept presented in Fig. 38.1, known as electrosynthesis, is the electrochemical cell. An electrochemical cell consists of a case, an anode, a cathode, an electrolyte, and current collectors (as shown in Fig. 38.1). Electrochemical cells allow the direct conversion of chemical energy into electrical energy (galvanic cells or spontaneous cells) or vice versa, the direct conversion of electrical energy into chemical energy (electrolytic or electrolysis cells). Traditional electrochemical energy storage devices, such as batteries, flow batteries, and fuel cells, are considered galvanic cells. The approach depicted in Fig. 38.1, electrosynthesis reactor, is defined as an electrolytic or electrolysis cell. Electrochemical cells can be electrically connected in series, in parallel, or other configurations according to the needs of particular processes or applications.

The basis for the direct conversion of chemical to electrical energy, or vice versa, is the electrochemical reaction. An electrochemical reaction is a heterogeneous chemical process involving the transfer of electrons to or from an electrode (generally a metal, carbon, or electrically conducting material). It is important to mention that electrochemical reactions take place only on the electrode surface; this distinguishes an electrochemical system from a traditional chemical reactor.

Within this context, other important definitions are presented in subsequent sections. Additional definitions, thermodynamics, transport mechanisms, and electrokinetic models can be found elsewhere [10–13].

Oxidation

The removal of electrons from an element or ion, even if the element is part of a compound, during a chemical reaction is known as an oxidation reaction. Corrosion of an iron sheet, also known as rusting, is an electrochemical process. The rusting of an iron sheet is a good example of an oxidation reaction, where the elemental iron is oxidized to ferric oxide (Fe_2O_3) by oxygen present in the electrolyte (38.1). The oxidation reaction for iron is shown in (38.2) where the elemental iron losses three electrons to form ferric (Fe^{3+}) ion. The rusting of iron will be accelerated in an acid electrolyte.

$$4Fe + 3O_2 \rightarrow 2Fe_2O_3 \tag{38.1}$$

$$Fe \rightarrow Fe^{3+} + 3e^- \tag{38.2}$$

Reduction

The reduction process is also related to transfer of electrons, where electrons are added to either the element or ion during the chemical reaction. For example, electroplating of copper metal on any electrically conducting surface is possible by the reduction of cupric (Cu^{2+}) ions found in the plating solution to elemental copper, by accepting two electrons provided by applied current supplied through the electrodes.

$$Cu^{2+} + 2e^- \rightarrow Cu \tag{38.3}$$

Anode

There are two established configurations for the electrochemical cells or systems: three-electrode and two-electrode systems. Generally, the two-electrode configuration is used in the commercial electrochemical devices, as shown in Fig. 38.1. In this two-electrode configuration, the electrochemical system consists of an anode, a cathode, and an electrolyte. The anode is the electrode where the oxidation reactions take place and it is connected to the positive terminal of the power supply unit. In order for the electrochemical oxidation reaction to take place on the anode, the anode has to be an electrically conducting material. Anode materials can be either insoluble or soluble substances depending on the electrochemical system. In the electroplating of precious metals, such as platinum, the anode material is made of platinum which is insoluble in the plating solution. On the other hand, for nickel electroplating, nickel sheet is used as the anode, which dissolves into the plating solution as nickel ions. The nickel ions eventually reduce as metallic nickel on the cathode electrode.

Cathode

On the other hand, the reduction reactions, such as (38.3), will take place at an electrode known as the cathode. The cathode electrode is connected to the negative terminal of the power supply unit. Similar to the anode material, the cathode material should also be electrically conducting so that the electrochemical reaction (reduction reaction) can take place. Generally, insoluble materials have been used as the cathode for the electrochemical systems. In the case of chrome plating on steel materials, the steel is the cathode electrode during the electroplating process.

The three-electrode system contains a working electrode, a counter electrode, and a reference electrode (will be discussed later). Either the working or the counter electrode could act as a cathode or an anode. The term working electrode is used to define the electrode that will be exposed

Table 38.2 Examples of different electrolytes used in electrochemical energy storage technologies and electrosynthesis

Electrolyte system	Supporting salt	Solvent	Conductivity (mS/cm)	Applications and details
1 M $LiPF_6$ in EC/DMC (50:50 wt. %)	$LiPF_6$	Organic solvents—EC and DMC	10.7 (at 25°C) [15]	Lithium ion batteries. No water should be present in the system. Nonaqueous electrolyte
H_2SO_4 (30 wt. %)	H_2SO_4	Water	730 [14]	Lead acid batteries and electroplating (low concentrations of H_2SO_4). Aqueous electrolyte
0.1 M KOH	KOH	Water	21.3 (at 25°C) [19]	Electrolytic systems and electrodeposition. Aqueous electrolyte
KOH (29.4 wt. %)	KOH	Water	540 [14]	Alkaline battery. Aqueous electrolyte
EC/PC/$LiClO_4$/PAN (38/33/8/21) (mol%) Solid polymer electrolyte	$LiClO_4$	EC, PC	1.7 (at 20°C) [16]	Lithium polymer battery. This is a solid polymer electrolyte
PC/$LiClO_4$/2p-MC-DBS/PEGPM (81.0/7.5/0.9/10.6) (wt. %)	$LiClO_4$	PC	4.6 (at 25°C) [16]	Lithium battery. Gel electrolyte in polymer matrix
$Na_3AlF_6/AlF_3/CaF_2/Al_2O_3$ (81/11/5/3) (wt. %)	Al_2O_3	Cryolite (Na_3AlF_6)	2,130 (at 950°C) [17]	Aluminum electrolysis. Molten electrolyte
$(Y_2O_3)_{0.08}(ZrO_2)_{0.92}$ Yttria-stabilized zirconia (YSZ)	Y_2O_3	ZrO_2	30 (at 800°C) [18]	Solid oxide fuel cells. This is a solid electrolyte, where yttria is doped in zirconia

EC ethylene carbonate; *DMC* dimethyl carbonate; *PC* propylene carbonate; *PAN* poly(acrylonitrile); *2p-MC-DBS* 1,3:2,4-di(*p*-methoxycarbonylbenzylidene) sorbitol; *PEGPM* poly(ethylene oxide)—grafted poly(methacrylate)

a. Schmidt M, Heider U, Kuehner A, Oesten R, Jungnitz M, Ignat'ev N, Sartori P (2001) Lithium fluoroalkylphosphates: a new class of conducting salts for electrolytes for high energy lithium-ion batteries. Journal of Power Sources 97–8:557–560[15]

b. Dobos D, (1975) Electrochemical Data. Akademiai Kiado, Budapest[14]

c. Haynes WM, ed. (2011) CRC Handbook of Chemistry and Physics. 92nd Edition (Internet Version 2012). CRC Press/Taylor and Francis, Boca Raton, FL[19]

d. Gores HJ, Barthel JMG (1995) Nonaqueous electrolyte solutions: New materials for devices and processes based on recent applied research. Pure and Applied Chemistry 67:919–930[16]

e. Hives J, Thonstad J, Sterten A, Fellner P (1996) Electrical conductivity of molten cryolite-based mixtures obtained with a tube-type cell made of pyrolytic boron nitride. Metallurgical and Materials Transactions B 27:255–261[17]

f. Yamamoto O (2000) Solid oxide fuel cells: fundamental aspects and prospects. Electrochimica Acta 45:2423–2435[18]

to a detailed electrochemical analysis through the use of different polarization techniques [11]. The counter electrode is used to carry enough current that will not prevent or limit any of the processes taking place at the working electrode. The three-electrode cell configurations are typically used for detailed and fundamental understanding of electrochemical reactions [11]. However, they also find applications in sensors.

Electrolyte

In an electrochemical cell, the electrodes—anode and cathode—are immersed in a solution (liquid, gel, or solid); the solution is known as the electrolyte. The electrolyte provides the medium for the movement of the ions (mobility) between the anode and the cathode. The operation of an electrochemical cell functions as a closed electrical system. Externally, the current travels between the power supply unit and the electrodes through the electrically conducting wires (see Fig. 38.1). In the electrolyte (internally), the movement of ions between the anode and the cathode (current flows internally through the transport of the ions) will complete the electrical circuit for the electrochemical cell. That is the anode/electrolyte/cathode, and power supply complete the electric circuit.

The flow of the electrons through the external circuit is relatively fast due to the use of highly electrically conducting materials. The limitation in the flow of current in an electrochemical system is experienced usually through the internal media: anode/electrolyte/cathode due to the diffusion and migration of ions across the electrolyte. The electrolyte can be aqueous, nonaqueous, or molten salt solutions. The electrolyte can also be made of a solid mixture. A solution of sulfuric acid is one of the most commonly used aqueous electrolytes. A nonaqueous electrolyte is lithium perchlorate ($LiClO_4$) dissolved in propylene carbonate, which is used in the lithium batteries. For the production of metallic sodium, molten salt of sodium chloride is electrolyzed, where molten sodium chloride is the molten electrolyte. A good example for a solid electrolyte is yttria-stabilized zirconia (YSZ), which is a ceramic material used as solid oxide electrolyte in solid oxide fuel cells (SOFCs).

The electrochemical reactions taking place on the surface of the electrodes depend on the transport properties of the ions in the electrolyte medium. One of the important properties of the electrolyte is its ionic conductivity, which is affected by factors, such as temperature, supporting salt concentration, solvent, and pressure. Table 38.2 lists some of the different electrolytes used in the energy storage systems

Table 38.3 Classical separators, membranes, and diaphragms used in electrochemical energy storage technologies and electrosynthesis

Material	Type	Applications and details
Nafion (sulfonated tetrafluoroethylene-based fluoropolymer-copolymer)	Membrane	Chlor-alkali cell and proton exchange membrane fuel cell. Allows transport of cations, especially protons
Cellophane	Membrane	Nickel cadmium and silver oxide batteries. Provides gas barrier along with ion permeability
Polyethylene	Separator	Lead acid batteries. These separators with small pore size are used in heavy duty batteries and they provide longer life to the battery
Polypropylene	Separator	Lithium-based batteries. The separator is used in primary and secondary types of lithium batteries
Asbestos	Diaphragm	Chlor-alkali cell. This separator was used in the original design of chlor-alkali cell. The health problems associated with asbestos has restricted its application as separator
Teflon (polytetrafluoroethylene)	Diaphragm	Chlor-alkali cells and aqueous electrolysis. Chemically inert to acids, alkalis, and organic solvents

Sources: Linden [20]; Kuhn [83]

along with their conductivity. Usually aqueous electrolytes possess the highest ionic conductivity at room temperature. As shown in Table 38.2, the 30 wt. % sulfuric acid solution (aqueous electrolyte) used in the lead acid battery (LAB) has an ionic conductivity of 730 mS/cm [14]. The nonaqueous electrolyte used in commercial lithium ion (Li-ion) batteries consisting of 1 M $LiPF_6$ dissolved in the organic solvents of ethylene carbonate and dimethyl carbonate (50:50 wt. %) has an ionic conductivity of 10.7 mS/cm at 25°C [15]. Lithium batteries also use gel type electrolytes, such as PC/$LiClO_4$/2p-MC-DBS/PEGPM (81.0/7.5/0.9/10.6) (wt. %), where PC stands for propylene carbonate, 2p-MC-DBS means 1,3:2,4-di(*p*-methoxycarbonylbenzylidene) sorbitol, and PEGPM is poly(ethylene oxide) (PEO)—grafted poly (methacrylate). This gel electrolyte has an ionic conductivity of 4.6 mS/cm at 25°C [16]. The molten electrolyte used in the aluminum electrolysis process has alumina (Al_2O_3) mixed with CaF_2 and AlF_3 additives, which are dissolved in cryolite (Na_3AlF_6) solvent. The ionic conductivity for this molten electrolyte is 2,130 mS/cm at 950°C [17]. In SOFCs, zirconia (ZrO_2) doped with 8% of yttria (Y_2O_3), commonly known as yttria-stabilized zirconia, is used as the solid electrolyte. YSZ has an ionic conductivity of 30 mS/cm at 800°C [18]. The readers are requested to review the following literature for more details about transport processes in electrochemical systems and conductivity [10, 11].

Separator

In a divided electrochemical cell, an electrically insulated layer material is used to separate the electrode, electrolyte, and any products formed between the anode and the cathode side. This electrically insulated material is called the separator. The separators should be permeable to the ions present in the electrolyte to facilitate the completion of the electrical circuit. The different types of separators are membranes and diaphragms. The membranes have a smaller pore size which will result in selective permeability of ions across the compartments. On the other hand, diaphragms are those separators that have larger pore size and permits transfer of many ions at higher flow rate across the compartment without leading to intermixing of electrolytes. In general, the distinction between membrane and diaphragm is blurred as some separators can be engineered to have different porosity and selectivity for permeation. A list of various commonly used separators (membranes and diaphragms) is listed in Table 38.3.

Reference Electrode

In a three-electrode cell system, a third electrode other than the anode and cathode, the reference electrode, is used. Reference electrodes are standard electrodes, which are chemically and mechanically robust, and their potential is expected to remain unchanged under most electrolytic conditions. In electrochemistry, one cannot measure the potential of an electrode, but the potential difference between two electrodes is measureable. If one of the two electrodes is a standard or reference electrode, then the potential of the other electrode can be determined in terms of the reference electrode. So, every time a potential value is assigned to an electrode, it is always expressed with respect to a reference electrode.

The standard hydrogen electrode (SHE) is the primary reference electrode and its potential is assigned a zero value (see section on “Standard Potential”). For example, consider an electrolytic system where deposition of copper is carried out using an electrolyte containing copper sulfate and sulfuric acid. At the anode, a water molecule is oxidized to oxygen and protons, whereas metallic copper is deposited at the cathode electrode. The following half cell reactions (anode and cathode reaction) with their standard reduction potential illustrate the electrolytic cell.

Table 38.4 Standard reduction potential for some electrochemical reactions

Reaction	$E°$(V)
$Li^+ + e^- \rightleftharpoons Li$	−3.0401
$Na^+ + e^- \rightleftharpoons Na$	−2.71
$Al(OH)_3 + 3e^- \rightleftharpoons Al + 3OH^-$	−2.31
$Te + 2e^- \rightleftharpoons Te^{2-}$	−1.143
$Ni^{2+} + 2e^- \rightleftharpoons Ni$	−0.257
$Pb^{2+} + 2e^- \rightleftharpoons Pb$	−0.1262
$Fe^{3+} + 3e^- \rightleftharpoons Fe$	−0.037
$2H^+ + 2e^- \rightleftharpoons H_2$	0.0000
$NO_3{}^- + H_2O + 2e^- \rightleftharpoons NO_2{}^- + 2OH^-$	0.01
$Cu^{2+} + e^- \rightleftharpoons Cu^+$	0.153
$Cu^{2+} + 2e^- \rightleftharpoons Cu$	0.3419
$Ag^+ + e^- \rightleftharpoons Ag$	0.7996
$Pt^{2+} + 2e^- \rightleftharpoons Pt$	1.18
$Au^{3+} + 3e^- \rightleftharpoons Au$	1.498
$F_2 + 2e^- \rightleftharpoons 2F^-$	2.866

Source: Haynes [19]

Anode reaction:

$$H_2O \rightarrow \frac{1}{2}O_2 + 2H^+ + 2e^- \quad E^\circ = 1.23 \text{ V vs. SHE} \quad (38.4)$$

Cathode reaction:

$$Cu^{2+} + 2e^- \rightarrow Cu \quad E^\circ = 0.34 \text{ V vs. SHE} \quad (38.5)$$

The secondary reference electrodes have been frequently used in the electrochemical systems because they are easier to construct and maintain than the SHE. One of the commonly used reference electrodes is the saturated calomel electrode (SCE), which has a potential of 0.242 V vs. SHE. Another reference electrode used in the acidic electrolytes has been the silver-silver chloride electrode and its potential is 0.197 V vs. SHE.

Standard Potential

The standard potential ($E°$) is related to an electrochemical reaction on the electrode surface at standard condition of 25°C and 1 atmospheric pressure. The standard potential for the reduction or oxidation reaction of the chemical substance present in the electrochemical system is valuable information principally to determine the possible reactions taking place in the system. In general, the standard potential is represented as the reduction potential for that reaction vs. the SHE. Table 38.4 contains the standard reduction potential of commonly used reactions.

The formation of hydrogen from protons is zero volts as the reaction is measured with respect to an SHE. The standard reduction potential for lithium ion (Li^+) is −3.0401 V which is the least valued potential in Table 38.4. The most negative value of the standard reduction potential means that lithium (Li) is more likely to get oxidized into Li^+ ion. On the other hand, the reduction of the platinum ion (Pt^{2+}) to Pt (1.18 V) is more favorable than the Li^+ ion reduction.

Table 38.4 lists only a few reactions and its standard reduction potential, but a complete list for the standard reduction potential of various reactions can be found in the CRC Handbook for Chemistry and Physics [19]. The list of the standard reduction potential suggests thermodynamically which reaction would occur, or is feasible, in an electrochemical system. For example, a platinum electrode immersed in a solution containing 0.1 M Ni^{2+} in 1 M HCl will result in an open circuit potential greater than 0 V. Once the potential of the Pt electrode is shifted in the negative direction with respect to SHE, the reaction with most positive reduction potential will occur first. So, in this case the evolution of hydrogen (0 V vs. SHE) will occur even before the deposition of Ni on the Pt electrode (−0.257 V vs. SHE). On the other hand, by shifting the potential of the electrode in the positive direction, the reaction with least positive potential will occur first. If a gold electrode is immersed in a Cu^+ ion solution and the potential of the Au electrode was increased in the positive direction, then the oxidation of Cu^+ to Cu^{2+} (0.153 V vs. SHE) would take place before the gold electrode oxidizes (1.498 V vs. SHE). The table covering standard reduction potentials provides an opportunity to perform controlled and specific reactions on the electrode surface because each electrochemical reaction can be controlled by its potential.

Nernst Equation

The potential for an electrochemical reaction can be represented in terms of a thermodynamic quantity, the Gibbs free energy change. Equations (38.6) and (38.7) display the relationship between the Gibbs free energy change and the potential for the reaction at normal and standard conditions. Standard conditions are referred to temperature of 25°C and 1 atmospheric pressure with unit activity for the chemicals involved in the electrochemical reaction. The Gibbs free energy for the reaction taking place at conditions other than the standard condition is represented in (38.6).

$$\Delta G = -nFE \quad (38.6)$$

$$\Delta G^\circ = -nFE^\circ \quad (38.7)$$

where ΔG° and E° are the standard Gibbs free energy change and the standard reaction potential, respectively. On considering a simple electrochemical reaction at equilibrium as given in (38.8), an oxidized species (O) takes up n electrons to form the reduced species (R).

$$O + ne^- \rightleftharpoons R \tag{38.8}$$

From the basic thermodynamics, the Gibbs free energy for the above reaction can be written as

$$\Delta G = \Delta G^\circ + RT \ln \frac{a_R}{a_O} \tag{38.9}$$

where R is the molar gas constant (8.314 J mol^{-1} K^{-1}), T is the temperature (K), a_R and a_O are the activities for the reduced and oxidized species, respectively. On substituting the Gibbs free energy change from (38.6) and (38.7) to the (38.9), we obtain

$$E = E^\circ - \frac{RT}{nF} \ln \frac{a_R}{a_O} \tag{38.10}$$

The Nernst equation for an electrochemical reaction is represented by rearranging (38.10).

$$E = E^\circ + \frac{RT}{nF} \ln \frac{a_O}{a_R} \tag{38.11}$$

The potential of the reaction is dependent on the standard reduction potential (E°), and the activities of the oxidized and reduced species. The activity is a chemical thermodynamic term meaning effective concentration of the species in a mixture. The activity of the species is expressed as the product between the activity coefficient (γ) and the concentration of the species. Since, it is difficult to determine the activity of the species in an electrochemical reaction, the Nernst equation can be defined in terms of the species concentration (38.12).

$$E = E^{\circ'} + \frac{RT}{nF} \ln \frac{[O]}{[R]} \tag{38.12}$$

where the concentrations of oxidized and reduced species are $[O]$ and $[R]$, respectively, and $E^{\circ'}$ is the formal potential of the reaction. The formal potential is defined as the measured potential of the reaction when the ratio of the concentrations of oxidized and reduced species is unity. The following equation displays the relationship between the formal potential and the activity coefficient of the species.

$$E^{\circ'} = E^\circ + \frac{RT}{nF} \ln \frac{\gamma_O}{\gamma_R} \tag{38.13}$$

An electrode following the Nernst equation is referred to as the Nernstian system when the electrode potential changes because of the changes observed in the concentration of the species present in the electrochemical system. Alternatively, if the potential of an electrode does not follow the Nernst equation, then it is considered to be non-Nernstian system.

Cell Voltage

In a two electrode system, the potential difference between the anode and cathode electrodes is defined as the cell voltage. The term “potential,” which has been used extensively earlier, is only referred to as the potential difference observed between an electrode and a reference electrode. The cell voltage is denoted by V; it also uses the same thermodynamic relation with the change in the Gibbs free energy as shown in (38.6). The units for cell voltage are volts (V), which is also the unit for the electrode potential.

As described earlier, there are two types of electrochemical cells, namely electrolytic and galvanic cells. In a galvanic cell, such as silver oxide-zinc battery, a primary battery system, the cell voltage can be determined using the anode and cathode electrode potentials. The anode, cathode, and the overall reactions for the silver oxide-zinc battery are given below [20].

Anode reaction:

$$\begin{aligned} &\mathrm{Zn} + 2\mathrm{OH}^- \rightarrow \mathrm{Zn(OH)_2} + 2\mathrm{e}^- \\ &E^\circ = -1.249 \text{ V vs. SHE} \end{aligned} \tag{38.14}$$

Cathode reaction:

$$\begin{aligned} &\mathrm{Ag_2O} + \mathrm{H_2O} + 2\mathrm{e}^- \rightarrow 2\mathrm{Ag} + 2\mathrm{OH}^- \\ &E^\circ = 0.342 \text{ V vs. SHE} \end{aligned} \tag{38.15}$$

Overall reaction:

$$\begin{aligned} &\mathrm{Zn} + \mathrm{Ag_2O} + \mathrm{H_2O} \rightarrow \mathrm{Zn(OH)_2} + 2\mathrm{Ag} \\ &E = 1.591 \text{ V} \end{aligned} \tag{38.16}$$

At the anode electrode, the Zn is oxidized to zinc hydroxide in the presence of hydroxyl ions from the alkaline electrolyte and the silver oxide is reduced to silver at the cathode electrode. The overall reaction is the formation of zinc hydroxide and silver from zinc and silver oxide, which results in a generating cell voltage of 1.591 V. The voltage for the overall reaction was obtained by subtracting the standard potential of the anode reaction from the cathode reaction. This cell voltage is also known as the theoretical voltage for the battery [20]. The theoretical voltage of a battery represents the maximum cell voltage that can be attained by the battery. In practical applications, the theoretical voltage is very difficult to reach due to various factors, such as concentration, temperature, and transport of ions.

The influence of these factors leads to ohmic limitations, concentration limitations, surface limitations, and diffusion limitations during the operation of any galvanic system. For the silver oxide-zinc battery, even though it has the theoretical voltage of 1.591 V, the nominal voltage or voltage during practical application is only 1.5 V [20].

In the case of an electrolytic system, such as the electrolysis of water, the theoretical cell voltage required to split water into hydrogen and oxygen gases is 1.23 V. In the electrolytic system, the theoretical voltage is the minimum required voltage to initiate the electrochemical reaction and due to various factors the cell voltage has to be maintained more than the theoretical value in order to sustain the reaction. The key factors affecting the operation of electrolytic systems are the solution resistance, diffusion limitations faced by the ions at the electrode surface, occurrence of secondary or undesired reactions at the electrode surface, and even the resistance to the flow of electrons along the wire or electrical contact. Even though the theoretical voltage required for water electrolysis is 1.23 V, the commercial alkaline water electrolyzers operate between 1.7 and 1.9 V to compensate for the losses due to the above mentioned factors [21].

Faraday's Law

In an electrochemical reaction, electron transfer results in the formation of a new product or loss of a compound. The mass of the new product formed or the loss of the existing material is directly proportional to the number of electrons passed through the electrode. This is known as Faraday's law. The time rate of flow of electrons is known as current, and the number of electrons is referred to as charge (Q), and its units are coulombs (C).

$$m = \frac{Q \times M}{n \times F} \tag{38.17}$$

where m is the mass of the substance (g) formed or lost during the electrochemical reaction, M is the molar mass of the substance (g/mole), n is the number of electrons transferred per mole during the reaction, and F is the Faraday constant, which equals 96,485 C.

The mass of the material deposited or lost is directly related to the current passed through the electrochemical cell. If all the current supplied to the electrochemical cell is utilized for the expected reaction, then the Faraday efficiency is 100%. In most systems, however, part of the applied current results in undesired or secondary reactions and generates other products, which leads to lower than 100% in Faraday efficiency. The ratio of actual weight of the material formed to the theoretical weight based on that current is known as Faraday efficiency. In some places, Faraday efficiency is also referred as Faradaic efficiency, coulombic efficiency, or current efficiency. Faraday's law is a valuable concept along with current efficiency to estimate the total current required for scaling up an electrochemical process, as it would lead to the calculation of the number of cells, electrodes, and energy consumption needed for the process.

The use of Faraday's law can be demonstrated with an electrolytic system, aluminum production. The production of aluminum is an energy consuming process and it is performed by electrolysis of molten salt of alumina (Al_2O_3). In the aluminum industry, the modern aluminum reduction cell system is supplied with a total current of 300 kA and a cell voltage of 4.2 V is maintained for every day operation [22]. The molar mass for aluminum (M) is 26.98 g/mol and the number of electrons (n) involved during the reduction reaction is 3. The total charge, the product of current and time, supplied for a day to the electrochemical cell for aluminum reduction is Q.

$$\begin{aligned} Q &= \text{current} \times \text{time} = 300\ \text{kA} \times 24\ \text{h} \times 60\ \text{min} \times 60\ \text{s} \\ &= 25,920 \times 10^6 \text{coulombs} \end{aligned}$$

Substituting the values for charge (Q), molar mass (M), number of electrons (n), and Faraday's constant in (38.17), we obtain the theoretical amount of aluminum produced per day using 300 kA of current. The theoretical aluminum calculated was 2,416 kg/day. According to the aluminum industry literature, the industry produces 2,308 kg of aluminum per day using a current supply of 300 kA [22].

The Faraday efficiency for the production of aluminum using the electrolysis process can be calculated using the above values. The current or Faraday efficiency for the aluminum reduction process was 95.5%. Energy consumption for the electrolysis process is defined as the product of the total charge supplied and the voltage maintained in the cell. For the production of aluminum, the energy consumption was calculated to be 13.1 kWh for every gram of aluminum produced.

Battery Capacity

For battery systems, the total current which can be obtained from an electrochemical system in 1 h is termed as capacity. The units for capacity are ampere-hour (Ah). The theoretical capacity of a battery is dependent on the active material available as electrode with the internal resistance of the electrolyte not being taken into consideration. Using Faraday's law, the theoretical capacity for the electrode material is given as

$$C_e = 26.8 \frac{\mathrm{A-h}}{\mathrm{eq-g}} \frac{1}{M_e} \tag{38.18}$$

where the electrode material capacity is C_e (ampere-hour per gram) and M_e is the equivalent weight of the electrode material (grams per equivalent gram). The equivalent weight is represented by the molecular weight of the electrode material divided by the number of electrons transferred during the reaction [13].

For example, the silver oxide-zinc battery is composed of zinc as active material in the anode electrode and silver oxide is the active material in the cathode. The overall reaction for the battery is given in (38.16).

$$\begin{aligned} &\mathrm{Zn} + \mathrm{Ag_2O} + \mathrm{H_2O} \rightarrow \mathrm{Zn(OH)_2} + 2\mathrm{Ag} \\ &E = 1.591\ \mathrm{V} \end{aligned} \tag{38.16}$$

The equivalent weight for Zn (M_{Zn}) is 32.68 g/eq−g and the cathode active material, silver oxide, has an equivalent weight (M_{Ag_2O}) of 116 g/eq−g. Using (38.18), the theoretical capacity for anode electrode of the silver oxide-zinc battery can be calculated as $C_{Zn} = 820$ mAh/g and the cathode material has the theoretical capacity of $C_{Ag_2O} = 231$ mAh/g.

Once the theoretical capacity of the electrode material is calculated, then the total capacity of the electrode can be obtained from (38.19).

$$C_T = C_e W_e \tag{38.19}$$

where C_T is the total capacity for the electrode (ampere-hour) and W_e is the weight of the electrode material (grams). The total capacity of both the electrodes will be calculated using the above equations; the lowest capacity is taken as the theoretical capacity of the battery system. If the weight of active material for the silver oxide-zinc battery is assumed to be 5 g, then the total capacity for Zn is $C_T = 4{,}100$ mAh and for Ag_2O is $C_T = 1{,}155$ mAh. Since the theoretical capacity of the cathode material, silver oxide, is the limiting value, then the theoretical capacity for the silver oxide-zinc battery will be 231 mAh/g.

There is another equation used to determine the theoretical capacity of the battery system, which is based on the total quantity of electricity used in both the anode and cathode electrode [20].

$$C_T = \frac{1}{1/C_{e,\mathrm{negative}} + 1/C_{e,\mathrm{positive}}} \tag{38.20}$$

On using the equation for the silver oxide-zinc battery system, the theoretical capacity of this battery was found to be 180 mAh/g.

Energy Density

Energy density of a battery is determined by taking the product of the cell voltage and capacity of the battery. The unit for the energy density is represented by watt-hour per gram (Wh/g). Energy density is also an important property as the capacity for the battery. The theoretical energy density for a battery system is represented as: [20]

$$\begin{aligned} &\text{Theoretical energy density(Wh/g)} = \text{Theoretical voltage(V)} \\ &\quad \times \text{Theoretical capacity(Ah/g)} \end{aligned} \tag{38.21}$$

In the case of the silver oxide-zinc battery, the theoretical voltage is 1.591 V and the theoretical capacity determined in the previous subsection was 180 mAh/g. Using these values in the above equation, we obtain the theoretical energy density for the silver oxide-zinc battery as 0.286 Wh/g or 286 Wh/kg.

Electrokinetics

In the electrochemical systems, the rate at which reactions take place on the electrode surface depends on different properties such as electrode material, concentration of ions present in the electrolyte, mobility of the ions, and the concentration of ions at the electrode surface. At equilibrium, the ions in the electrolyte will cover the electrode surface based on their charge and form a layer known as double layer. The double layer is described as the combination of adsorbed ionic species and the layer of solvated ions over the surface of metal. Figure 38.2 shows the illustration of the metal/solution interface based on the geometric model. There are three established models, namely Helmholtz theory, Gouy-Chapman theory, and Stern theory, to elucidate the formation of the double layer over the metal surface.

The metal electrode immersed in an electrolyte is covered by absorbed ions due to the excess charge on the metal surface. An imaginary line running through the centers of the adsorbed ions form the inner Helmholtz layer or plane. The outer Helmholtz layer or plane is the locus of centers of the nearest solvated ions. Beyond the outer Helmholtz layer, there are nonspecifically adsorbed ions which form a diffuse layer. The double layer in the electrolyte solution is composed of the inner and outer Helmholtz layers, and the diffuse layer. The thickness and composition of double layer will determine the surface overpotential. For ionic concentrations more than 0.01 M, the double layer thickness is expected to be less than ~100 Å [11]. Surface overpotential (η_s) is expressed as the potential of the working electrode with respect to the reference electrode, which is placed just outside of the double layer. The term overpotential is defined as the difference between the measured potential and the equilibrium potential for that reaction.

Electrochemical reaction rates are defined in terms of the current density of the reactions, the concentration of ions in

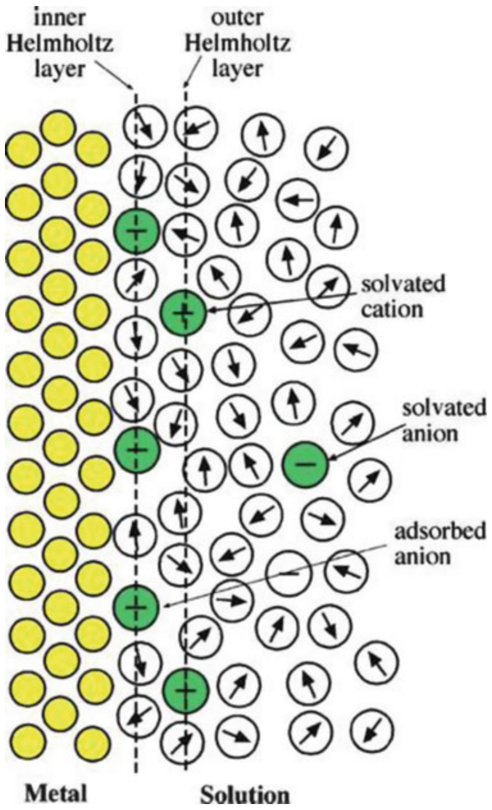


Fig. 38.2 Illustration of metal/solution interface to depict the presence of double layer [23]. The double layer is composed of adsorbed and solvated ions over the metal surface. The double layer has many layers; among those are the inner and outer Helmholtz layer

the electrolyte, and the overpotential. A classical way to relate these parameters is using the Butler-Volmer equation:

$$i = i_o \left[e^{((\alpha_a F/RT)\eta_s)} - e^{(-(\alpha_c F/RT)\eta_s)} \right] \tag{38.22}$$

where i_o is the exchange current (A), α_a and α_c are the transfer coefficients for the anodic and cathodic reactions, respectively, i is the current for the reaction (A), and η_s is the surface overpotential for the electrode.

For electrochemical reactions which are irreversible in nature, an approximation of the Butler-Volmer equation will lead to the Tafle equations (38.23) and (38.24). In the case where the anodic reactions are more applicable than cathodic reactions, then $a_a F\eta_s \gg RT$ and the Butler-Volmer equation is reduced to

$$i = i_o e^{((\alpha_a F/RT)\eta_s)} \tag{38.23}$$

Similarly, if the cathodic reactions are more applicable than the anodic reactions, then $a_c F\eta_s \gg RT$ and the Butler-Volmer equation is modified to

$$i = -i_o e^{(-(\alpha_c F/RT)\eta_s)} \tag{38.24}$$

Table 38.5 Various dielectric materials and their constants

Dielectric material	Dielectric constant (ε)
Air (dry, CO_2 free)	1.0005360
Aluminum oxide	9.34–11.54
Glass (Pyrex 7740)	5.0
Lead magnesium niobate	10,000
Lead titanate	~200
Mica (muscovite)	5.4
Polypropylene	2.3
Polystyrene	2.6
Potassium tantalate niobate (0°C)	34,000
Quartz (fused)	3.75
Tantalum pentoxide (α phase)	30–65
Teflon	2.1
Vacuum	1.000 (exact)
Water (liquid, 20°C)	80.1

Source: Haynes [19]

Capacitor

An electronic device, which is capable of storing electric charge, is defined as a capacitor. Capacitors can be used in place of batteries to store electrical energy and can also discharge the electrical energy during a relatively short duration. This device has two electrically conducting electrodes, which are separated by a nonconducting material. The nonconducting material is known as a dielectric, which is made of an insulator material.

The capacitor is classified as polarized and non-polarized based on the characteristic of the dielectric material. The non-polarized capacitors have ceramics and polymers as the dielectric materials, whereas oxides of aluminum and tantalum are used in the polarized capacitors. Polarized capacitors are generally referred to as electrolytic capacitors, which are electrochemical devices [23]. Each dielectric material has a characteristic value known as dielectric constant. Vacuum is dielectric and it has been used in the old-fashioned capacitors. The dielectric constant for vacuum is assigned a value of unity. Technically, the dielectric constant of a material is the ratio of the relative permittivity of the material to store the electrical energy, when voltage is applied, with respect to the storage capability under vacuum. Various dielectric materials and their constants are listed in Table 38.5.

Capacitance is the measure of the capacitor's ability to store the electrical energy for a certain potential difference across the electrodes. When a voltage is applied between the electrodes, the positive charges are collected in one electrode and the negative charges will be generated on the other electrode.

$$C = \frac{Q}{V} \tag{38.25}$$

Table 38.6 Comparison of the characteristics of few primary batteries

Characteristics	Zinc chloride (zinc-carbon)	Alkaline manganese dioxide	Zinc-air
Anode	Zn	Zn	Zn
Cathode	MnO_2	MnO_2	O_2 (air)
Electrolyte	$ZnCl_2$ (aqueous solution)	KOH (aqueous solution)	KOH (aqueous solution)
Open-circuit voltage (V)	1.6	1.5–1.6	1.45
Nominal voltage (V)	1.5	1.5	1.5
End voltage (V)	0.9	0.9	0.9
Operating temperature (°C)	−10 to 50	−20 to 55	0–50
Energy density at 20°C: button size (Wh/kg)	–	38	340
Energy density at 20°C: cylindrical size (Wh/kg)	85	125	–
Power density	Low to moderate	Moderate	Low
Self-discharge rate at 20°C (% loss per year)	7	4	3 (if sealed)

Source: Linden [20]

The ratio of electric charge divided by the potential difference is referred to as the capacitance of the capacitor. The units for the capacitance are Farads (*F*). Applying a potential difference of 1 V across the electrode plates when 1 coulomb of charge is stored on the electrodes is referred to a capacitance of 1 farad for that capacitor. Capacitors for typical industrial use are manufactured in the range of μF to mF.

Electrochemical Energy Storage Technologies

Classical electrochemical energy storage technologies include batteries, flow batteries, and fuel cells. This section provides an overview of the different technologies; additional literature is recommended [13, 20, 24–32]. In addition, this section also includes a synopsis of super capacitors or electrochemical double layer capacitors (EDLCs), which could be considered advanced electrochemical energy storage systems.

Batteries

The most commonly known electrochemical energy storage device is a battery, as it finds applications in all kinds of instruments, devices, and emergency equipment. A battery's principal use is to provide immediate power or energy on demand. A battery is an electrochemical device where energy from a chemical reaction of the reactants is directly converted into electrical energy. There are two types of batteries: primary and secondary. In primary batteries, the chemical compounds as reactants are consumed to form electrical energy and products, but the products cannot be transformed back into reactants with the use of external electric power. On the other hand, secondary batteries can be recharged by producing the reactants from the products with the help of external electrical energy. Secondary batteries are also known as rechargeable batteries.

The batteries that come under the primary batteries are zinc chloride, alkaline (zinc-manganese dioxide), and zinc-air. Secondary batteries consist of LAB, nickel-cadmium (Ni-Cd), nickel metal hydride (Ni-MH), and lithium ion (Li-ion) battery. Other kinds of batteries, such as a flow battery, redox battery, and sodium nickel chloride battery, are the latest improvements in the area of batteries. In batteries, the convention is that the anode is considered to be the negative terminal and the cathode is considered to be the positive terminal. The electrochemical reactions of these batteries at both the positive and negative terminals, along with a brief description of their properties, are presented next.

Primary Batteries

The characteristics of the few primary batteries that are discussed in this section are listed in Table 38.6. The table provides information on the electrodes, electrolyte, cell voltage, operating temperature, energy and power density, self-discharge rate, advantages, and limitations for each battery.

Zinc-Chloride

This battery is one of the first kinds to be commercially sold, and it is an improved version of the zinc-carbon battery. A can, or the container of the battery that is made of Zn, will act as the negative terminal (anode) of the battery. A carbon rod is the positive terminal (cathode), which is covered by a manganese dioxide paste. Next to the manganese dioxide paste is a layer of zinc chloride electrolyte, which is separated by a paper separator. There are two types of separators used in the Zn-chloride batteries, gelled paste and a special paper coated with gelled paste or cereal. The electrochemical reactions for zinc chloride battery are given below:

Anode rection

$$Zn_{(s)} \rightarrow Zn^{2+}{}_{(aq)} + 2e^- \quad E^\circ = -0.76 \text{ V vs. SHE} \tag{38.26}$$

Cathode reaction

$$MnO_{2(s)} + H_2O_{(l)} + e^- \rightarrow MnO(OH)_{(s)} + OH^-{}_{(aq)} \\ E^\circ = 0.36 \text{ V vs. SHE} \tag{38.27}$$

Overall reaction

$$Zn_{(s)} + 2MnO_{2(s)} + ZnCl_{2(aq)} + 2H_2O_{(l)} \\ \rightarrow 2MnO(OH)_{(s)} + 2Zn(OH)Cl_{(aq)} \quad E = 1.1 \text{ V} \tag{38.28}$$

At the anode electrode, Zn is oxidized to Zn^{2+} ions and manganese dioxide is reduced to manganese oxyhydroxide (MnO(OH)) at the cathode electrode. The cell voltage for the overall reaction is 1.5 V.

The zinc chloride batteries are used in flashlights, penlight, photoflash, smoke detectors, and electronic applications. Commercially, these batteries are manufactured in cylindrical form in different sizes, such as AA, C, D, and F. Household flashlights use sizes C and D Zn-chloride batteries, whereas size AA is normally reserved for penlight, photoflash, and electronic applications [20].

One of the limitations of the Zn-chloride batteries is its high self-discharge rate which means that the Zn-chloride battery will lose its capacity even if it were left idle. Additionally, these batteries are known to have high gassing rates; that is, the Zn-chloride batteries produce more gas from the secondary reactions due to the presence of contamination in the Zn anode [20]. In general, the Zn-chloride battery is not able to keep up with the growing electronic industry because the Zn-chloride battery has lower power density due to the small size and the lack of rechargeable capabilities.

Alkaline Batteries

The electrode material used for the alkaline battery is similar to the Zn-chloride battery with the exception of the potassium hydroxide (KOH) electrolyte. The anode electrode for the alkaline battery is zinc powder, which is mixed with KOH to form a gel. The cathode electrode is manganese dioxide. The following equations illustrates the electrochemical reactions taking place in an alkaline battery: [33]
Anode reaction

$$Zn + 2OH^- \rightarrow Zn(OH)_2 + 2e^- \\ E^\circ = -1.25 \text{ V vs. SHE} \tag{38.29}$$

Cathode reaction

$$MnO_2 + H_2O + e^- \rightarrow MnOOH + OH^- \\ E^\circ = 0.36 \text{ V vs. SHE} \tag{38.30}$$

Overall reaction

$$Zn + 2MnO_2 + 2H_2O \rightarrow Zn(OH)_2 + MnOOH \\ E = 1.61 \text{ V} \tag{38.31}$$

The theoretical cell voltage for an alkaline battery is 1.61 V, but the optimum voltage will be closer to the zinc-based primary batteries (1.5 V). The purity of the electrode material and its activity determines the initial open-circuit voltage of the alkaline battery. These aspects lead to the lowering of the optimum open-circuit voltage of the alkaline battery to 1.5 V from the theoretical cell voltage (1.61 V).

Alkaline batteries have higher energy density and longer operation time in comparison with the inexpensive zinc-carbon and zinc chloride batteries. These batteries are popular for usage in various applications because of their relatively low cost, higher energy and power density, and good portability for providing power to portable electronic devices. Alkaline batteries are manufactured in both cylindrical and miniature sizes.

The alkaline batteries experience similar kinds of limitations as those of zinc chloride batteries. The corrosion of the Zn anode material is also prevalent in the alkaline battery, which essentially lowers the battery capacity and eventually degrades the life of the battery. Zinc in the environment of an alkaline electrolyte is active and it can reduce water into hydrogen gas. The presence of hydrogen gas will increase the internal pressure of the battery, and can cause leakage of the electrolyte and even possible explosion of the battery [34]. In addition, these alkaline batteries are the primary batteries which cannot be recharged to extend their life. In recent years, various research activities have been focused to convert the alkaline battery into a rechargeable battery by the addition of different active materials to the alkaline manganese dioxide cathode [35–38].

Zinc-Air

Zinc-air batteries are well known as the batteries used in hearing aid devices. A cross-sectional view of a Zn-air battery is displayed in Fig. 38.3. The voltage for this battery is produced from the oxidation of zinc in air. As the oxygen from the air is the active material on the cathode side, the cathode can (Fig. 38.3) has holes for the air to flow into the Zn-air battery. The anode electrode is made up of porous zinc powder which is mixed with KOH electrolyte. Air is dispersed on to the cathode electrode with the help of an air distribution membrane and air diffusion Teflon layer, as shown in Fig. 38.3. The current collector for the cathode is the carbon-pressed nickel-plated screen, which is separated from the anode material by a separator. The oxygen from the dispersed air is converted into hydroxide ions on the cathode electrode (38.33) and the hydroxide ion migrates to the

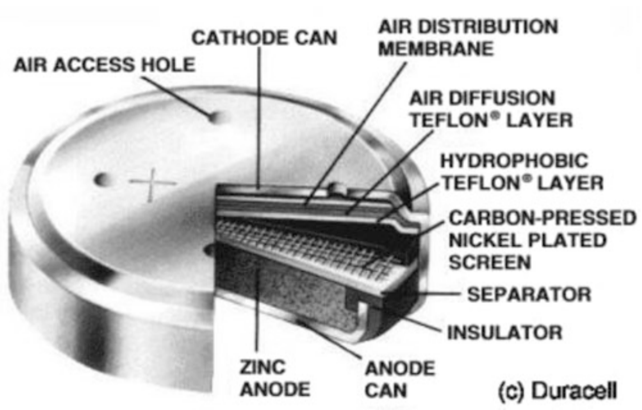


Fig. 38.3 Illustration of Zinc-air button cell [20]. The Zn-air battery has holes for air to pass through the cathode can. The oxygen present in the air is the active material for the cathode side and the Zn is the active material in the anode side. The spacing on the cathode side is smaller than any other battery

anode side to react with Zn. In this reaction, water or electrolyte is not consumed; only the Zn paste is consumed, which implies that replacement of the Zn paste could extend the operation of Zn-air batteries. A sectional view of zinc-air battery is displayed in Fig. 38.3.

The electrochemical reactions at the anode and the cathode electrodes of the Zn-air battery are listed below [33].
Anode reaction

$$\begin{array}{c} \mathrm{Zn} + 2\mathrm{OH}^- \rightarrow \mathrm{ZnO} + \mathrm{H_2O} + 2\mathrm{e}^- \\ E^\circ = -1.25\ \mathrm{V\ vs.\ SHE} \end{array} \tag{38.32}$$

Cathode reaction

$$\begin{array}{c} \mathrm{O_2} + 2\mathrm{H_2O} + 4\mathrm{e}^- \rightarrow 4\mathrm{OH}^- \\ E^\circ = 0.40\ \mathrm{V\ vs.\ SHE} \end{array} \tag{38.33}$$

Overall reaction

$$\mathrm{Zn} + \mathrm{O_2} \rightarrow 2\mathrm{ZnO} \quad E = 1.65\ \mathrm{V} \tag{38.34}$$

Theoretically, a zinc-air battery will produce 1.65 V, and it has the highest specific energy density (340 Wh/kg) of all of the primary batteries [20, 39]. The high specific energy density, flat discharge voltage, long shelf life, environmentally safe nature, and low cost of the Zn-air batteries make them attractive storage devices for many applications [20, 40]. Even though the initial application of Zn-air batteries was restricted to button cells for hearing aids, the recent developments on Zn-air batteries have found applications in cell phones, heavy equipments, and electric vehicles [41].

Zn-air batteries are limited by the amount of zinc present as active material in the anode electrode. The shelf life of the Zn-air battery will be affected by the self-discharge of the battery, meaning the corrosion of zinc will degrade the battery capacity. Other factors affecting the capacity or shelf life of a Zn-air battery are based on the gas transfer mechanism which takes place on the cathode side. The issues associated with gas transfer mechanism are direct oxidation, carbonation of electrolyte, and effect of water vapor transfer [20]. The process of the zinc anode's being directly oxidized by the oxygen that has passed into the cathode side after diffusing through the electrolyte is known as the direct oxidation mechanism. Carbonation of the electrolyte is observed when the alkaline electrolyte reacts with the carbon dioxide present in the atmosphere, and the electrolyte transforms into metal carbonate and bicarbonate. The degradation of the Zn-air battery is caused by the formation of carbonate crystals on the cathode structure which affects cathode performance. The presence of carbonate in the electrolyte will also increase the vapor pressure and enhance the loss of water vapor. Finally, any deregulation in the transfer of water vapor between the electrolyte and the environment will significantly affect the shelf life of the Zn-air battery.

Secondary Batteries

The secondary batteries that are described in this section are the LAB, nickel-cadmium, nickel metal hydride, and Li-ion batteries. The characteristics of the selected secondary batteries are provided in Table 38.7.

Lead Acid Battery

The LAB, which is based on electrochemical principles, was first invented in 1860 by Planté. Over the years, this battery has been thoroughly studied, researched, and developed. In the category of medium to large sized batteries, the LAB is the most commonly manufactured rechargeable battery. The negative electrode for this battery is a lead paste covered grid, and the positive electrode is the plate or grid which has lead dioxide. In secondary or rechargeable batteries, an electrode which behaves as an anode during the discharge process will act as cathode during the charging process because of the reduction reactions taking place at this electrode. The electrodes are therefore referred to as negative and positive terminals instead of as anode and cathode. The grid is made of a lead-antimony alloy and the electrolyte for the LAB is sulfuric acid. The discharging and charging reactions for both the terminals are given below: [42]
Negative terminal reactions

$$\begin{array}{l} \mathrm{Pb} + \mathrm{HSO_4}^- \xrightleftharpoons[\text{Charge}]{\text{Discharge}} \mathrm{PbSO_4} + \mathrm{H}^+ \\ +2\mathrm{e}^- \\ E^\circ = -0.358\ \mathrm{V\ vs.\ SHE} \end{array} \tag{38.35}$$

Table 38.7 Comparison of the characteristics of some secondary batteries

Characteristics	Lead acid (SLI design)	Nickel-cadmium (sealed design)	Nickel metal hydride	Lithium ion ($C/LiCoO_2$ system)
Anode	Pb	Cd	MH	C
Cathode	PbO_2	NiOOH	NiOOH	$LiCoO_2$
Electrolyte	H_2SO_4 (aqueous solution)	KOH (aqueous solution)	KOH (aqueous solution)	Organic solvent[a]
Open-circuit voltage (V)	2.1	1.29	1.4	4.2
Nominal voltage (V)	2.0	1.2	1.2	4.0
Operating voltage (V)	2.0–1.8	1.25–1.00	1.25–1.10	4.0–2.5
End voltage (V)	1.75	1.0	1.0	2.5
Operating temperature (°C)	−40 to 55	−40 to 45	−20 to 50	−20 to 55
Energy density at 20°C (Wh/kg)	35	30–35	50	90
Power density	High	Moderate to high	Moderate to high	Moderate
Self-discharge rate at 20°C (% loss per month)	2–3[b]	15–20	20	5–10
Cycle life	200–700	300–700	300–600	500–1,000

Source: Linden [20]

[a] Organic solvents such as propylene carbonate, ethylene carbonate, ethyl methyl carbonate, and so on

[b] Maintenance-free design

Positive terminal reactions

$$PbO_2 + HSO_4^{-} + 3H^{-} + 2e^{-} \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} PbSO_4 + 2H_2O$$

$$E^\circ = 1.69\ \text{V vs. SHE} \tag{38.36}$$

Overall reactions

$$Pb + PbO_2 + 2H_2SO_4 \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} 2PbSO_4 + 2H_2O$$

$$E = 2.048\ \text{V} \tag{38.37}$$

During the discharge process, the electrolyte (sulfuric acid) reacts with lead and lead dioxide to form lead sulfate and water. The cell voltage observed for an LAB is slightly more than 2V, and its specific energy density ranges between 25 and 35 Wh/kg [25]. The density of the sulfuric acid solution is often measured to estimate the status of charge for the battery. A fully charged LAB should possess 1.27–1.30 relative density for the electrolyte.

Since the LAB is a rechargeable battery, the external electricity is used to convert $PbSO_4$ to PbO_2 and Pb at the positive and negative terminals, respectively. This charging process also results, however, in secondary reactions, such as the evolution of oxygen at the positive terminal and the liberation of hydrogen gas at the negative terminal. These secondary reactions are the main cause for the loss of water in the electrolyte. As a result, the lead acid batteries have to be periodically replenished with water.

Research studies indicate that the presence of antimony on the lead grid electrodes was the cause for the secondary reaction which includes water electrolysis and that resulted in loss of water from the electrolyte. In order to avoid periodic replenishment of water and to retain a low maintenance LAB, the chemical composition of the electrode was changed to lead-calcium or lead-calcium-tin alloy instead of lead-antimony. The amount of water lost from the electrolyte in the low maintenance LAB is less than that lost in traditional lead acid batteries. Recent advances in lead acid batteries have focused on developing completely sealed systems. The new LAB is called a valve-regulated lead acid (VRLA) battery. In a VRLA battery, the oxygen gas generated at the positive terminal is transported internally to the negative terminal, where the oxygen is reduced with protons to form water. This process is known as the oxygen cycle [25]. The advantage of the VRLA battery is that no maintenance is required, and no gases (H_2 and O_2) escape from the battery system.

The lead acid batteries have been mainly used for automotive applications. Those batteries which are used for automotive purposes are referred to as an SLI (starting, lighting, and ignition) system. The SLI system of lead acid batteries finds application in automotive vehicles, golf carts, tractors, and lawn mowers. Additionally, the LABs are used in industrial trucks and electric vehicles for traction or motive power. The VRLA battery has been used for stationary applications, such as emergency and standby power systems, telecommunications, uninterrupted power supply (UPS), railroad signaling, and energy storage systems. The advancement in the LAB research has lead to portable systems which are sealed and maintenance-free systems. The portable lead acid batteries are mainly used in small appliances and devices, portable tools, and electronic equipments [20].

The generation of hydrogen and oxygen gas during the operation of the LAB limits its shelf life and capacity. These gases are released at levels which can lead to explosions, as in the case of SLI type lead acid batteries. Even in the sealed lead acid batteries, certain amounts of hydrogen gas are released during its operation; therefore, care must be taken while planning for charging and location of the battery for safety reasons. The lead acid batteries fail mostly due to the corrosion of the positive and negative grids or terminal. Lead and lead oxide, the active materials in negative and positive terminals, respectively, will react with sulfuric acid even if the battery is in idle condition. Therefore, grid corrosion is a major problem for the failure of lead acid batteries [20].

Nickel-Cadmium

The nickel-cadmium (Ni-Cd) battery is one of the rechargeable alkaline batteries because it uses KOH as its electrolyte. Ni-Cd battery was invented in the early twentieth century by the Swedish scientist Waldemar Jüngner. Almost at the same time, Thomas Edison in the USA invented another kind of nickel-based battery known as the nickel-iron battery, which also had similar electrochemical properties to those of the Ni-Cd battery [42, 43].

The negative terminal of an Ni-Cd battery has cadmium as the active material, whereas nickel oxyhydroxide (NiOOH) is the active material in the positive terminal. During the discharge process, cadmium, in the negative terminal, is oxidized by the hydroxide ion present in the electrolyte to form cadmium hydroxide. Nickel oxyhydroxide, in the positive terminal, is converted into nickel hydroxide. During the charging process, cadmium hydroxide is reduced to metallic cadmium at the negative terminal, and Ni^{2+} in the nickel hydroxide is converted into the Ni^{3+} species as nickel oxyhydroxide at the positive terminal [44, 45].

Negative terminal reactions

$$\mathrm{Cd} + 2\mathrm{OH}^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} \mathrm{Cd(OH)_2} + 2\mathrm{e}^- \qquad E^\circ = -0.80\ \mathrm{V\ vs.\ SHE} \tag{38.38}$$

Positive terminal reactions

$$2\mathrm{NiOOH} + 2\mathrm{H_2O} + 2\mathrm{e}^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} 2\mathrm{Ni(OH)_2} + 2\mathrm{OH}^- \qquad E^\circ = 0.49\ \mathrm{V\ vs.\ SHE} \tag{38.39}$$

Overall reactions

$$2\,\mathrm{NiOOH} + \mathrm{Cd} + 2\mathrm{H_2O} \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} 2\mathrm{Ni(OH)_2} + \mathrm{Cd(OH)_2} \qquad E = 1.29\ \mathrm{V} \tag{38.40}$$

Thermodynamically, the cell voltage for the Ni-Cd battery is 1.30 V. However, in practice, the voltage output is 1.20 V. Over the years, the specific energy density for Ni-Cd has improved from 30 to 40 Wh/kg to 60 Wh/kg [42]. Overcharging of the Ni-Cd battery will produce oxygen gas on the positive terminal (38.41). Since the positive terminal always gets fully charged before the negative terminal charges, generation of oxygen gas is a frequent occurrence at the positive terminal. Oxygen formed at the positive terminal will be transferred to the negative terminal with the help of oxygen permeable separator, for recombination with Cd to form cadmium hydroxide (38.42) [45]. Finally, the cadmium hydroxide will be converted into metallic cadmium during the overcharging process (38.43) [29]. These reactions are similar to the oxygen cycle found in the low maintenance and VRLA batteries.

$$4\mathrm{OH}^- \rightarrow 2\mathrm{H_2O} + \mathrm{O_2} + 4\mathrm{e}^- \tag{38.41}$$

$$\mathrm{Cd} + \frac{1}{2}\mathrm{O_2} + \mathrm{H_2O} \xrightarrow{\text{Gas Recombination}} \mathrm{Cd(OH)_2} \tag{38.42}$$

$$\mathrm{Cd(OH)_2} + 2\mathrm{e}^- \xrightarrow{\text{Charge}} \mathrm{Cd} + 2\mathrm{OH}^- \tag{38.43}$$

Sealed Ni-Cd battery was developed to avoid gassing (oxygen) from the positive terminal. Until nickel-metal hydride batteries came along in 1991, the Ni-Cd batteries were the main choice as the power supply for all cordless or portable electronic equipment [44]. The longer cycle life of the Ni-Cd battery has resulted in its extensive use in the spacecraft applications, such as satellites in low earth orbit (LEO) and geo-synchronous earth orbit (GEO) [45].

The limitations of using the Ni-Cd batteries have been associated with the generation of gases on the electrodes during the charge–discharge process. The operation of the Ni-Cd batteries is affected by the generation of hydrogen gas on the positive terminal due to over discharging. Even in a sealed Ni-Cd battery, the evolved hydrogen gas is consumed at a low rate, which leads to a buildup of internal pressure and eventually bursting of the Ni-Cd battery [45]. At the same time, during the charging process of a sealed Ni-Cd battery, the water formed at the positive terminal will dilute the electrolyte concentration on the pores of the active material, thereby causing a drop in electrode voltage or overvoltage. While operating the Ni-Cd battery in a low temperature environment with high rate of charging and discharging, the formation of water at the pores of the positive terminal will lead to the dilution of the electrolyte and freezing. However, on the negative terminal the electrolyte concentration will increase resulting in precipitations on the pores [45]. In addition to these issues, the Ni-Cd battery has the inherent problem of using a carcinogenic material, cadmium, which is also a pollutant that will harm the environment.

Nickel Metal Hydride

The Ni-MH battery is similar to the nickel-cadmium battery in most aspects, except for the active material in the negative terminal. The active material in the negative terminal is hydrogen, which is stored in the form of metal hydride. The hydrogen atoms are stored in metal alloys and these alloys are of two kinds, AB_5-type and AB_2-type. The AB_5-type alloy is made up of nickel and rare earth metals, whereas the AB_2-type alloy is a mixture of nickel, titanium, vanadium, and zirconium [45]. Typically, the negative terminal electrode is made of either $LaNi_5$ (AB_5-type) or $TiNi_2$ (AB_2-type) alloy and the AB_5-type alloys have better corrosion resistance than the AB_2-type [25]. In recent commercial development of the Ni-MH battery, the alloy composition for the negative electrode has been $MmNi_{3.2}Co_{1.0}Mn_{0.6}\,Al_{0.11}Mo_{0.9}$ (AB_5-type) or $Ti_{0.51}Zr_{0.49}V_{0.70}Ni_{1.18}Cr_{0.12}$ (AB_2-type), where Mm is misch-metal, which is composed of 25 wt. % La, 50 wt. % Ce, 7 wt. % Pr, and 18 wt. % Nd [45].

NiOOH is the active material for the positive terminal and the aqueous KOH is the electrolyte. The reactions at the terminals are listed below: [33]

Negative terminal reactions

$$MH + OH^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} M + H_2O + e^- \qquad E^\circ = -0.83\ \text{V vs. SHE} \tag{38.44}$$

Positive terminal reactions

$$NiOOH + H_2O + e^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} Ni(OH)_2 + OH^- \qquad E^\circ = 0.49\ \text{V vs. SHE} \tag{38.45}$$

Overall reactions

$$NiOOH + MH \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} Ni(OH)_2 + M \qquad E = 1.32\ \text{V} \tag{38.46}$$

The Ni-MH battery has been replacing the Ni-Cd battery in various types of electronic equipment because the specific energy density of the Ni-MH battery is 60–80 Wh/kg, which is higher than that of the Ni-Cd battery. The cell voltage of the Ni-MH battery is similar to that of the Ni-Cd battery, ranging between 1.2 and 1.3 V.

One of the advantages of using the Ni-MH battery over the Ni-Cd battery is the absence of a toxic pollutant (cadmium) in the battery system. The evolution of hydrogen and oxygen gases, even in the sealed Ni-MH battery during overcharging and discharging conditions, has been effectively eliminated by the design of the battery [45]. During the overcharging step, oxygen gas is produced at the positive terminal; the oxygen gas then migrates to the negative terminal, where it reacts with metal hydride electrode to form water (38.47), (38.48), and (38.49):

$$4OH^- \rightarrow O_2 + 2H_2O + 4e^- \tag{38.47}$$

$$M + H_2O + e^- \rightarrow MH + OH^- \tag{38.48}$$

$$4MH + O_2 \rightarrow 4M + 2H_2O \tag{38.49}$$

Similarly, during the deep discharging process, hydrogen gas is evolved from the positive terminal (38.50). The hydrogen gas migrates to the negative terminal, where it is oxidized into water in the presence of the metal hydride electrode (38.51).

$$2H_2O + 2e^- \rightarrow H_2 + 2OH^- \tag{38.50}$$

$$H_2 + 2OH^- \rightarrow 2H_2O + 2e^- \tag{38.51}$$

The Ni-MH batteries do not show any memory effect caused by recrystallization on the electrodes. Moreover, the Ni-MH batteries have higher energy density in terms of weight and volume in comparison with the Ni-Cd batteries, which has resulted in the replacement of the Ni-Cd batteries by the high energy density Ni-MH batteries. In addition to the application of the Ni-MH batteries for all of the electronic devices, the prismatic cells with 250 Ah capacity are manufactured for electric vehicle applications. The Ni-MH batteries are capable of operating only in the temperature range of −20 to +45°C; beyond 45°C, the charging efficiency of the battery has been observed to decline [45].

The swelling of the electrode material from absorption of hydrogen gas and the leaching of the metal ions present in the electrode into the electrolyte are some of the issues related to Ni-MH batteries. The hydrogen storage materials have the problem of severe volume expansion with more intake of hydrogen gas. This excessive absorption of hydrogen gas can occur from the charging—discharging process, and the absorption of hydrogen gas can lead to cracking and

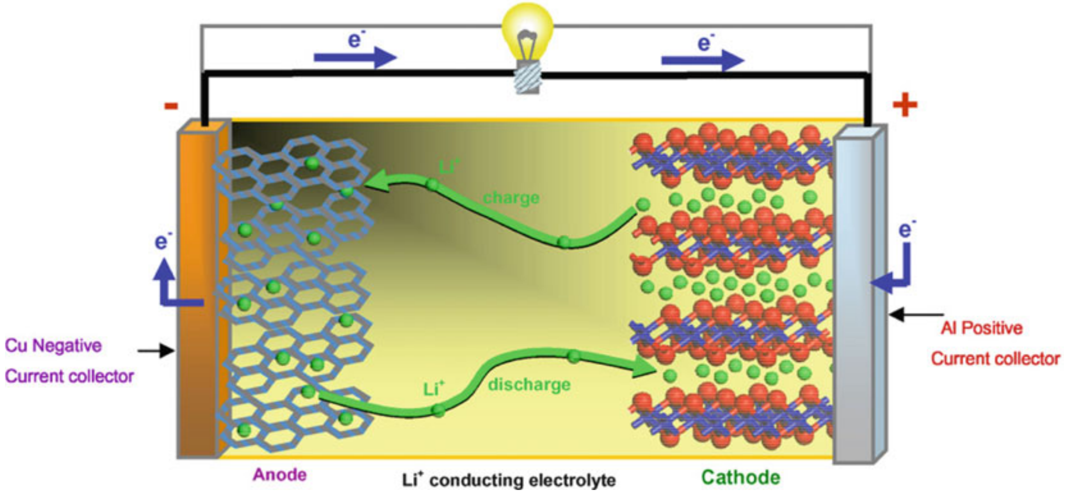


Fig. 38.4 Illustration of discharge and charge processes of Li-ion battery [46]. Copper is the negative current collector and it has graphite on the surface. Aluminum is the positive terminal with $LiCoO_2$ compound on the surface. The intercalation of Li-ions during charging and discharging is shown in this figure

pulverization of alloy present in the electrode [45]. The leaching of the metal ions from the metal alloy electrodes, due to the presence of organic impurities in the electrolyte, will lead to self-discharge of the battery and eventual decay of the battery capacity. The separators used in the Ni-MH battery can also decompose and disperse the organic impurities in the electrolyte, which is another problem related to Ni-MH batteries.

Lithium Ion Battery

Lithium is the lightest element among all the materials used in batteries. The standard reduction potential for Li is −3.04 V vs. SHE, which makes it a suitable component for batteries. Lithium batteries, commonly used in cameras, have an average cell voltage of 3.5 V. Lately, however, another kind of lithium battery, the lithium ion battery (LIB), has demonstrated higher cell voltages in the range of 4 V and a specific energy density of 100–150 Wh/kg, which translates into a longer cycle life than any other commercial battery [27]. So, the LIB finds application in mobile phones, laptops, portable music players, and even as a power supply unit for electric vehicles [43].

The electrochemistry for the LIB is not substantially different from other batteries. Graphite is the active material in the negative terminal and the positive terminal is made of lithium metal oxide compounds, such as $LiCoO_2$ and $LiNiO_2$. The element lithium is highly reactive and can catch fire, or even explode, when it comes in contact with water. So, the electrolyte used in the LIB is a nonaqueous solution, such as a combination of lithium salt ($LiPF_6$) with organic solvents, namely ethylene carbonate-dimethyl carbonate (EC-DMC). The electrochemical reactions for the LIB with $LiCoO_2$ as the active material in the positive terminal are listed below: [33]

Negative terminal reactions

$$\begin{aligned} &Li_xC \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} C + xLi^+ + xe^- \\ &E^\circ = -2.90\,\text{V vs. SHE} \end{aligned} \tag{38.52}$$

Positive terminal reactions

$$\begin{aligned} &Li_{1-x}CoO_2 + xLi^+ + xe^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} LiCoO_2 \\ &E^\circ = 1.20\,\text{V vs. SHE} \end{aligned} \tag{38.53}$$

Overall reactions

$$\begin{aligned} &Li_xC + Li_{1-x}CoO_2 \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} C + LiCoO_2 \\ &E = 4.10\text{ V} \end{aligned} \tag{38.54}$$

During the charging process, the Li-ions present inside the structures of the $LiCoO_2$ compound are removed or withdrawn, and transported via the mixed organic electrolyte into the graphite structure of the negative terminal. This process is reversed during the battery discharge by removing Li-ions from the graphite structure, and it is intercalated into the $LiCoO_2$ structure. The phenomena of Li-ion transported back and forth between anode and cathode material has led to the naming of the Li-ion battery as a “rocking-chair” cell [42]. The current collector used for the anode or the negative terminal (graphite) is copper, and aluminum is the current collector on the cathode or the positive terminal side (Fig. 38.4).

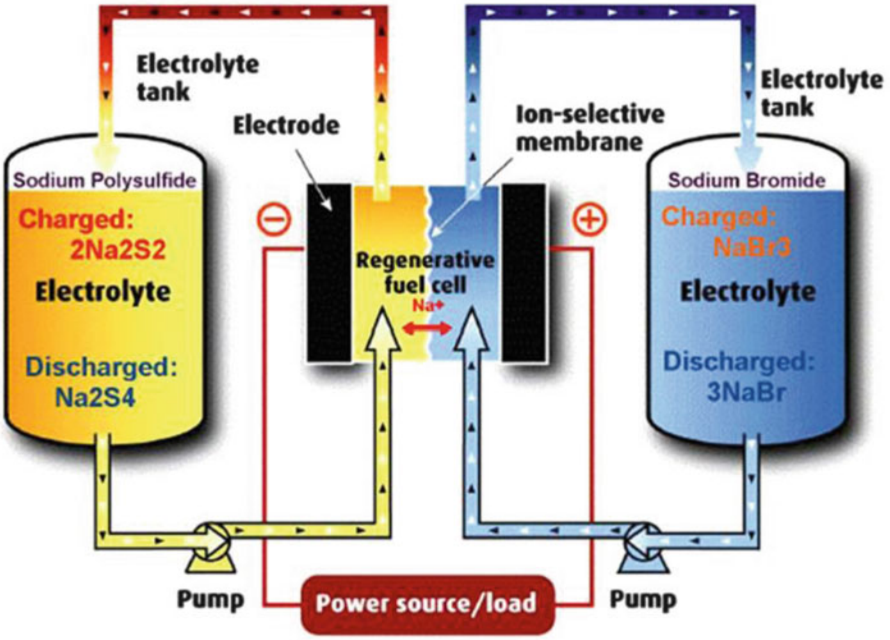


Fig. 38.5 Schematic diagram of polysulfide bromide battery (flow battery) [24]. The flow battery shown in this figure has reservoirs containing charged and discharged species of the compounds for each terminal. Sodium polysulfide solution is circulated on the negative side and sodium bromide is circulated through the positive side of the flow battery

Many researchers are working to improve and develop better Li-ion batteries for various applications. Electrolytes used in the first generation of Li-ion batteries have been liquid organic solvents, but the use of polymer and ceramic electrolytes has simplified the battery design resulting in improved safety and durability [46, 47]. Polymer electrolytes being studied are PEO, a solid polymer electrolyte, and poly(vinylidene fluoride) (PVdF) with hexafluoropropylene (HFP), a polymer gel electrolyte. In the case of ceramic electrolytes, $Li_{1+x}Al_xGe_{2-x}(PO_4)_3$, or LAGP, is considered to have relatively high conductivities [47]. In the same manner, different electrode materials for the LIB are being investigated as well. Anode materials that have caught the researcher's attention are carbon nanotubes (CNTs) with alloys, such as Sn_2Sb and SnNi, and Li alloys, such as LiAl, Li_3Sb, Li_2Si_5, and Li_2Sn_5 [46, 48]. In the case of cathode materials for the Li-ion batteries, $LiMn_2O_4$, vanadium oxides (V_2O_5), olivines ($LiFePO_4$), and layered lithium metal oxides—$LiMO_2$ ($M = Mn$, Ni, Co) have been actively considered [46, 48].

Flow Batteries

Other than the primary and secondary batteries, there are new types of batteries, which constitute the emerging technologies in the field of battery, also being investigated. The research activity surrounding these new types of batteries are focused mainly to improve the energy density of the battery system and also to develop advanced energy storage device capabilities. Among the new kinds of batteries, redox flow and high temperature batteries have demonstrated promising future.

The redox flow batteries are suited as energy storage units for the distributed power supply installations. The outlook for using the redox flow batteries is in stand-alone power applications and distributed energy installations for electric utility service. The salient features of the redox flow batteries are its modular design, ease in transportability, low cost compared to other power generation systems, long life, and high reliability [26]. A few examples of the redox flow battery systems are polysulfide-bromide, zinc-bromine, and vanadium-bromide redox flow cells [24, 49]. Recent advances and further reading on the flow batteries can be found in the literature [50–59].

The design and structure of a redox flow battery, as shown in Fig. 38.5, resemble a fuel cell rather than the traditional rechargeable battery. The functioning of the redox flow battery, however, is different from that of a fuel cell. A simple diagram of the polysulfide bromide flow battery is shown in Fig. 38.5. The positive and negative compartment of the redox flow cell is separated by an ion-selective membrane which provides the ionic conductivity required for the electrochemical reactions at the electrodes. The electroactive species for the positive and negative sides are dissolved in their respective electrolyte solutions. The electrolyte tank for the negative side of the polysulfide bromide battery has both disodium disulfide (Na_2S_2) and disodium tetrasulfide (Na_2S_4), which are the charged and discharged compounds, respectively. Similarly on the positive side, the electrolyte has sodium tribromide ($NaBr_3$) as the charged species and sodium bromide (NaBr) as the discharged species. The presence of charged and discharged species in an electrolyte is not used in the fuel cell system. These electrolyte solutions are circulated into their respective negative or positive side of the flow cell so that the electroactive species will be reduced or oxidized, and produce the electric voltage. The energy of the redox flow battery system depends on the volume of the electrolyte tank, so the energy component of this battery is determined independently from the battery power, which is based on the flow cell.

The electrochemical reactions for the $ZnBr_2$ system are listed below, where bromine gas combined with organic reagent forms polybromide compound in the positive electrode. The thermodynamic cell voltage for the $ZnBr_2$ system is 1.83 V, but only 1.3 V is typically observed during its operation [43].

Negative terminal reactions

$$\mathrm{Zn} \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} \mathrm{Zn}^{2+} + 2e^- \qquad E^\circ = -0.763\ \mathrm{V\ vs.\ SHE} \tag{38.55}$$

Positive terminal reactions

$$\mathrm{Br_2} + 2e^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} \mathrm{Br}^- \qquad E^\circ = 1.065\ \mathrm{V\ vs.\ SHE} \tag{38.56}$$

Another flow battery is based on the vanadium-bromide redox system. Normally, the specific energy density for the vanadium-bromide redox flow battery, for an ion concentration of 2 mol/L, is 25–35 Wh/kg. When the ion concentration of the redox species is increased to 3–4 mol/L, the specific energy density of the flow battery also increases to 60 Wh/kg. The negative electrode of this redox flow battery is the VBr_2/VBr_3 redox couple and the $Cl^-/ClBr_2^-$ redox couple is present at the positive electrode [26]. The reactions for the vanadium-bromine redox flow battery are as follows:

Negative terminal reactions

$$\mathrm{VBr_2} + \mathrm{Br}^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} \mathrm{VBr_3} + e^- \tag{38.57}$$

Positive terminal reactions

$$\mathrm{ClBr_2}^- + 2e^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} 2\mathrm{Br}^- + \mathrm{Cl}^- \tag{38.58}$$

High Temperature Batteries

New types of batteries, such as sodium (Na) beta batteries (NBBs), utilize high temperatures for their effective functioning. These NBBs have beta alumina as their solid electrolyte, in order to handle high temperatures. One of the beta batteries is the sodium-nickel chloride battery, also known as ZEBRA (ZEolite Battery Research Africa), so named because it was first developed in South Africa. The active material in the anode, or the negative terminal, for the ZEBRA battery is molten Na. The active material at the positive terminal is nickel chloride. The primary electrolyte used to separate anode and cathode is β″-alumina. In addition to $\beta''\text{-}Al_2O_3$, the battery also has a secondary electrolyte in the form of molten sodium aluminum chloride ($NaAlCl_4$), which provides good mobility for sodium ions. The sodium-nickel chloride battery is operated between 175 and 400°C. The electrochemical reactions at the electrodes are as follows: [49]

Anode, or negative, terminal reactions

$$2\mathrm{Na} \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} 2\mathrm{Na}^+ + 2e^- \qquad E^\circ = -2.71\ \mathrm{V\ vs.\ SHE} \tag{38.59}$$

Cathode, or positive, terminal reactions

$$\mathrm{NiCl_2} + 2\mathrm{Na}^+ + 2e^- \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} \mathrm{Ni} + 2\mathrm{NaCl} \tag{38.60}$$

Overall reactions

$$2\mathrm{Na} + \mathrm{NiCl_2} \underset{\text{Charge}}{\overset{\text{Discharge}}{\rightleftharpoons}} 2\mathrm{NaCl} + \mathrm{Ni} \qquad E = 2.58\ \mathrm{V\ at\ 300^\circ C} \tag{38.61}$$

Figure 38.6 displays a schematic diagram of the sodium-nickel chloride battery. In this battery, liquid sodium is enclosed in a metal shim and held in an anode end plate. The anode or the negative terminal side is separated from the cathode or the positive terminal side by copper wool. The copper wool separator comes in contact with the solid electrolyte (β″-alumina). In the cathode, which is the positive terminal side, the current collector is an Ni mesh which is surrounded by the active material for the cathode—$NaAlCl_4$. The α-alumina ring is the secondary electrolyte of the sodium-nickel chloride battery.

The sodium beta battery technology based on the beta alumina solid electrolyte has recently attracted interest for applications such as renewable energy storage, as well as

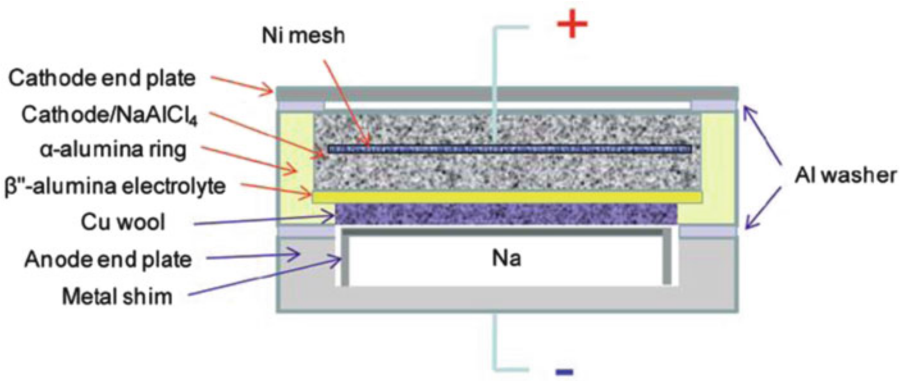


Fig. 38.6 Schematic diagram of a Sodium-nickel chloride battery [140]. Cross-section view of the high temperature battery system, where solid electrolyte is β″-alumina, anode is molten sodium, and molten $NaAlCl_4$ is the cathode material

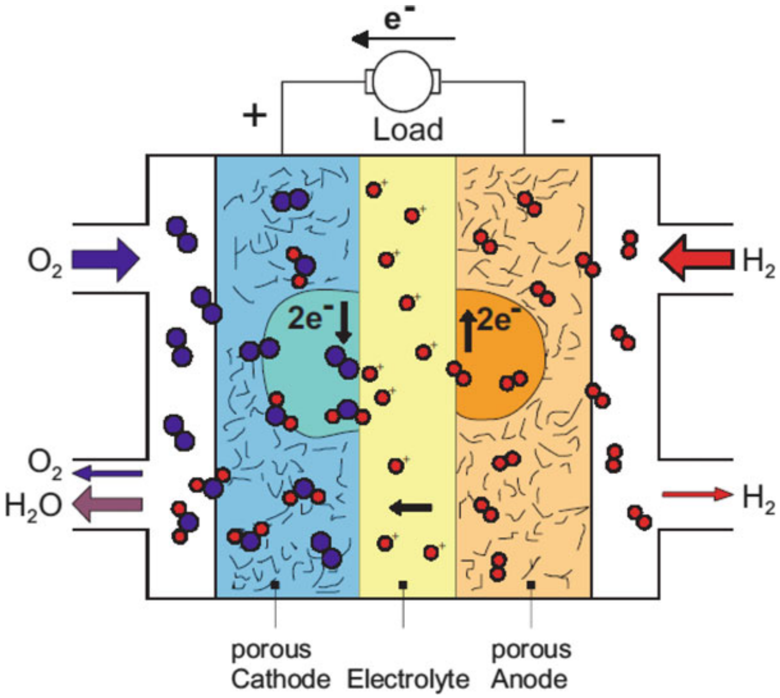


Fig. 38.7 Schematic drawing of a proton exchange membrane fuel cell (PEMFC) [31]. The electrolyte is capable of conducting or exchanging protons between anode and cathode. On the anode side H_2 is oxidized to protons on Pt surface over the porous anode. In the cathode side, the proton combines with O_2 to form water which is removed along with unreacted O_2

electric vehicles, because of its high energy density, high round-trip efficiency, and capacity to store energy for long durations [49].

Fuel Cells

In 1839, Sir William Grove invented an electrical energy conversion system known as fuel cells. The fuel cell is an electrochemical device which can directly convert chemical energy bound up in a chemical compound into electrical energy. Figure 38.7 is the schematic diagram of a proton exchange membrane fuel cell (PEMFC). Similar to the battery system, the fuel cell has anode and cathode compartments. The electrolyte used in the PEMFC is solid polymer electrolyte, Nafion (sulfonated tetrafluoroethylene-based fluoropolymer-copolymer). Hydrogen gas (the fuel) is supplied to the anode side of the fuel cell, and the cathode side needs oxygen or the oxygen present in the air for the electrochemical reaction. The reactants supplied to the fuel cell are in gaseous form. Both the anode and cathode side have a porous electrode with platinum (Pt) as the catalyst for the electrochemical reactions. On the anode side, hydrogen gas is oxidized to become a proton, which travels through the solid electrolyte to the cathode side. On the cathode side, the proton, along with electrons from an external circuit, reacts with oxygen to form water and produce an electric current.

According to Fig. 38.7, the hydrogen gas is fed to the anode side of the PEMFC and the platinum catalyst on the porous anode oxidizes a hydrogen molecule into a proton and two electrons (38.62). The electrons travel through the external circuit and the protons are transported through the proton exchange membrane (PEM), or the electrolyte, to the cathode side. On the cathode side, the protons combine

Table 38.8 Characteristics of the different types of fuel cells

Characteristics	AFC: alkaline fuel cell	PEMFC: proton exchange membrane fuel cell	DMFC: direct methanol fuel cell	PAFC: phosphoric acid fuel cell	MCFC: molten carbonate fuel cell	SOFC: solid oxide fuel cell
Electrolyte	Potassium hydroxide (8–12 N)	Proton exchange membrane	Proton exchange membrane	Phosphoric acid (85–100 %)	Molten carbonate (Li, K, Na)	Solid oxide (ZrO_2–Y_2O_3)
Operating temperature	50–250°C	50–90°C	50–90°C (≤130°C)	180–200°C	650°C	750–1,050°C
Charge carrier	OH^-	H^+	H^+	H^+	CO_3^{2-}	O^{2-}
Electrocatalyst	Pt, Ni/NiO_x	Pt	Pt	Pt	Ni/$LiNiO_x$	Ni/Perovskites
Fuel	H_2	H_2 (pure or reformed)	CH_3OH	H_2 (reformed)	H_2 and CO (reformed), CH_4	H_2 and CO (reformed), CH_4
Poisons	CO, CO_2	CO > 10 ppm	Adsorbed intermediates	CO > 1 % H_2S > 50 ppm	H_2S > 0.5 ppm	H_2S > 1 ppm
Applications	Transportation, space, military, energy storage systems			Combine heat and power for decentralized stationary power systems	Combined heat and power for stationary decentralized systems and for transportation (trains, boats, etc.)	
Realized power	Small plants 5–150 kW Modular	Small plants 5–250 kW Modular	Small plants 5 kW	Small—medium plants 50 kW–11 MW	Small power plants 100 kW–2 MW	Small power plants 100–250 kW

Sources: Srinivasan et al [32].; Carrette et al [31].

with the electrons and the oxygen gas on the porous cathode in the presence of Pt catalyst to form water (38.63) and produce an electric current. The oxygen gas stream is supplied to the cathode side and the water produced from the electrochemical reaction is carried away by the exiting oxygen stream. The electrochemical reactions at the anode and the cathode electrodes of the PEMFC are listed below.

Anode reaction

$$H_2 \rightarrow 2H^+ + 2e^- \quad E^\circ = 0 \text{ V vs. SHE} \tag{38.62}$$

Cathode reaction

$$\frac{1}{2}O_2 + 2H^+ + 2e^- \rightarrow H_2O \quad E^\circ = 1.23 \text{ V vs. SHE} \tag{38.63}$$

Overall reaction

$$\frac{1}{2}O_2 + H_2 \rightarrow H_2O \quad E = 1.23 \text{ V} \tag{38.64}$$

Fuel cells are capable of generating power as long as the fuel (hydrogen) and the oxidant (oxygen) are supplied to the system. On the other hand, even a rechargeable battery is limited in generating power; as all secondary batteries have a definite number of rechargeable cycles, known as cycle life, a rechargeable battery can be recharged (see Table 38.7). In the case of the redox flow battery, the charged and discharged species of the electrochemical reaction are stored in the electrolyte. There are two different electrolytes, one for the anode and one for the cathode. Conversely, the fuel cell has a single electrolyte with the ability to conduct ions between the anode and the cathode. The fuel cell is not designed to be recharged like in the secondary batteries to extend the battery life. A list of different types of fuel cell and its characteristics has been complied in the following Table 38.8.

Notice that traditionally the names for the fuel cells were based on the type of the electrolyte used in that system, except for the Direct Methanol Fuel Cell (DMFC). The DMFC is a PEMFC which uses methanol as its fuel. The most significant problems with adopting fuel cell technologies have to do with: hydrogen storage, transportation, and production. Alternative solutions for these problems will be discussed later in the chapter (see section on “Trends: Novel Electrosynthesis Processes”).

Alkaline Fuel Cell

The alkaline fuel cell (AFC) has the highest electrical efficiency (>60%) among all the types of fuel cells. In the USA, the National Aeronautics and Space Administration (NASA) has used the AFC power plant in the Apollo missions and Space Shuttle program due to its higher efficiency and proven reliability. Initially, nickel-based catalysts were used in the AFCs before switching to a Pt electrocatalyst.

However, one of the advantages of the AFC is that it can employ other metals as the catalyst, such as Pt-Co alloy for oxygen reduction reactions [60]. At the beginning, the AFCs used in the Apollo missions were operated at 200–230°C, whereas now the AFCs are normally operated less than 100°C. The electrochemical reactions at the anode and the cathode electrodes of the AFC are given below:
Anode reaction

$$H_2 + 2OH^- \rightarrow 2H_2O + 2e^- \quad E^\circ = -0.83 \text{ V vs. SHE} \tag{38.65}$$

Cathode reaction

$$\frac{1}{2}O_2 + 2H_2O + 2e^- \rightarrow 2OH^- \quad E^\circ = 0.40 \text{ V vs. SHE} \tag{38.66}$$

The alkaline electrolyte provides better ionic conductivity than acid electrolytes for the fuel cell [32]. Additionally, the fuel cells with alkaline electrolyte, such as KOH used in the AFC, has displayed better kinetics for the oxygen reduction reaction compared to the acid electrolytes [31]. In space applications, the KOH electrolyte for the AFCs is stabilized in an asbestos matrix, which has provided limited life for the fuel cell. In 1970, researchers found that circulating the KOH electrolyte increased the life of the AFC. Carbon dioxide (CO_2) found in the atmosphere and the alkaline electrolyte have been found to be detrimental to the operation of the AFC. CO_2 will react with the hydroxyl ion (OH^-) to form carbonate, which would precipitate on the electrode and the electrolyte, thereby decreasing the performance of the AFC (38.67). In this manner, electrolyte circulation will help in removing the CO_2 gas and reduce the build-up of carbonate in the electrolyte, which extends the life of the AFC. Additionally, a CO_2 scrubber can be used for the hydrogen and air streams in order to reduce the amount of CO_2 in the fuel cell.

$$CO_2 + 2OH^- \rightarrow (CO_3)^{2-} + H_2O \tag{38.67}$$

The AFCs have used in Apollo space craft and other shuttle programs by NASA. Both the USA and Russia have extensively used the AFCs in their space flights [32]. In addition to the space industry, the highly efficient fuel cell (AFC) also found application in the transportation sector. In 1970s, Kordesch built the first fuel cell car using the AFC. In one of the projects, the London cabs were fitted with AFCs modules to provide up to 10 kW power. The cab also had a battery to start and operate the car for the first 10 min before the fuel cell could generate enough power to control the vehicle [31]. Energy storage was the other application of the AFCs. Around the world various organizations, such as Allis-Chalmers, Siemens, Union Carbide, Exxon/Asthom, Fuji, Varta, IFP, CGE, and ELENCO, have installed and tested 1–10 kW capacity of AFC systems [32].

Proton Exchange Membrane Fuel Cell

The PEMFC was formerly known as solid polymer electrolyte membrane fuel cell. These fuel cells use polymer electrolyte as a membrane, which is capable of transporting protons between the anode and the cathode side. The membrane used in these fuel cells is Nafion, which is a perfluorocarbon sulfonic acid membrane. The Nafion membrane is highly acidic in nature with good conductivity for protons in the presence of water. Since Nafion structure is similar to Teflon, it is highly stable in oxidizing and reducing environments [31, 32]. Water management is very important for the efficient functioning of the PEMFC. The gas (hydrogen and oxygen, or air) streams are hydrated before being supplied to the fuel cell, so that ionic conduction of proton across the Nafion membrane can be maintained.

The operating temperature for the PEMFC is in the range of 50–90°C. Pt is the catalyst used in both the anode and cathode electrodes. The electrodes are made of porous gas diffusion electrodes (GDE) so that the gases can have better contact with the Pt catalyst. The electrodes for the PEMFC are prepared in a delicate manner in order to obtain the least resistance between anode, electrolyte, and cathode. This method to combine Pt, catalyst support, binders, and the Nafion membrane is known as membrane electrode assembly (MEA). Dispersion of Pt on carbon support enhances the oxygen reduction reaction on the cathode electrode. If air was used in the cathode side of the fuel cell, then the lower partial pressure of oxygen in the air will decrease the performance of the fuel cell.

In the anode electrode, Pt is the best catalyst for oxidizing pure hydrogen gas. However, the presence of carbon monoxide (CO) in the hydrogen stream will have damaging effect on the electrode and also the PEMFC. Carbon monoxide (CO) will adhere to Pt and block the active sites from the desired electrochemical reaction, evidencing the fact that CO is the major poison for the PEMFC. Cleaning the hydrogen stream off of the CO is one of the ways to avoid CO contamination for the PEMFC. Research on developing CO tolerant catalysts, such as Pt-Ru, is also being investigated [61–65].

Even though the PEMFCs are famously known for its application in cars or for transportation, it has also been used in the Gemini spacecraft. The PEMFC provided a power of 1 kW to the spacecraft [31]. Some of the well-known companies involved in the production of PEMFCs are Ballard Power Systems, Inc. in Canada; IFC, Energy Partners, General Motors, Plug Power, and H-Power is United States; Toyota, Honda, Mazda, and Fuji in Japan;

Daimler Chrysler in association with Ballard Power Systems, Inc. in Germany; and DeNora in Italy [32]. The PEMFCs are used in the modern electric vehicles to avoid the usage of gasoline in the cars. The hybrid electric vehicles are powered by batteries and fuel cells are used to sustain the power demand. Presently, most of the big car companies are involved in the development of fuel cell-based cars. For sometime fleets of PEMFC-powered buses have been operated and tested by Chicago Transit Authority and BC (British Columbia) Transit, for public transportation, with the help from Daimler Chrysler/Ballard [32].

In the area of low level power generation, the PEMFC systems are increasingly finding a good fit. Especially in the portable power applications, the PEMFCs with 35–250 W power have found use in video cameras, electric wheelchairs, laptops, intelligent transportation systems, and military communications [32]. On the other hand, the PEMFC systems are also crossing into the stationary applications. Interestingly, Plug Power has installed a 7 kW residential power system based on the PEMFCs to provide a single home in upstate New York with electricity, heat, and hot water [31].

Direct Methanol Fuel Cell

The PEM technology has been used in another kind of fuel cell known as the DMFC. The DMFC operates in the same temperature range as the PEMFC, and it also transports protons across the membrane to produce water and generate electric current. There are several advantages of the DMFC over the PEMFC. One of them is that methanol can be more efficiently stored than hydrogen. Therefore, operating directly with methanol minimizes a significant problem for fuel cell technology: hydrogen storage. Additionally, the use of 1 M methanol directly in the fuel cell helps the water and thermal management of the fuel cell with the PEM. The PEM requires water for the conductivity of protons. However, at the operating temperature, the membrane dries up and the presence of liquid methanol helps in hydrating the PEM. The electrochemical reactions on the anode and cathode electrode of the DMFC are listed below:

Anode reaction

$$\mathrm{CH_3OH + H_2O \rightarrow CO_2 + 6H^+ + 6e^-} \tag{38.68}$$

Cathode reaction

$$\frac{3}{2}\mathrm{O_2 + 6H^+ + 6e^- \rightarrow 3H_2O} \quad E^\circ = 1.23\ \mathrm{V\ vs.\ SHE} \tag{38.69}$$

At the anode electrode, the methanol is oxidized to CO_2, protons, and electrons. The electrons are transported through the external circuit and the protons are migrated through the membrane to the cathode side. At the cathode electrode, oxygen reacts with the protons and the electrons to form water. Pt is not an efficient catalyst for the electrochemical oxidation of methanol, so other metals, as alloys of Pt, are investigated as possible candidates for the anode electrode of the DMFC. The most commonly used catalyst for the methanol oxidation at the anode electrode is Pt-Ru. In the cathode electrode, similar to the PEMFC, Pt is the best catalyst for the oxygen reduction reaction.

One of the issues with the DMFC system is methanol crossover, where methanol migrates through the membrane from the anode to the cathode side of the fuel cell. The membranes used in the DMFC systems were mainly developed for the PEMFC, which was designed to transport protons along with water molecules. Methanol has similar properties to the water which results in easy diffusion as well as electroosmotic drag of the methanol molecules across the PEM. The oxidation reaction of methanol at the Pt catalyst surface interferes with the oxygen reduction reaction and degrades the performance of the DMFC. Membrane development to prevent methanol crossover is the key area of research for the DMFC systems. At the same time, investigation has also been focused on a methanol resistant catalyst for the cathode electrodes. This is a different approach to solve the presence of methanol on the cathode side of the DMFC. In this approach, the methanol crossover across the membrane is not prevented but methanol tolerant catalysts are used to avoid methanol oxidation at the cathode [66–69].

The DMFCs are typically PEMFCs, which uses easily transportable liquid methanol as fuel. Similar to PEMFCs, the DMFCs have the strong potential to be used for the portable power systems. The use of methanol in these fuel cells, which has low noise and thermal signature, shows the possibility in military and defense applications for the DMFC systems. Presently, the potential applications for the DMFCs are focused on the portable power systems and military, and later it could be extended to stationary power generation.

Phosphoric Acid Fuel Cell

One of the most commercialized fuel cells is the phosphoric acid fuel cell (PAFC). The PAFC systems are used as stationary power plants. They have been installed worldwide in the range of 5–10 MW capacity. Phosphoric acid is an excellent acid electrolyte for the fuel cell system and the PAFC electrolyte, phosphoric acid, is stabilized in an SiC matrix. The hydrogen gas obtained from steam reforming or other reformation process can be handled efficiently by the phosphoric acid electrolyte [32]. The PAFCs operating temperature is around 200°C, which is suitable for the electrolyte to provide good ionic conductivity.

The phosphoric acid electrolyte is inert and stable in a reducing and oxidizing environment. The electrolyte also has low volatility in the operating temperature. Among other advantages, the PAFC system has a relatively simple design construction, and it is also a thermally, chemically, and electrochemically stable system [31]. The high temperature operation of the PAFC has resulted in tolerances of up to 1% CO and 40 ppm H_2S. The electrochemical reactions in the anode and cathode electrodes are same as the ones observed in the PEMFCs (38.62) and (38.63). The electrocatalyst used on the anode electrode of the PAFC is Pt, whereas Pt-Co and Pt-Cr-Co supported on carbon has been studied as catalysts for the oxygen reduction reaction on the cathode electrode [32]. The expensive manufacturing cost for the PAFC systems is the main drawback from being widespread installations. The cost of the power from the PAFC power plant is still more than $4,000/kW.

The PAFC systems are mainly applied for stationary power generation. Since the PAFC systems operate at 200°C, the heat energy from the PAFC systems is recovered by integrating it with combined heat and power (CHP) plant. The major company involved in the manufacturing of the PAFC power plants is United Technologies Corporation (UTC). UTC along with its subsidiaries has installed more than 75 MW of PAFC power plants over 19 countries. UTC has manufactured and installed at least 250 PAFC units of each 200 kW power capacity. Presently, one of the UTC's subdivision, UTC Power, Inc., is marketing 400 kW units under the name of PureCell400. A 250 kW PAFC power plant has been used to power the police station at the New York City's Central Park [70]. Toshiba has installed a factory of PAFCs at Hamakawasaki works in Japan [32]. A major power plant of 11 MW capacity was commissioned by UTC along with Toshiba to provide power for Ichihara in Tokyo [31]. Military application of the PAFC systems is being investigated by the US Army Engineering Research and Development Center, Construction Engineering Research Laboratory (ERDC/CERL) with 30 units of 200 kW PAFCs [71].

Molten Carbonate Fuel Cell

The molten carbonate fuel cell (MCFC) is one of the high temperature fuel cells. The operating temperature for the MCFC has been in the range of 600–700°C. The advantages of the MCFC systems are: the use of non-noble metal catalyst for oxygen reduction reactions at the cathode electrode due to a relatively high operating temperature; carbon dioxide is supplied as fuel and not as a poison; the waste heat generated during MCFC operation can be recovered to increase the overall efficiency of the MCFC system; and the MCFCs are capable of combining fuel reforming of natural gas in order to produce hydrogen gas for fuel cell operation [32].

The electrochemical reactions taking place on the anode and cathode electrodes are listed below:

Anode reaction

$$H_2 + {CO_3}^{2-} \rightarrow H_2O + CO_2 + 2e^- \tag{38.70}$$

Cathode reaction

$$\frac{1}{2}O_2 + CO_2 + 2e^- \rightarrow {CO_3}^{2-} \tag{38.71}$$

At the cathode, oxygen reacts with carbon dioxide in order to form the carbonate ion, which is the charge carrier across the electrolyte to the anode side. The carbonate ion reacts with hydrogen on the anode to form water, carbon dioxide, and electric current. Carbon dioxide is recycled back into the cathode stream.

The high operating temperature of the MCFC helps in the use of a non-noble metal catalyst, such as NiO, on the cathode electrode for oxygen reduction reaction. Alternate cathode electrodes investigated are lithium oxide based, such as $LiFeO_2$, Li_2MnO_3, and $LiCoO_2$. A double-layered catalyst like $NiO/LiCoO_2$ has also been used in the MCFC systems to improve the stability of the oxygen reduction reaction [72]. At the anode, Ni is used as the electrode and Ni-Al or Ni-Cr has also been used as the anode to prevent Ni creeping into the molten carbonate electrolyte.

Molten carbonate stabilized in an alumina-based matrix is the electrolyte for the MCFC. The molten carbonate is obtained from the $LiCO_3/K_2CO_3$ (Li/K) mixture supported on the alumina. The sodium carbonate (Li/Na) in the above mixture (Li/K) has increased the alkalinity of the electrolyte and decreased the Ni dendrite formation in the electrolyte. The issues associated with the MCFC technology include the migration of Ni from anode into the electrolyte, corrosion of the NiO at the cathode, and electrolyte migration [32].

Stationary power generation is the area of focus for the MCFC technology. Among various demonstrations of this technology, the stand out and best outcome has been a 2 MW MCFC power plant at Santa Clara, California, USA. In 1996, the 2 MW MCFC power plant was built, commissioned, and tested by the Energy Research Corporation (ERC). It used natural gas as its fuel. This demonstration plant reached a maximum electrical power of 1.93 MW with 2 ppm NO_x, undetectable level of SO_x, and operated within the noise limits. In Bielefeld, Germany, a 250 kW power plant was tested by MTU Friedrichshafen, an affiliate of the Daimler Chrysler group. This power plant operated at 50% efficiency using natural gas, and it also provided heat and hot water to the university at Bielefeld. In 1997, the MC Power Corporation demonstrated a 250 kW power plant in Miramar (San Diego, USA) [31, 32].

Solid Oxide Fuel Cell

The other high temperature fuel cell is the SOFC, which utilizes solid electrolyte made of ceramic material for the electrochemical generation of power. The SOFC operates in the range of 750–1,050°C, which is higher than the MCFC. The high operating temperature of the fuel cell helps in tolerating carbon monoxide, carbon dioxide, and, possibly, reform hydrocarbon fuels. The electrochemical reactions observed on the anode and cathode electrode of the SOFC are listed below:

Anode reaction

$$H_2 + O^{2-} \rightarrow H_2O + 2e^- \tag{38.72}$$

Cathode reaction

$$\frac{1}{2}O_2 + 2e^- \rightarrow O^{2-} \tag{38.73}$$

At the cathode electrode, the oxygen molecule from the air is oxidized to an oxide (O^{2-}) ion, which is transported through the solid oxide electrolyte to the anode side. The oxide ions combine with the hydrogen molecule on the anode electrode to form water and generate electric current. The solid oxide electrolyte commonly used in the SOFC system is made of 85–90% zirconia (ZrO_2) and 10–15% yttria (Y_2O_3), which is referred to as YSZ. The defect in the perovskite (ZrO_2-Y_2O_3) structure as well as the high operating temperature of the SOFC helps in the movement of the oxide ion in the solid YSZ electrolyte. The solid form of the electrolyte results in only two phases (solid–gas) being available (instead of the solid–liquid–gas phases in other types of fuel cells) for the reactions. The two phase situation eliminates the problems, such as water management, flooding of the catalyst, and slow kinetics, due to mass transport of the reactant and products in the liquid phase.

The cathode materials employed for the SOFC are lanthanide-based perovskites, especially lanthanum strontium manganate ($LaSrMnO_3$—LSM). The lanthanide-based perovskites have displayed higher ionic conductivity for the oxide ion and good catalytic activity for the oxygen reduction reaction. Two phase cathode materials are investigated for the SOFC, where a layer of YSZ will be added along with LSM to form the two phase cathode. For many years, the material used as the anode for the SOFC has been Ni-YSZ. The Ni cermet is more stable than the pure Ni as the anode material for the SOFC.

The stack or planar design of the SOFC system has been well studied and documented. The most advanced design of the SOFC system has been a tubular design developed by Siemens-Westinghouse. The thermal stability of the SOFC was improved by the tubular configuration due to the presence of self-sealing structures. Incorporation of good sealing material for the SOFC system has been one of the many challenges faced by the SOFC technology. Additionally, the development of an efficient solid electrolyte, anode, and cathode material for the SOFC has always been emphasized.

Similar to the MCFC systems, application for the SOFC technology is in the stationary power generation or distributed power systems. Siemens-Westinghouse is pioneer in the field of SOFC technology. A 25 kW SOFC unit was installed by Siemens-Westinghouse and tested using natural gas at the University of California's National Fuel Cell Research Center situated in Irvine, California [32]. In 1998, Siemens-Westinghouse collaborated with NUON and ELSAM to develop a 100 kW power production plant in Westvoort, Netherlands. This power plant was able to produce 54 kW of heat from the system at normal operation conditions (400 V). This SOFC power plant has been providing electric current to the grid and hot water to the local heating system [31]. Siemens-Westinghouse has also investigated combining the SOFC systems with gas turbine to increase the overall electrical efficiency. In 1999, Siemens-Westinghouse built a 250 kW prototype of the SOFC gas turbine hybrid power plant. This hybrid power plant was tested by Edison Technology Systems at National Fuel Cell Research Center in Irvine, California [32].

Fuel Processing

The gas streams required for the fuel cells are hydrogen and oxygen, as oxygen can be obtained from air. Most of the fuel cells are supplied with hydrogen and air streams. As mentioned previously, hydrogen production, storage, and transportation represents a significant problem for fuel cell technology. Therefore, the production of hydrogen gas for the fuel cells has been an important field of research and development. Presently, approximately 95% of hydrogen produced in the world comes from reforming of fossil fuels (gasoline, natural gas, and oil) and the rest are from other methods [73, 74].

Reforming fossil fuels for hydrogen production has been achieved by steam reforming, partial oxidation, and coal gasification [75–80]. Hydrocarbons are used in the reforming processes to produce hydrogen, especially methane (CH_4) and methanol (CH_3OH). In the steam reforming method any hydrocarbon (C_nH_m) is capable of being reformed to produce hydrogen and carbon dioxide. The carbon dioxide produced is further converted into hydrogen using a water gas shift reaction. In the partial oxidation method, methane is the main raw material to be reformed into hydrogen gas. As the name suggests, the coal gasification process combusts coal, carbon material, and biomass in order to produce gaseous products. This gas product is composed of carbon dioxide, carbon monoxide, hydrogen, nitrogen, methane, and hydrogen sulfide. In all of the reforming methodologies, purification of hydrogen gas is necessary to meet the requirements of the fuel cells. Purification process is an additional, and energy consuming, step for hydrogen production. So, certain fuel cells, such as DMFC, MCFC, and SOFC, can use methanol as the fuel as well as tolerate some percentage of carbon monoxide and carbon dioxide along with a hydrogen stream.

One of the alternate methods used for the production of hydrogen is electrolysis, which has been primarily focused on water electrolysis. The electrolysis of water or other alternate hydrogen sources, such as ammonia, urea, and coal, are discussed in detail in the later section, "Trends: Novel Electrosynthesis Processes." As described in Fig. 38.1, the synthesis of chemicals using electrochemical methods, where the chemicals synthesized has the capability to produce energy or power. These chemicals store energy in their chemical bonds, so the electrochemical synthesis of the hydrogen gas can also contribute towards electrochemical energy storage systems.

Supercapacitors

After batteries, supercapacitors are considered the next most important device in the area of electrochemical storage. Supercapacitors are also known as ultracapacitors, EDLC, or electric double-layer capacitors. The EDLC has electrodes similar to the conventional electrostatic capacitor, but the capacitance for the EDLC is measured due to the presence of a liquid electrolyte (Fig. 38.8). The electrostatic capacitor is the conventional form of the capacitor where the electrodes are separated by non-polarized materials, such as vacuum or air or mica. A supercapacitor is composed of two electrodes and an electrolyte, such as a battery, but the electrode/electrolyte interface is a capacitor, which is the equivalent of a complete cell with two capacitors in series [29].

Ultracapacitors fall in-between electrostatic capacitors and batteries as they have higher power density (10 kW/kg) than batteries, and also more specific energy density than conventional capacitors. On the contrary, supercapacitors have less energy density (about 5 Wh/kg) when compared to 30–40 Wh/kg from a typical LAB [81]. Figure 38.9 displays a plot of specific power density vs. specific energy density for some energy storage devices.

Ultracapacitors have a longer cycle life than any kind of battery and they can also be more quickly charged than most of the batteries. Supercapacitors are generally divided into three different types based on the charge storage process. The types are double layer capacitors, redox-based electrochemical capacitors, and hybrid capacitors. Carbon derived from charcoal, such as activated carbon, is used as the active material in the electrode for supercapacitors. The electrolyte used for double layer capacitors is either sulfuric acid or organic liquid. The capacitance for double layer capacitors in organic electrolytes has reached 100–120 F/g, but it can be 150–300 F/g in aqueous electrolytes, with the exception that a lower cell voltage has to be maintained to avoid water decomposition [81].

Redox-based electrochemical capacitors are the devices in which fast and reversible redox reactions takes place on the surface of the active material. The active materials used in these capacitors are metal oxides, such as RuO_2, Fe_3O_4, MnO_2, and conducting polymers. Most commonly, RuO_2 has been investigated because it has three distinct oxidation states within 1.2 V. The redox reaction for the RuO_2 active material is fast and reversible, where RuO_2 reacts with protons by electro-adsorption of protons on its surface (38.74). RuO_2-based electrochemical capacitors have displayed a capacitance of more than 600 F/g.

$$RuO_2 + xH^+ + xe^- \rightleftharpoons RuO_{2-x}(OH)_x \tag{38.74}$$

A hybrid capacitor is the most recent electrochemical capacitor which combines a capacitor-like electrode with a battery-like electrode in order to obtain both high power and energy density systems. This concept of hybrid capacitors is based on Li-ion capacitors. In 1999, Amatucci's group developed a hybrid capacitor with nanostructured lithium titanate ($Li_4Ti_5O_{12}$) anode and activated carbon cathode, which was able to deliver a specific energy density of at least 10 Wh/kg at 2.8 V [82]. Recently, this material, $Li_4Ti_5O_{12}$, has been investigated by combining the material with carbon nanofibers in order to produce nanohybrid capacitors [28].

Applications of supercapacitors have been mainly in the consumer electronics area, but more research is currently focused on combining supercapacitors with battery packs in order to provide stable and on-demand power supplies for electric vehicles.

Industrial Electrosynthesis Processes

As discussed previously, under the vision described in Fig. 38.1, electrochemical synthesis can be viewed as a way to store energy, as renewable and/or intermittent electrical power could be used to store energy as chemicals, products, commodities, fuels, etc. Industrial electrosynthesis had been used in different large-scale processes, including: chlorine and sodium hydroxide production, aluminum production, metal winning and refining, hydrogen and oxygen production, and specialized synthesis of organics. This section describes some of these processes. Further reading can be found in the literature [8, 12, 21, 83–95].

Chlor-Alkali Process

One of the largest electrochemical operations in the world is the chlor-alkali process or technology. The industry based on the chlor-alkali technology has been around for more than

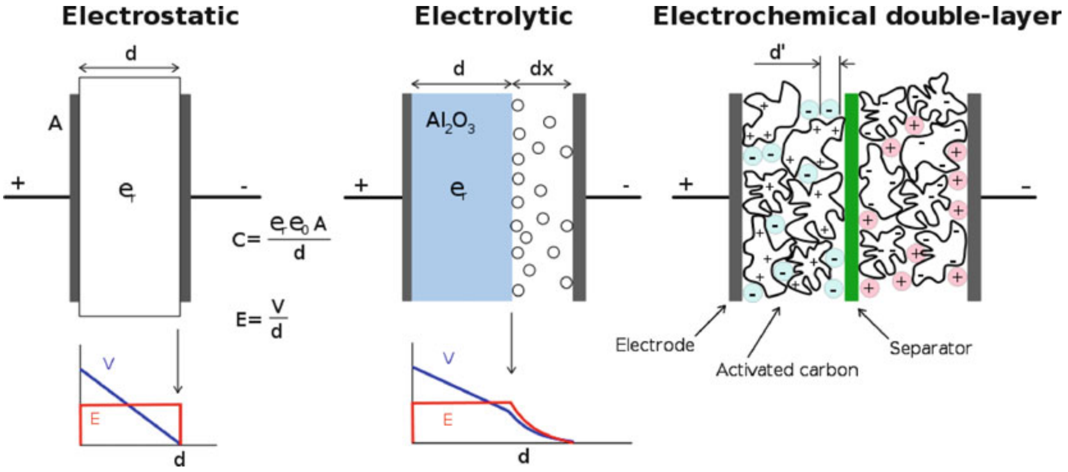


Fig. 38.8 Comparison of three kinds of capacitors [141]. The three types of capacitors are electrostatic, electrolytic, and electrochemical double-layer. The material in-between the positive and negative terminal determines the type of capacitor

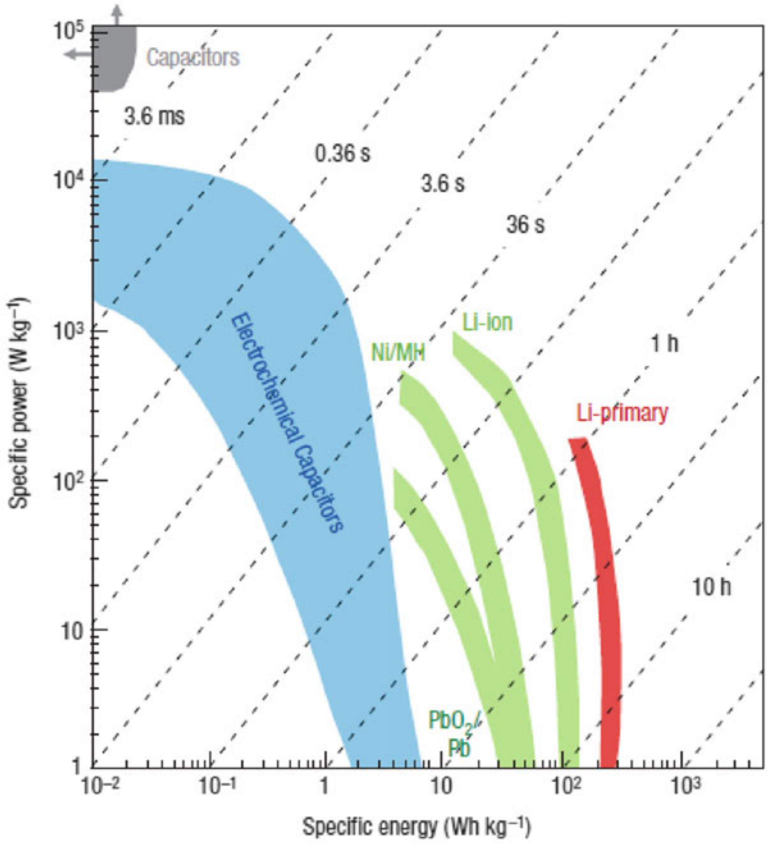


Fig. 38.9 Specific power vs. specific energy (Ragone plot) for energy storage devices [81]. Batteries have high specific energy density but low specific power density. Conversely, the traditional capacitors have very high power density and very low energy density. Electrochemical capacitors have the potential to possess optimum level power and energy density. *Legends*—*PbO_2/Pb* lead acid batteries; *Ni/MH* nickel metal hydride batteries; *Li-ion* lithium ion batteries; *Li-primary* lithium-based primary batteries

100 years. The chlor-alkali process produces chlorine and sodium hydroxide (caustic soda) from the electrolysis of sodium chloride or brine solution. A detailed description of the chlor-alkali technology with its accompanying history, advancements in different electrochemical cells used in the process (including their design), and engineering principles needed to establish chlor-alkali processing plants is available for further reading in the literature [12, 83, 84].

Among the electrochemical methods for producing chemicals, chlor-alkali technology consumes the second largest amount of electricity or electrical energy [86]. Chlorine and caustic soda are critically important chemicals in the world, and they are used in the production of various other chemicals and products which are used in everyday life. Production of more than 95% of chlorine and 99.5% of caustic soda are from the electrochemical method. Chlorine production capacity in the world was 22 million tons in 1970 and it increased to 53 million tons in 2002 [84]. The production of chlorine and caustic soda using this process is another example of the vision described in Fig. 38.1, storing electrical energy in the form of valuable chemicals. The chlorine is used in the manufacture of the polymer polyvinylchloride (PVC), and in the preparation of organic solvents (e.g., methylene chloride, chloroform, carbon tetrachloride), organic compounds (e.g., chlorobenzenes, alkyl chlorides), and inorganic compounds (e.g., hypochlorites of sodium, calcium). The bleaching property of the chlorine is used in the paper and pulp industry, and employed as an oxidizing agent in water treatment operations [12, 84]. The use of sodium hydroxide or caustic soda has been in the paper and pulp industry, soap and detergents manufacturing, and textile industry, as well as in the aluminum extraction process [12, 84].

The electrolysis of sodium chloride solution results in the formation of chlorine and hydrogen gases, and of sodium hydroxide. In 1800, Cruikshank discovered the electrolysis of brine in order to produce chlorine, but it took a relatively long time to scale up this process to manufacturing scale [84]. Development of the chlor-alkali process is attributed to the improvement of the electrolysis cell and its operation. There are three different types of cells used in the chlor-alkali technology, namely the diaphragm cell, the mercury cell, and the membrane cell. In 1855, the diaphragm cell process (Griesheim Cell) was introduced. At the same time, development of the mercury cell process led to the industrialization of the Castner-Kellner cell, which took place in 1892. Finally, in the early 1970s, the membrane cell process was developed as the modification to the diaphragm cell [85]. The current efficiency for the chlor-alkali process has been in the range of 85–98% depending on the type of the cell, operation of the electrolyzer, and purity of the brine solution. The electrical energy demand for performing electrolysis under different types of chlor-alkali technologies are 3,100–3,400 kWh/ton of Cl_2 for the mercury cell process, 2,300–2,900 kWh/ton of Cl_2 for the diaphragm cell process, and 2,100–2,900 kWh/ton of Cl_2 for the membrane cell process [85]. The schematic diagram of the three different cell processes of the chlor-alkali technology is displayed in Fig. 38.10. Brief descriptions of these three different cell designs are presented next.

Mercury Cell

The mercury cell is an undivided cell, as shown in Fig. 38.10, where the base of the container is made of steel in order to provide the electrical contact for the cathode. The cathode for the mercury cell process is liquid mercury (Hg), which is kept flowing during electrolysis. The anode electrode for the mercury cell process is a coated expanded titanium dimensionally stable anode (DSA). The anodes are left in a suspending manner on the top of the container or electrolyzer such that they are parallel to the pool of the Hg cathode. The electrolyte for the mercury cell is a brine (NaCl) solution. The saturated brine solution (25% NaCl), at a temperature of 60°C, is supplied to the top of the electrolyzer. With the electrolysis, the concentration of the brine solution is reduced in the mercury cell. The depleted brine solution (17% NaCl) is removed from the container just above the top of the mercury electrode. The electrochemical reactions taking place at the electrodes in the mercury cell are given below: [12]

Anode reaction

$$2\mathrm{Cl}^- \rightarrow \mathrm{Cl}_2 + 2\mathrm{e}^- \quad E^\circ = 1.31\ \mathrm{V\ vs.\ SHE} \tag{38.75}$$

Cathode reaction

$$\mathrm{Na}^+ + \mathrm{Hg} + \mathrm{e}^- \rightarrow \mathrm{NaHg} \quad E^\circ = -1.85\ \mathrm{V\ vs.\ SHE} \tag{38.76}$$

At the anode, the chloride ions are oxidized to a chlorine gas, which at first saturates the electrolyte solution before escaping the mercury cell. The Na^+ ions present in the electrolyte migrates to the Hg cathode where they are reduced to form an amalgam. The overall voltage for the reactions at the electrodes is 3.16 V but typical operation of the mercury cell requires 4.50 V due to the resistance from electrical contacts. This amalgam containing liquid Hg, from the mercury cell, is taken into a container known as a decomposer or denuder. In the decomposer, the mercury is recovered from the Na(Hg) amalgam and sent back to the electrolyzer in order to use it as cathode. The Na(Hg) amalgam, in the decomposer, trickled along with water over the graphite particles containing transition metals (e.g., Fe and Ni) to obtain sodium hydroxide and mercury (38.77) [86].

$$2\mathrm{NaHg} + 2\mathrm{H_2O} \rightarrow 2\mathrm{NaOH} + \mathrm{H_2} + 2\mathrm{Hg} \tag{38.77}$$

One of the issues with the mercury cell electrolyzer is the health and environmental damages associated with mercury. The mercury cell was widely used in many countries for the production of chlorine and caustic soda during the early days of the chlor-alkali industry. In 1984, 45% of the chlor-alkali plants in the world employed mercury cell processes and its

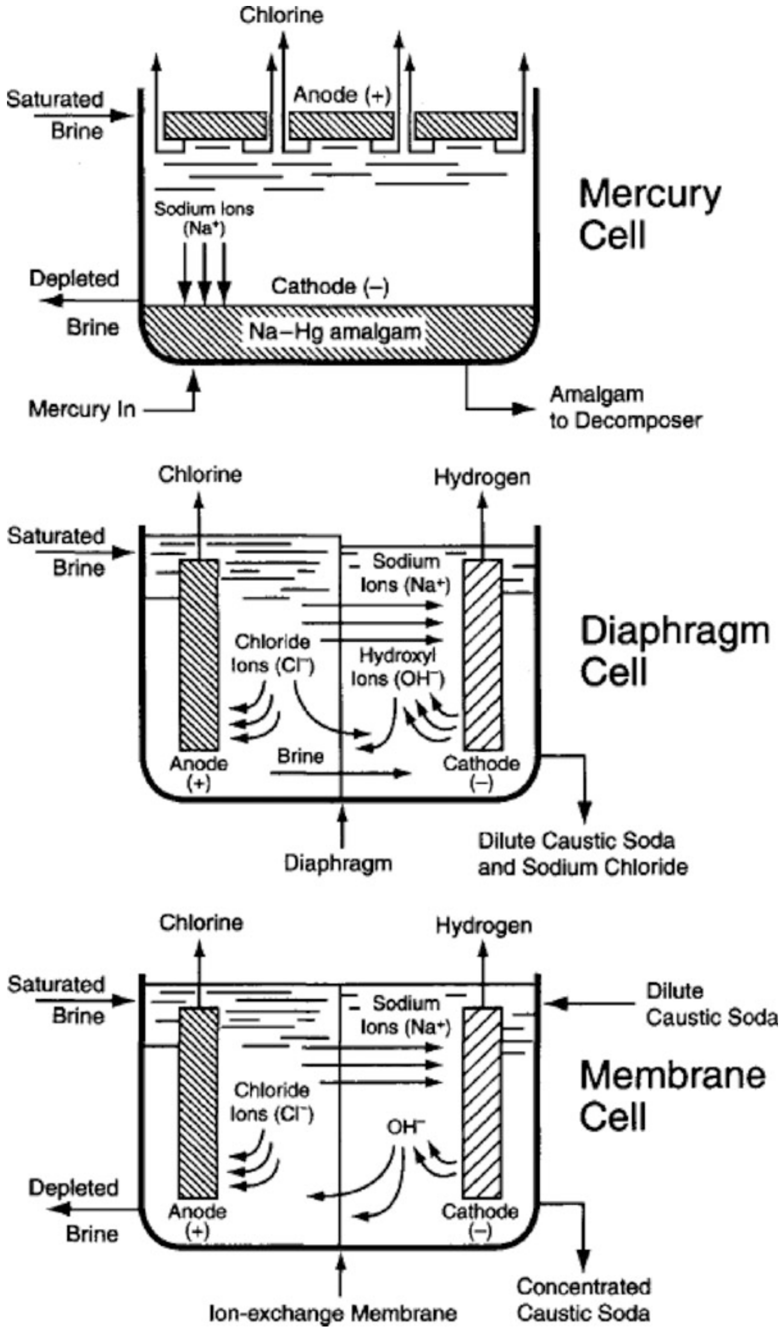


Fig. 38.10 Schematic diagram of the three different cells used in the chlor-alkali technology [84]. The mercury, diaphragm, and membrane cells are depicted with the flow of brine solution, flow of ions across the cells, and production of chlorine, hydrogen, and sodium hydroxide

percentage declined to only 18% in 2001 due to the health and environmental concerns of using mercury [84]. In fact, presently there are no functioning mercury cell type chlor-alkali plants in Japan and many countries are planning to gradually phase out the mercury cell process.

Diaphragm Cell

This process uses a divided cell, where a diaphragm is used to separate the anode side from the cathode side. Asbestos as diaphragm was used in the earlier days before the knowledge of the detrimental health effects of asbestos was known. Now, non-asbestos-based diaphragm are used in the construction of the diaphragm cells and plans are being taken to convert all the diaphragms to non-asbestos material.

Saturated brine solution is fed to the anode compartment of the diaphragm cell, where the chloride ions are oxidized to a chlorine gas as shown in the (38.75). The brine solution flows from the anode to cathode side through the diaphragm. The anodes used in the diaphragm cell process have either ruthenium-based or platinum/iridium-based coated titanium electrode (DSA) [86]. During the electrolysis, the temperature of the electrolyte is maintained between ~95 and 100°C.

On the cathode side, hydrogen gas is generated from the reduction of water (38.78) even using a nickel electrode. The Na^+ ion from the anode compartment migrates through the diaphragm into the cathode side to combine with hydroxyl ion from water reduction reaction to form sodium hydroxide (38.79).

Cathode reaction

$$2H_2O + 2e^- \rightarrow H_2 + 2OH^- \quad E^\circ = -0.84 \text{ V vs. SHE} \tag{38.78}$$

$$2Na^+ + 2OH^- \rightarrow 2NaOH \tag{38.79}$$

Overall reaction

$$2NaCl + 2H_2O \rightarrow 2NaOH + Cl_2 + H_2 \quad E = 2.15 \text{ V} \tag{38.80}$$

The voltage for the electrolysis of brine to produce caustic soda, chlorine, and hydrogen is 2.15 V. However, the normal operational cell voltage for this process is in the range of 3.2–3.8 V, so as to overcome the resistance encountered during the electrolysis [12]. In this process, the cathode electrolyte solution with concentration of 12% NaOH and 15% NaCl is taken for further processing to separate the caustic soda from brine solution for commercial application [86]. The purpose of the diaphragm in this cell is to allow the passage of Na^+ ions from the anode to cathode side but also to avoid the mixing of caustic soda with chlorine. Still there are two parasitic reactions possible on the anode electrode, namely the formation of oxygen along with chlorine (38.81) and oxidation of hypochlorite to chlorate (38.82) [86].

$$2H_2O \rightarrow O_2 + 4H^+ + 4e^- \quad E^\circ = 1.23 \text{ V vs. SHE} \tag{38.81}$$

$$6ClO^- + 3H_2O \rightarrow 2ClO_3{}^- + 4Cl^- + 6H^+ + \frac{3}{2}O_2 + 6e^- \tag{38.82}$$

Membrane Cell

The operation of the membrane cell electrolyzer is similar to the diaphragm cell process except for the ion exchange membrane being used in place of a diaphragm. The ion exchange membrane used in the electrolyzer is commercially known as Nafion, which is sulfonated tetrafluoroethylene-based fluoropolymer-copolymer. The electrochemical reactions at the anode and cathode electrodes for the membrane cell process are same as in the diaphragm cell process. The brine solution used as electrolyte for this process should have less than 0.1 ppm Ca^{2+} and Mg^{2+} ions to avoid precipitation of salt over the electrodes and the ion exchange membrane [86]. The saturated brine solution is added to the anode compartment and only the Na^+ ions along with water molecules are migrated across the Nafion membrane to the cathode side. The depleted brine solution is removed from the anode compartment. On the cathode side of the membrane cell, solution containing dilute concentration of caustic soda is added to increase the concentration of caustic soda in the product stream.

High purity of caustic soda, up to 50% NaOH, can be produced in the cathode compartment using this membrane cell process as the chloride ions are not migrated to the cathode side. However, in practice, 32–35% NaOH is normally produced on the cathode compartment from this process. Another interesting feature of this process is the distance between the anode and cathode electrode can be reduced to almost zero because of the use of ion exchange membrane. This option of sandwiching the electrodes with Nafion membrane has resulted in operating cell voltage of ~2.7 V compared to 3.2–3.8 V for diaphragm cell process. This means the membrane cell process will consume less energy than diaphragm cell process [12].

A new approach is being investigated in the membrane cell process for reducing the energy consumption of the electrolysis and also to avoid the production of hydrogen gas at the cathode. In the chlor-alkali industry, importance is given towards the production of caustic soda and chlorine from the electrolysis of brine (NaCl), but not as much importance as is given for hydrogen generation. So, the evolution of hydrogen gas at the cathode electrode is considered an undesired reaction. The approach is to use oxygen-depolarized cathode (ODC) electrode to avoid hydrogen formation at cathode by reduction of oxygen in the presence of water to form hydroxyl ions (38.83). The ODC is made of GDE which is normally used in the fuel cells [85].

$$O_2 + 2H_2O + 4e^- \rightarrow 4OH^- \quad E^\circ = 0.40 \text{ V vs. SHE} \tag{38.83}$$

Now combining the above cathode reaction with the oxidation of chloride ion to chlorine at the anode (38.75) and formation of NaOH with Na^+ ions (38.79), we obtain the overall reaction for the electrolysis of brine using ODC electrode as

$$2NaCl + \frac{1}{2}O_2 + H_2O \rightarrow 2NaOH + Cl_2 \quad E = 0.91 \text{ V} \tag{38.84}$$

From the (38.84), the voltage for the overall reaction is only 0.91 V compared to 2.15 V for the traditional membrane cell process. The use of ODC electrode has resulted in

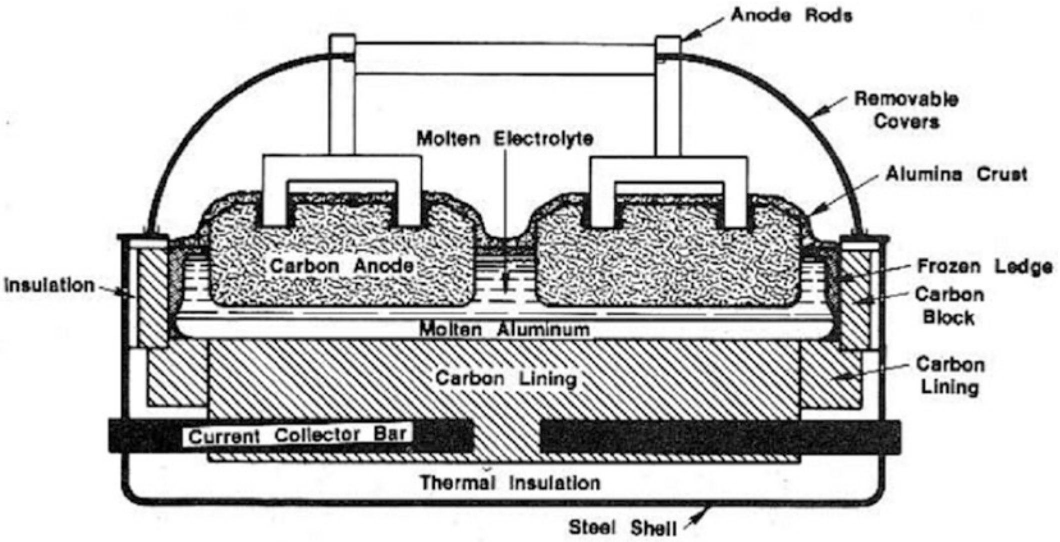


Fig. 38.11 Cross-sectional view of the electrolytic cell to produce aluminum [87]. Prebaked anodes are used in this type of electrolytic cell. Molten aluminum is produced from the Hall-Heroult process

the operation cell voltage of a mere 1.23 V for the new membrane cell process, which translates to almost 30% less electric energy consumption than the traditional approach of membrane cell where hydrogen gas is produced [85]. A detailed description on the development of the ODC electrodes with different catalyst and design of the electrolyzer can be found in the literature [85].

Aluminum Production

Aluminum is the third most abundant element in the Earth's crust. Since aluminum can easily react with oxygen, it does not exist in metallic form in the Earth's crust but in the form of ores, such as bauxite. The valuable properties of aluminum are its being soft, durable, lightweight, electrically conducting, nonmagnetic, stiffness equivalent to steel, and corrosion resistant. These favorable characteristics have resulted in various uses for aluminum. Aluminum has found application in the areas of transportation, packaging (e.g., beverage cans), construction, electronics, and electrical wires.

Electrochemical method of producing aluminum was not effective until 1886, when Charles Martin Hall from Ohio, USA, and Paul L. T. Heroult from Paris, France, independently invented the electrolytic way of extracting aluminum from its molten salt. The Hall-Heroult process, which is widely used throughout the world, for aluminum production has been a very reliable and efficient method for the past 125 years. Over these years, the process has undergone technological improvements mostly in the anode material [22]. This process involves electrolysis of alumina (Al_2O_3), which is dissolved in molten cryolite-based electrolyte. Alumina for the Hall-Heroult process was obtained from the aluminum ore, bauxite, using the Bayer process. In the Bayer process, the bauxite ore, which has almost 50% Al_2O_3, was digested in sodium hydroxide (NaOH) solution under pressure resulting in the formation of sodium aluminate and the impurities were left behind as red mud (38.85).

$$Al_2O_3 \cdot 3H_2O + 2NaOH \xrightleftharpoons{\text{extraction}} 2NaAlO_2 + 4H_2O \tag{38.85}$$

The aluminate solution was seeded with hydrated alumina crystals to re-precipitate alumina, which can easily be washed. In the final step, the hydrated alumina was calcined at 1,200°C to remove water vapor and obtain more than 99% pure alumina [12, 83].

In the Hall-Heroult process, the electrolytic cell is called as pot and a cross-sectional view of the electrolytic cell is shown in Fig. 38.11.

In this process, the anode electrode, mainly made of carbon, is consumed during the electrolysis of alumina. The electrolytic cells are differentiated by the kind of anode electrodes used. There are two kinds of anodes in the Hall-Heroult process, prebaked and self-baking, or Søderberg anodes. Both kinds of anodes are made of petroleum coke and coal tar, where the prebaked anodes are baked in brick lined pits before using in the electrolytic cells. However, the Søderberg anodes are allowed to bake in the cell using the heat from the electrolysis of alumina. In Fig. 38.11, the electrolytic cell has a steel shell container which is lined with carbon to act as the cathode for the electrolysis.

The electrolyte has alumina, from the Bayer process, mixed with molten cryolite (Na_3AlF_6) solution. Calcium fluoride is also included as an additive. The ratio of the sodium fluoride to aluminum fluoride is maintained at 1.5 in the molten electrolyte. The temperature of the electrolytic cell was maintained in the range of 920–980°C in order to keep the electrolyte in molten condition. Even though the Hall-Heroult process is 125 years old, the explanation for the electrode reaction mechanism has not been clearly documented due to the complex nature of the electrolyte and the possibility of having various species in the electrolyte. The readers are advised to check this literature to understand the complex nature of the ions present in the electrolyte and their reactions [87]. A simplistic explanation would be that the alumina dissociates into aluminum ion and oxide ion in the molten electrolyte. At the cathode, the aluminum ion is reduced to liquid aluminum metal at the operating temperature and collected on top of the cathode electrode. At the anode electrode, oxide ion is oxidized to oxygen gas, which reacts with the carbon anode to form carbon dioxide. So, the overall reaction would be alumina reacts with carbon to form metallic aluminum and carbon dioxide [12].

Anode reaction

$$C + 2O^{2-} \rightarrow CO_2 + 4e^- \tag{38.86}$$

Cathode reaction

$$Al^{3+} + 3e^- \rightarrow Al \tag{38.87}$$

Overall reaction

$$2Al_2O_3 + 3C \rightarrow 4Al + 3CO_2 \tag{38.88}$$

The cell voltage for the overall reaction is 1.18 V, but due to the electrolyte resistance and other voltage drops, the operating voltage required for aluminum production ranges from 4.0 to 4.5 V [12]. In 1914, on average 59 kg of Al was produced per day from a plant operating at 10 kA and 8 V with 73.3% as current efficiency. Over the years, the production capacity of the Hall-Heroult process has improved to 2,308 kg of Al per day, in 2000, at 300 kA and 4.2 V operating condition with a current efficiency of 95.6%. Still, among the electrolytic processes, the production of aluminum is the highest energy consuming process, which translates into high demand for electricity by this technology. During this period, the energy requirement has become better, from 32.53 kWh in 1914 to 13.10 kWh for every kilogram of Al produced in 2000 [22].

Electrowinning and Refining Process

The process of extracting metals from its ores using the electrochemical method is known as electrowinning. The electrowinning method is the last step in the entire extraction process of a metal from its ore. The ores of a metal are generally in the form of metal sulfides or oxides or mixture of both. For simplicity, let us consider the sulfide form of the metal as the ore to be used in the extraction process. The sulfide ore will be converted into oxide form by roasting the metal ore in the air [12].

$$2MS + 3O_2 \rightarrow 2MO + 2SO_2 \tag{38.89}$$

The oxide salt of the metal is leached or dissolved in acid solution especially with sulfuric acid to form acid solution containing metal ions [83].

$$MO_x + 2xH^+ \rightarrow M^{2x+} + xH_2O \tag{38.90}$$

In fact, the acid solution contains different metal ions from the ore along with the desired metal ion, so the solution is purified and subsequently concentrated with the desired metal ions. The electrowinning process is basically the electrolysis of the acid solution, where the metal ions are reduced into metal on the cathode electrode (38.91) and the oxygen gas is evolved at the anode (38.92).

$$M^{2x+} + 2xe^- \rightarrow M \tag{38.91}$$

$$xH_2O \rightarrow \frac{1}{2}xO_2 + 2xH^+ + 2xe^- \tag{38.92}$$

Along with the oxygen gas, the protons or acid is formed at the anode electrode, which will be recirculated for leaching the metal oxide ores. The overall reaction of the electrowining process is the deposition of the metal and the release of oxygen gas (38.93). There is no loss of water or the acid in this electrowinning process.

$$MO_x \rightarrow M + \frac{1}{2}xO_2 \tag{38.93}$$

The few elements that can be electrowon under aqueous conditions are copper (Cu), silver (Ag), nickel (Ni), cobalt (Co), chromium (Cr), and zinc (Zn). However, the highly reactive elements, such as lithium (Li), sodium (Na), magnesium (Mg), and aluminum (Al), needs molten electrolyte to extract them using the electrowinning process. Electrolysis of alumina to produce aluminum is also considered as the electrowinning process for aluminum metal, and a detailed description of this process was provided in the above section, “Aluminum Production.” The electrochemistry of extracting the above listed elements can be found in the literature [12, 83]. In addition to extracting the metal from its ores, the electrowinning process is used to recover the metals from industrial wastewater in order to eliminate or reduce the metal contamination in the wastewater.

The electrolytic process to refine or increase the purity of the metal is defined as electrorefining, and this method is employed in the metallurgy industry. The impure metal is taken as anode of the electrolytic cell and the electrolyte is the solution of that metal salt. During the electrolysis process, the desired metal is oxidized into metal ion at the anode (38.94), which dissolves into the electrolyte. At the same time, the pure metal is deposited over the cathode by reducing the metal ion from the electrolyte (38.95). The anode and cathode reactions ((38.94) and (38.95), respectively) are reversible reactions, which should result in a net zero cell voltage.

$$\mathrm{M} \rightarrow \mathrm{M}^{n+} + ne^{-} \tag{38.94}$$

$$\mathrm{M}^{n+} + ne^{-} \rightarrow \mathrm{M} \tag{38.95}$$

But in practical applications, the electrolytic cells used for the electrorefining processes require certain cell voltage. The cell voltage was necessary to overcome the solution resistance, voltage drops at the electrical contacts for the anode and cathode, and the overpotentials for both reactions (anodic and cathodic) to take place at a reasonable rate. At the anode electrode, metals nobler than M will be oxidized and retained in the electrode as residue, which is commonly known as the anode slime. But those metals which are less noble than M, after undergoing oxidation will be dissolved into the electrolyte as contaminants. These contaminants present in the electrolyte are not reduced onto the cathode before the reduction of metal M, and the undesired metal can build up during the electrorefining process thereby increasing the impurity.

In the electrorefining process, the use of aqueous electrolyte has been applied for elements, such as copper, nickel, cobalt, lead, silver, gold, and tin. The refining process for the copper metal is extensively used in the production of high purity copper metal for electronics applications. Similarly in the jewelry industry, the electrorefining process is used to remove or extract the noble metal (e.g., gold) from the anode slime or contaminated solutions. Electrorefining of aluminum is performed using the molten electrolyte, which is mainly applied for the recycling of waste aluminum scrap metal. More detailed information on the electrorefining process can be obtained in the literature [12, 83].

Hydrogen and Oxygen Production

Manufacturing of gases like hydrogen and oxygen is also accomplished using electrochemical methods. Hydrogen gas is formed as a by-product during the electrolysis of a brine solution in the chlor-alkali process. Depending on the type of the chlor-alkali cell, the hydrogen gas is purified for commercial applications. However, the water electrolysis process is entirely committed for the production of both hydrogen and oxygen gases with highest purity. A detailed description on the science and the electrochemical reactions associated with the electrolysis of water is explained in the following section, "Electro-Synthesis of Hydrogen."

In 1800, Nicholson and Carlisle discovered the splitting of water into hydrogen and oxygen [21, 83]. The development and commercialization of water electrolyzers for the industrial scale began in the 1920s and 1930s [90]. During that time several 100 MW electrolyzers were built around the world for ammonia fertilizers and petroleum refining [21, 90]. Among the notable installations of the water electrolyzers was at the Aswan High Dam site in the late 1980s with 144 electrolyzers, which was rated at 162 MW and a capacity to produce 32,400 m^3/h of hydrogen gas [21]. The cell designs used for the electrolyzer were monopolar electrodes in tank electrolyzers and bipolar electrodes in a filter press cells. High pressure electrolyzers were also tested and commercially produced. One of the industrial water electrolyzers manufactured by DeNora Impianti Electtrochimici S.p.A. is shown in Fig. 38.12.

The details of other commercially available water electrolyzers with their operation parameters are listed in Table 38.9 [21].

Organic Electrosynthesis

The synthesis of organic compounds was traditionally favored through chemical routes. From the beginning of the last century, however, the use of electrochemical reactions to synthesize organic compounds has been practiced for various reasons. The advantages on using the electrochemical synthesis method include: (1) product selectivity and purity is higher than for chemical synthesis, (2) less reaction steps are needed than for the chemical method, (3) inexpensive starting materials are used, (4) least polluting chemicals or by-products are formed, (5) less hazardous routes are involved, and (6) less energy is consumed [93, 96].

The electrosynthesis processes for organic compounds also have disadvantages or limitations in terms of selecting the electrode material and the electrolyte composition, because generally aqueous electrolytes are used in electrochemical synthesis, but solubility of organic compounds in aqueous solutions has to be sorted out. Some of the shortcomings of the electrochemical processes to synthesize organic compounds have been the discovery of an alternate catalytic route (sorbitol and propylene oxide), an alternate chemical route was discovered to avoid the electrochemical intermediate product (pinacol), the formation of a pollutant product (tetraalkyl lead), and the product had lost its market or the market never existed [93].

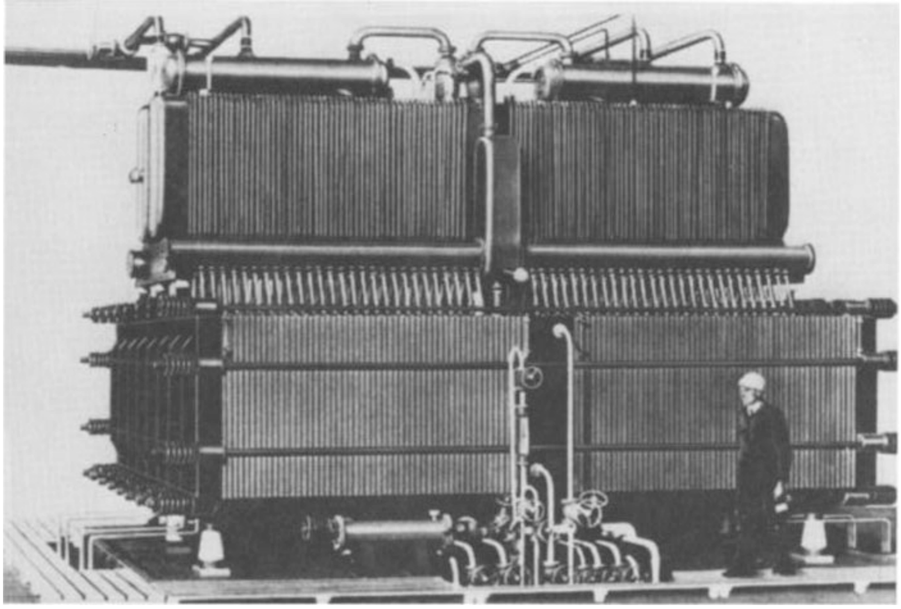


Fig. 38.12 Bipolar electrolyzer manufactured by DeNora Impianti Electtrochimici S.p.A. This unit has 54 cells, operated with a current of 2.5 kA, power consumption was approximately 4.6 kWh N/m^3 and the total capacity of the electrolyzer was 250 kW [90]

Table 38.9 Water electrolyzers and other their operating conditions

Parameter	De Nora S.p.A.	Norsk hydro	Electrolyzer Corp. Ltd.	Teledyne energy systems	General electric
Cell type	B-FP	B-FP	M-T	B-FP	B-FP
Anode	Expanded Ni-plated mild steel	Activated Ni-coated steel	Ni-coated steel	Ni screen	PTFE-bonded noble metal
Cathode	Activated Ni-plated steel	Activated Ni-coated steel	Steel	Ni screen	PTFE-bonded noble metal
Pressure (MPa)	Ambient	Ambient	Ambient	0.2	0.4
Temperature (°C)	80	80	70	82	80
Electrolyte	29 % KOH	25 % KOH	28 % KOH	35 %	Nafion
Current density (A/m^2)	1,500	1,750	1,340	2,000	5,000
Cell voltage (V)	1.85	1.75	1.9	1.9	1.7
Current efficiency (%)	98.5	98.5	>99.9	NR	NR
Oxygen purity (%)	99.6	99.3–99.7	99.7	>98.0	>98.0
Hydrogen purity (%)	99.9	98.9–99.9	99.9	99.99	>99.0

B-FP bipolar electrodes in filter press cells; *M-T* monopolar electrodes in tank cells
Source: Zeng and Zhang [21]

One of the well-known organic compounds electrochemically synthesized is adiponitrile. The chemical adiponitrile is an intermediate in the production of Nylon 66. In 1959, Manuel Baizer at Monsanto discovered the process to produce adiponitrile by the electrochemical method of hydrodimerization of acrylonitrile at the cathode. This method was later known as the Monsanto process. "More details and progress on the Monsanto process can be gathered from the literature [12]. In the Monsanto process, oxygen gas is evolved at the anode, while at the cathode acrylonitrile is electrohydrodimerized to adiponitrile, as shown below: [93]

$$2CH_2 = CHCN + 2H^+ + 2e^- \rightarrow NC(CH_2)_4CN \quad (38.96)$$

The Monsanto plant at Decatur, Alabama, USA, is capable of producing 200 million kg of adiponitrile per year. The cathode used for this process is cadmium and the anode is a steel electrode, which are both assembled in a bipolar way in an undivided cell. The electrolyte comprises acrylonitrile, a quaternary salt (hexamethylene(bisethyltributyl)ammonium phosphate), phosphate buffer, and the anode anticorrosive additives, borax, and EDTA. A current density of 2 kA/m^2 is maintained at 55°C operating temperature and the process has a power consumption of 2.4 kWh/kg of adiponitrile produced [93, 96]. Table 38.10 provides a list of various organic compounds manufactured in the world using the electrosynthesis approach; some of these processes are commercialized, others in the pilot stage, and some of them are decommissioned.

Table 38.10 List of some organic compounds produced using industrial electrosynthesis

Product	Starting material	Company	Operating status
Acetoin	Butanone	BASF	C
1-Acetoxynaphthalene	Naphthalene	BASF	P
Acetylenedicarboxylic acid	1,4-Butynediol	BASF	C
Adipoin dimethyl acetal	Cyclohexanone	BASF	C
Adiponitrile	Acrylonitrile	Monsanto (Solutia), BASF, Asahi Chemical	C
2-Aminobenzyl alcohol	Anthranilic acid	BASF	P
4-Aminomethylpyridine	4-Cyanopyridine	Reilly Tar	C
Anthraquinone	Anthracene	L. B. Holliday, ECRC	C
Anthraquinone	Naphthalene, butadiene	Hydro Quebec	P
Arabinose	Gluconate	Electrosynthesis Co.	P
Azobenzene	Nitrobenzene	Johnson Matthey Company	C
Calcium gluconate	Glucose	Sandoz, India	C
Calcium lactobionate	Lactose	Sandoz, India	C
S-Carbomethoxymethylcysteine	Cysteine + chloroacetic acid	Spain	C
Diacetone-2-ketogulonic acid	Diacetone-*L*-sorbose	Hoffman-LaRoche	C
Dialdehyde starch	Starch	CECRI	C
3,6-Dichloropicolinic acid	3,4,5,6-Tetrachloro-picolinic acid	Dow	P
1,4-Dihydronaphthalene	Naphthalene	Clariant	C
1,2-Dihydrophthalic acid	*o*-Phthalic acid	BASF	D
2,5-Dimethoxy-2,5-dihydrofuran	Furan	BASF	C
Ethylene glycol	Formaldehyde	Electrosynthesis Co.	P
Gluconic acid	Glucose	Sandoz, India	C
Glyoxylic acid	Oxalic acid	Rhone Poulenc, Steetley	P
Hexafluoropropyleneoxide	Hexafluoropropylene	Clariant	C
Monochloroacetic acid	tri- and di-Chloroacetic acid	Clariant	P
Mucic acid	Galacturonic acid	EDF	C
Nitrobenzene	*p*-Aminophenol	India, Monsanto	P
5-Nitronaphthoquinone	1-Nitronaphthalene	Hydro Quebec	P
Partially fluorinated hydrocarbons	Alkanes and alkenes	Phillips Petroleum	P
Perfluorinated hydrocarbons	Alkyl substrates	3M, Bayer, Clariant	C
Pinacol	Acetone	BASF, Diamond Shamrock	P
Piperidine	Pyridine	Robinson Bros.	D
Polysilanes	Chlorosilanes	Osaka Gas	C
Salicylaldehyde	*o*-Hydroxybenzoic acid	India	C
Sorbitol	Glucose	Hercules	D
Substituted benzaldehydes	Substituted toluenes	Hydro Quebec, W.R. Grace	P
Succinic acid	Maleic acid	CERCI, India	C
Tetraalkyl lead	Alkyl halide	Nalco	D

C commercial process; *P* pilot process; *D* discontinued commercial process
Source: Sequeira and Santos [93]

Trends: Novel Electrosynthesis Processes

The latest processes involved in the electrosynthesis of chemicals which have potential for power generation, and water and air remediation are discussed in this section. The new interest in utilizing solar energy to manufacture chemicals capable of being used as fuels has caught the attention of many researchers. Additionally, the electrochemical approach to synthesize hydrogen from various chemicals (water, ammonia, urea, and coal) is also described in this section. The main focus of the electrolysis of ammonia and urea has been to remediate the ammonia and urea rich wastewater. The hydrogen gas produced during the ammonia and urea electrolysis process is considered as high value by product. The advantage of coal electrolysis is to reuse the coal waste for the production of hydrogen, and also to generate other chemicals which have potential to be classified as liquid fuels.

Solar Fuels

In recent years, a new area of research has been devoted towards producing liquid fuels, which can be used for

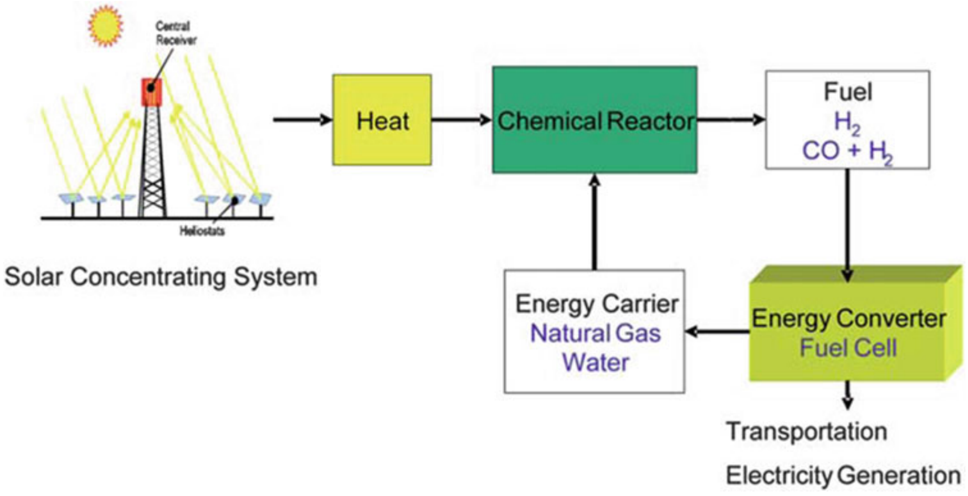


Fig. 38.13 Schematic diagram for the concept of solar fuel production [142]. The heat from the concentrated solar power plant is used in the chemical reactor to produce fuels, such as H_2 and syn gas ($CO + H_2$). The fuels produced from the chemical reactor will be utilized in fuel cells for power generation or transportation applications. Alternatively, the hydrogen can be combined with CO and CO_2 to form hydrocarbons, such as methane

generating energy or power from solar energy. These fuels are referred to as "solar fuels," because the energy needed to produce them comes from the sun. Solar fuels can be used as raw materials for manufacturing different kinds of chemical products as well as generating power. In this way, solar fuels can be considered energy carriers.

The methodology adapted to produce solar fuels was to utilize solar energy thorough solar cells in order to split water into oxygen and hydrogen. The hydrogen gas produced from water electrolysis is combined with carbon dioxide and/or carbon monoxide found in industrial effluents to produce hydrocarbons, which are referred to as "solar fuels." This method not only uses renewable energy sources to produce chemicals that have properties as fuel, but it also converts pollutants present in industrial effluents into useful products.

The solar energy is collected in large power plants, known as concentrating solar power plants (CSP). These systems are being built around the world in regions of relatively high solar radiation. A schematic diagram illustrating the idea of using solar energy obtained from the concentrated solar power plant for the production of solar fuels is shown in Fig. 38.13. The concept involves solar thermochemical cycles to produce hydrogen from solar steam forming of natural gas and methane cracking.

The splitting of water using solar energy is also achieved by photoelectrochemical (PEC) and photocatalytic methods [97]. In the photocatalytic method, the catalyst material is dispersed in pure water where it uses the solar energy to split water, and produce hydrogen and oxygen together in the solution. However, in the PEC method, the photoactive material is on the electrode and the process of splitting water is performed in an electrolytic cell with the ability to separate hydrogen and oxygen gases. In the PEC method of water splitting, titanium dioxide (TiO_2) is used as the active material on the electrodes [98, 99]. Other materials, including semiconductors, have been investigated as a photoelectrode for the conversion of solar energy to hydrogen gas.

An inexpensive and stable oxide photoanode material has been WO_3. This material is used along with dye sensitized solar cells (DSC) to capture solar photons, because the WO_3 is incapable of separately harvesting the solar photons. The efficiency to convert the solar energy to produce hydrogen gas is known as solar to hydrogen efficiency. The WO_3 photoanode has a maximum solar to hydrogen efficiency of only 8% [100]. A promising photoanode material, which is inexpensive, stable, and capable of harvesting solar photons, is iron (III) oxide (Fe_2O_3), especially α –Fe_2O_3 (hematite) [97]. The iron (III) oxide has the capacity to achieve 16.8% efficiency towards the solar to hydrogen conversion [100]. The Solar Fuel Center for Chemical Innovation (CCI Solar) at the California Institute for Technology has been working in the area of solar fuels for many years. They have developed an electrochemical fuel cell with nanorods as photoelectrodes, which can oxidize and reduce water to produce oxygen and hydrogen, sandwiched by a proton conducting membrane [101]. CCI Solar has been undertaking leading research in the area of solar fuels and more information on their solar fuels research can be obtained from their website [102].

Electro-Synthesis of Hydrogen

During the production of solar fuels, hydrogen gas is used as the main raw material to produce hydrocarbons. Hydrogen gas has, however, been used as a fuel for many years to generate electric power, in various types of fuel cells (see section on "Fuel Cells"). Hydrogen gas is also referred to as an energy carrier. So, the production of hydrogen gas by electrochemical methods and its storage should be considered as one of the methods for electrochemical energy storage. Traditionally, electrolysis is used to split a chemical compound into its elemental forms and water electrolysis has been utilized to produce hydrogen gas. In addition, other starting materials, such as ammonia, urea, and coal particles, are also available to produce hydrogen via electrolysis.

Water Electrolysis

The electrochemistry of water electrolysis is discussed in this subsection (For more details on industrial scale electrolyzers, see section on "Hydrogen and Oxygen Production"). The splitting of a water molecule into hydrogen and oxygen gases with the help of electric current is known as water electrolysis. Initially, nickel electrodes were used as the cathode and anode material for the electrolysis of water because of its low cost. Recent research indicates the use of nickel-based alloys, steel, and nickel-coated or noble metal-coated steel as effective electrode materials [21]. KOH solution is used as the electrolyte for water electrolysis, and is also responsible for the name "alkaline water electrolysis."

The anode and cathode reactions for the electrochemical splitting of water to produce hydrogen and oxygen are given below:

Anode reaction

$$4OH^- \rightarrow O_2 + 2H_2O + 4e^- \quad E^\circ = 0.40\ \text{V vs. SHE} \tag{38.97}$$

Cathode reaction

$$4H_2O + 4e^- \rightarrow 2H_2 + 4OH^- \quad E^\circ = -0.83\ \text{V vs. SHE} \tag{38.98}$$

Overall reaction

$$2H_2O \rightarrow O_2 + 2H_2 \quad E = 1.23\ \text{V} \tag{38.99}$$

The theoretical voltage required to dissociate water into hydrogen and oxygen gases is 1.23 V, whereas the thermoneutral voltage is 1.48 V. Commercially available water electrolyzers operate between 1.6 and 2.0 V. The anode and cathode sides of the electrolyzer are separated by a membrane that allows the ions to pass through but keep the gases separate. Therefore, the hydrogen gas produced using this method does not require further purification as it is almost 100% pure.

The hydrogen gas produced from water electrolysis can be supplied to a fuel cell to generate electric power and water. The reactions taking place inside the fuel cell are opposite to the ones observed in the water electrolyzer. So, combining the water electrolyzer with the fuel cell will result in a net energy or power loss because more energy is required to split water in the electrolyzer. Additionally, deionized water is necessary for the optimum functioning of the water electrolyzer [103].

Ammonia Electrolysis

Water is not the only source of hydrogen. There are other chemicals which have a higher atomic weight of hydrogen present, such as ammonia and urea. Ammonia is a colorless gas with a pungent smell. Ammonia is produced by the Haber-Bosch process, in which molecular nitrogen is reacted with hydrogen at high temperature and pressure in the presence of a catalyst to produce ammonia. Ammonia has, however, been found in the emitted gases associated with the fertilizer industry, livestock management, coke manufacturing industry, refrigeration methods, and combustion of fossil fuels [104]. In 2009, the United States was the fourth largest producer of ammonia in the world, as it produced 7,704 thousand metric tons of ammonia [105]. In 2002, the US Department of Energy suggested that 65% of the emitted ammonia comes from fertilizer, livestock, and domestic sources, which accounts for nearly 3.25 million tons of ammonia in the atmosphere [106]. Additionally, ammonia is a pollutant as it can have an impact on vegetation and other organisms [107, 108].

The wastewater from some industries as well as residential community is rich in ammonia. Dissociation of ammonia present in the emitted gases as well as wastewater will help in cleaning up the environment along with the production of hydrogen gas. This rationale motivated the researchers at the Center for Electrochemical Engineering Research (CEER) at Ohio University to study ammonia electrolysis [109]. In 2005, Gerardine Botte, director of the CEER, proposed this new technology for the production of hydrogen gas from the electrolysis of ammonia and since then, there have been a consistent development in this technology [110–118]. During ammonia electrolysis, water present in the cathode electrolyte, a 5 M KOH solution, is reduced to hydrogen gas. On the anode side, 1 M ammonia dissolved in the KOH solution is oxidized to nitrogen gas:

Anode reaction

$$\begin{gathered} 2NH_3 + 6OH^- \rightarrow N_2 + 6H_2O + 6e^- \\ E^\circ = -0.77\ \text{V vs. SHE} \end{gathered} \tag{38.100}$$

Cathode reaction

$$2H_2O + 2e^- \rightarrow H_2 + 2OH^- \quad E^\circ = -0.829\ \text{V vs. SHE} \tag{38.101}$$

Overall reaction

$$2NH_3 \rightarrow N_2 + 3H_2 \quad E = 0.059\ \text{V} \tag{38.102}$$

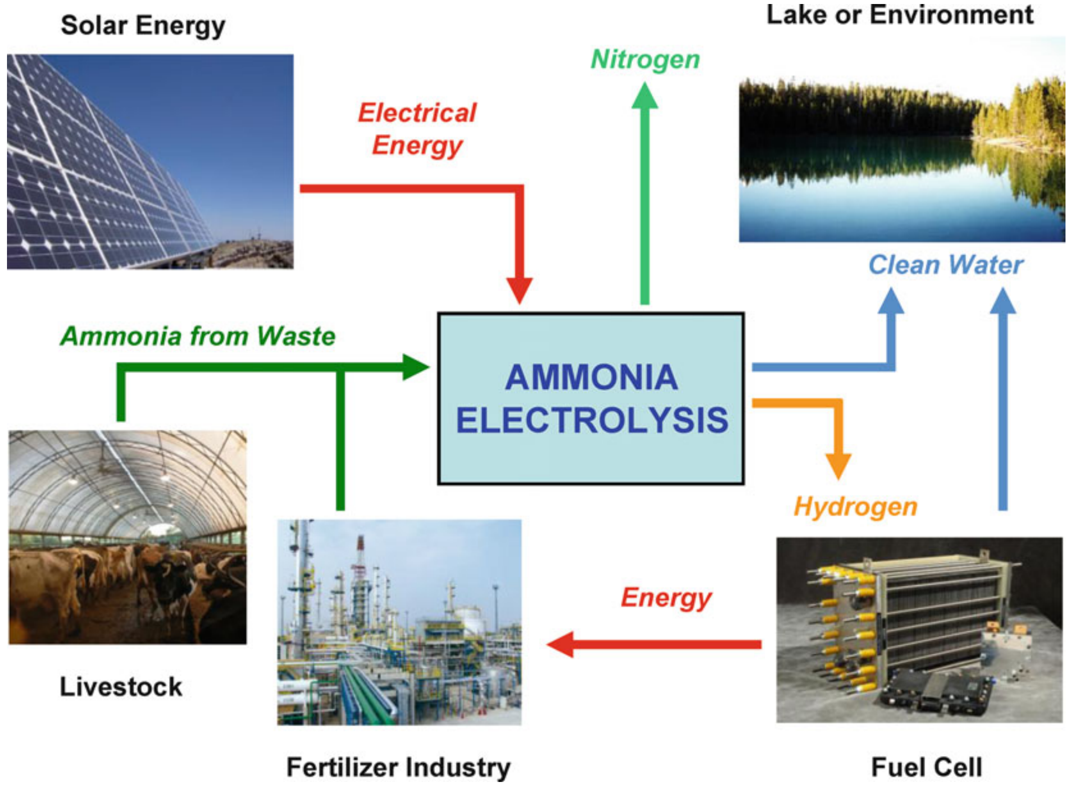


Fig. 38.14 Scheme of ammonia electrolysis along with renewable energy (e.g., solar energy) to remediate ammonia rich wastewater into hydrogen and clean water [143]. The idea of the scheme is to remediate the wastewater using electrolysis to produce hydrogen gas, which could generate additional energy, and also clean water for eventual use in other applications

The electrolysis of ammonia requires a theoretical cell voltage of 0.059 V, which is miniscule in comparison with that of water electrolysis (1.23 V). If solar energy ($0.21/kWh) is used to provide the electric current needed for ammonia electrolysis, then the cost for 1 kg of hydrogen produced from ammonia electrolysis will be $0.90 based on $175/ton of ammonia. In contrast, water electrolysis using electricity from solar energy would produce hydrogen gas at a cost of $7.10/kg. So, the production of hydrogen gas from ammonia electrolysis is less energy consumptive, as well as cheaper, than water electrolysis.

One of the goals of the CEER is to develop a scenario in which an ammonia electrolyzer can be used to clean ammonia rich wastewater in order to produce clean water and hydrogen gas for power generation (Fig. 38.14). In Fig. 38.14, the ammonia rich wastewater from industry, livestock, and residential communities are electrolyzed using the electricity from solar energy or any available energy to produce clean water, which can then be returned to bodies of water, and to produce pure hydrogen gas along with nonpolluting nitrogen gas. The industry or residential community can use the energy obtained from the hydrogen gas from fuel cells thereby reducing the demand for oil and natural gas. In addition to hydrogen production, the electrolysis of ammonia rich wastewater eliminates a pollutant from wastewater and produces clean water, which could be used for various activities. The technology represents a groundbreaking solution for the denitrification of water in wastewater treatment plants.

Urea Electrolysis

Urea is an important chemical for the mankind as it is used in the production of fertilizer. At the same time, urea is found in wastewater, such as agricultural runoff, as well as in human and animal urine. Even industrial wastewater, especially the wastewater from the fertilizer industry, has relatively high concentrations of urea. The chemical structure of urea (NH_2–CO–NH_2) has two amine groups attached by a carbonyl group. The presence of the amine group in urea was beneficial for CEER researchers in order to extend the ammonia electrolysis principle to the electrochemical oxidation of urea [119, 120]. Cleaning the urea rich wastewater is important for the disposal of the wastewater in bodies of water. Using the electrochemical approach to remove urea from wastewater has an added advantage of producing high value chemicals, such as hydrogen.

The anode and cathode reactions for urea electrolysis are given below, where 0.33 M urea along with 5 M KOH forms the anode side electrolyte solution and 5 M KOH forms the cathode electrolyte.

Anode reaction

$$CO(NH_2)_2 + 6OH^- \rightarrow N_2 + 5H_2O + CO_2 + 6e^- \quad E^\circ = -0.46 \text{ V vs. SHE} \tag{38.103}$$

Cathode reaction

$$2H_2O + 2e^- \rightarrow H_2 + 2OH^- \quad E^\circ = -0.83 \text{ V vs. SHE} \tag{38.104}$$

Overall reaction

$$CO(NH_2)_2 + H_2O \rightarrow N_2 + 3H_2 + CO_2 \quad E = 0.37 \text{ V} \tag{38.105}$$

The electrocatalyst used in the anode electrode for the urea oxidation is nickel, which is oxidized to nickel hydroxide ($Ni(OH)_2$) in the alkaline solution, which can be KOH. The theoretical cell voltage for urea electrolysis is 0.37 V, which is 70% less energy consuming than water electrolysis [121]. The urea rich wastewater can be remediated using this electrolysis method, thereby producing hydrogen gas, an energy carrier and environmentally harmless nitrogen gas. CEER researchers are working towards developing an electrolyzer which can directly oxidize urine into hydrogen and nitrogen gas. The urea electrolyzer can be used in place of the ammonia electrolyzer in the scheme displayed in Fig. 38.14 in order to demonstrate the possibility of cleaning urea rich wastewater into clean water and producing valuable product as hydrogen gas [119–125].

Coal Electrolysis

In the world today, hydrogen gas is typically being produced from fossil fuels, including coal, using the gasification technique. An alternate method to utilize coal for hydrogen production is coal electrolysis, which was proposed by Coughlin and Farooque in the late 1970s, but their research was discontinued due to low oxidation currents [126–130]. Recently, Botte and her collaborators at CEER revitalized and significantly improved the electrolysis of coal for the production of hydrogen [131–139].

The production of hydrogen gas using the electrolysis technique described until now has been based on the use of an alkaline electrolyte. However, coal electrolysis uses sulfuric acid as its electrolyte; it therefore can be considered an acid electrolysis system. Electrolysis of coal is an excellent alternative method to the coal gasification technique mainly because of the high energy required under the gasification process (800°C) along with the required purification steps in order to obtain highly pure hydrogen gas. On the other hand, coal electrolysis uses 4 M sulfuric acid circulating in the cathode side, which is separated from the anode side by a membrane. The anode side electrolyte is composed of coal particles mixed with 4 M sulfuric acid as coal slurry. The electrochemical reactions at the anode and cathode electrodes take place at intermediate temperatures (40–108°C):

Anode reaction

$$C + 2H_2O \rightarrow CO_2 + 4H^+ + 4e^- \quad E^\circ = -0.21 \text{ V vs. SHE} \tag{38.106}$$

Cathode reaction

$$4H^+ + 4e^- \rightarrow 2H_2 \quad E^\circ = 0 \text{ V vs. SHE} \tag{38.107}$$

Overall reaction

$$C + 2H_2O \rightarrow 2H_2 + CO_2 \quad E = 0.21 \text{ V} \tag{38.108}$$

The theoretical cell voltage for the overall reaction from coal electrolysis is 0.21 V, which is less than the 1.23 V value associated with water electrolysis. The electrochemical cell used at CEER is known as a continuous electrolytic cell (CEC). Electrolytes for both the anode and cathode side are constantly circulated in order to oxidize a high percentage of the coal particles. The research at CEER has found that the addition of Fe(II)/Fe(III) redox couple to the coal slurry solution increases the rate for the electrochemical oxidation of coal [134, 136, 139]. The CEC used for coal electrolysis has a membrane separating the anode side from the cathode, which has resulted in almost 100% pure hydrogen gas from the cathode compartment. This approach eliminates the need for gas purification equipments.

Once the coal particles have been used in the electrolysis process, the electrolyzed coal is subjected to extraction of liquid chemicals. These liquid chemicals are expected to have high amounts of various hydrocarbons with properties of fuels so as to be termed as liquid fuels. The electrolysis of coal helps in the production of hydrogen gas, a high value chemical, and the solid waste from the electrolysis is further processed to obtain liquid fuels [133].

Summary

The concept of storing energy is as important as generation of the energy itself. Even though there are various methods of storing energy, the most promising methods will be based on electrochemical storage systems. Electrochemical energy storage devices are not limited to batteries, flow batteries, fuel cells, and supercapacitors. The advancement in rechargeable batteries, such as lithium-based batteries, has improved their usage as electrochemical energy storage devices for varying applications. There are new types of batteries being developed, namely high temperature and flow batteries, which has the potential to be used for large-scale power generation or storage units. Similarly, there are promising developments in the field of fuel cells and supercapacitors to improve the capacity to generate power and store energy, respectively, from these systems.

The idea of storing energy can be extended to the synthesis of the chemicals as well, in which the chemicals are generally used for producing power or energy. In this manner, during the synthesis of the chemicals the energy is stored in the chemical bonds of the compounds formed. On that basis, the electrochemical synthesis of the chemical compounds has been around for more than a century. Synthesis of chemical compounds using electrochemical processes provides precise control over the reaction, which results in higher yield and purity.

In the beginning, the manufacturing of chemicals using electrochemical methods was highly energy consuming and less efficient. These challenges meant that industrial electrosynthesis could hardly be considered a process for energy storage. In the earlier days, the focus of the industry and even the accompanying research was to improve the yield and purity of the chemical compounds produced. Over the years, there has been development and advancement in every aspect of the electrochemical and electrochemical engineering field as well as the other areas of science and technology. With the passage of time, improvements from all of the related fields have been incorporated at different levels (including nanotechnology) of the industrial scale electrosynthesis technology mainly to lower energy requirement and also to increase the yield of the products.

In the modern era, energy requirement for the synthesis processes along with the health and environmental impact of the synthesis routes have equal or even more importance than yield of the product as in earlier times. Consequently, more time and money are spent in the research and development process to improve the synthesis technique using the knowledge, including information obtained from other fields of science, which has resulted in a true multidisciplinary research. This research has focused on understanding the reaction mechanism using available analytical methods even under *in-situ* conditions, developing advance materials through nanotechnology and nanoscience, and systematic use of experimental studies at the laboratory, bench scale, and pilot level, along with theoretical investigations using computational and multiscale modeling methods. This approach to using multidisciplinary research and the involvement of new technologies reflect an encouraging and promising future for electrochemical energy storage systems.

Acknowledgments The authors will like to acknowledge the financial support provided by the Center for Electrochemical Engineering Research (CEER) at Ohio University.

References

1. ENERDATA (2010) World energy demand down for the first time in 30 years. Available at: http://www.enerdata.net/enerdatauk/press-and-publication/publications/. Accessed 1 Aug 2011
2. IEA (2010) Key World Energy Statistics 2010. International Energy Agency. Available at: http://www.iea.org/textbase/nppdf/free/2010/key_stats_2010.pdf. Accessed 1 Aug 2011
3. ENERDATA (2011) Enerdata releases its 2011 edition of its Global Energy Statistical Yearbook. Available at: http://www.enerdata.net/enerdatauk/press-and-publication/publications/. Accessed 1 Aug 2011
4. Energimyndigheten (2010) Energy in Sweden—facts and figures 2010. Swedish Energy Agency. Available at: http://www.energimyndigheten.se/en/. Accessed 1 Aug 2011
5. The World Watch Institute (2009) State of the World 2009: Into a Warming World. Available at: http://www.worldwatch.org/. Accessed 1 Aug 2011
6. Wikipedia (2011) Energy Storage. Available at: http://en.wikipedia.org/wiki/Energy_storage. Accessed 7 Aug 2011
7. Botte GG (2009) Vision of the Center for Electrochemical Engineering Research, Ohio University. Available at: http://www.ohio.edu/ceer/research/index.cfm. Accessed 1 Aug 2011
8. McIntyre J (2002) 100 years of industrial electrochemistry. J Electrochem Soc 149:S79–S83
9. Richards JW (1902) A University Course in Electrochemistry, Trans Am Electrochem Soc 1:42
10. Newman JS, Thomas-Alyea KE (2004) Electrochemical systems, 3rd edn. Wiley—Interscience, New York
11. Bard AJ, Faulkner LR (2001) Electrochemical methods: fundamentals and applications, 2nd edn. Wiley, New York
12. Pletcher D, Walsh F (1990) Industrial electrochemistry, 2nd edn. Chapman and Hall, New York
13. Botte GG (2007) Batteries: basic principles, technologies, and modeling. In: Bard AJ, Stratmann M (eds) Encyclopedia of electrochemistry: electrochemical engineering, vol 5. Wiley-VCH, New York, pp 377–423
14. Dobos D (1975) Electrochemical data. Akademiai Kiado, Budapest
15. Schmidt M, Heider U, Kuehner A, Oesten R, Jungnitz M, Ignat'ev N et al (2001) Lithium fluoroalkylphosphates: a new class of conducting salts for electrolytes for high energy lithium-ion batteries. J Power Sources 97–8:557–560
16. Gores HJ, Barthel JMG (1995) Nonaqueous electrolyte-solutions—new materials for devices and processes based on recent applied-research. Pure Appl Chem 67:919–930
17. Hives J, Thonstad J, Sterten A, Fellner P (1996) Electrical conductivity of molten cryolite-based mixtures obtained with a tube-type cell made of pyrolytic boron nitride. Metall Mater Trans B Proc Metall Mater Proc Sci 27:255–261
18. Yamamoto O (2000) Solid oxide fuel cells: fundamental aspects and prospects. Electrochim Acta 45:2423–2435
19. Haynes WM (2011) CRC handbook of chemistry and physics (Internet Version 2012), 92nd edn. CRC Press/Taylor and Francis, Boca Raton, FL
20. Linden D (1995) Handbook of batteries, 2nd edn. McGraw-Hill, Inc., New York
21. Zeng K, Zhang DK (2010) Recent progress in alkaline water electrolysis for hydrogen production and applications. Prog Energy Combustion Sci 36:307–326

22. Tarcy GP, Kvande H, Tabereaux A (2011) Advancing the industrial aluminum process: 20th century breakthrough inventions and developments. JOM 63:101–108
23. Bard AJ, Inzelt G, Scholz F (2008) Electrochemical dictionary. Springer, Berlin
24. Ibrahim H, Ilinca A, Perron J (2008) Energy storage systems—characteristics and comparisons. Renew Sustain Energy Rev 12:1221–1250
25. Garche J (2001) Advanced battery systems—the end of the lead-acid battery? Phys Chem Chem Phys 3:356–367
26. de Leon CP, Frias-Ferrer A, Gonzalez-Garcia J, Szanto DA, Walsh FC (2006) Redox flow cells for energy conversion. J Power Sources 160:716–732
27. Scrosati B, Garche J (2010) Lithium batteries: status, prospects and future. J Power Sources 195:2419–2430
28. Naoi K (2010) 'Nanohybrid capacitor': the next generation electrochemical capacitors. Fuel Cells 10:825–833
29. Osaka T, Datta M (eds) (2000) Energy storage systems for electronics. Gordon and Breach Science Publishers, Singapore
30. Larminie J, Dicks A (2003) Fuel cell systems explained, 2nd edn. Wiley, West Sussex, England
31. Carrette L, Friedrich KA, Stimming U (2001) Fuel cells—fundamentals and applications. Fuel Cells 1:5–39
32. Srinivasan S, Mosdale R, Stevens P, Yang C (1999) Fuel cells: reaching the era of clean and efficient power generation in the twenty-first century. Annu Rev Energy Environ 24:281–328
33. Chen J, Cheng FY (2009) Combination of lightweight elements and nanostructured materials for batteries. Acc Chem Res 42:713–723
34. Zhou HB, Huang QM, Liang M, Lv DS, Xu MQ, Li H et al (2011) Investigation on synergism of composite additives for zinc corrosion inhibition in alkaline solution. Mater Chem Phys 128:214–219
35. Bailey MR, Donne SW (2011) Electrochemical impedance spectroscopy study into the effect of titanium dioxide added to the alkaline manganese dioxide cathode. J Electrochem Soc 158: A802–A808
36. Minakshi M, Ionescu M (2010) Anodic behavior of zinc in Zn-MnO_2 battery using ERDA technique. Int J Hydrogen Energy 35:7618–7622
37. Pan JQ, Sun YZ, Wang ZH, Wan PY, Fan MH (2009) Mn_3O_4 doped with nano-$NaBiO_3$: a high capacity cathode material for alkaline secondary batteries. J Alloys Compd 470:75–79
38. Raghuveer V, Manthiram A (2006) Role of TiB_2 and Bi_2O_3 additives on the rechargeability of MnO_2 in alkaline cells. J Power Sources 163:598–603
39. Beck F, Ruetschi P (2000) Rechargeable batteries with aqueous electrolytes. Electrochim Acta 45:2467–2482
40. Wen YH, Cheng J, Ning SQ, Yang YS (2009) Preliminary study on zinc-air battery using zinc regeneration electrolysis with propanol oxidation as a counter electrode reaction. J Power Sources 188:301–307
41. Goldstein J, Brown I, Koretz B (1999) New developments in the Electric Fuel Ltd zinc air system. J Power Sources 80:171–179
42. Dell RM, Rand DAJ (2004) Clean energy. Royal Society of Chemistry, Cambridge, UK
43. Rand DAJ, Woods R, Dell RM (1998) Batteries for electric vehicles. Research Studies Press Ltd., Somerset, England
44. Morioka Y, Narukawa S, Itou T (2001) State-of-the-art of alkaline rechargeable batteries. J Power Sources 100:107–116
45. Shukla AK, Venugopalan S, Hariprakash B (2001) Nickel-based rechargeable batteries. J Power Sources 100:125–148
46. Patil A, Patil V, Shin DW, Choi JW, Paik DS, Yoon SJ (2008) Issue and challenges facing rechargeable thin film lithium batteries. Mater Res Bull 43:1913–1942
47. Fergus JW (2010) Ceramic and polymeric solid electrolytes for lithium-ion batteries. J Power Sources 195:4554–4569
48. Shukla AK, Kumar TP (2008) Materials for next-generation lithium batteries. Curr Sci 94:314–331
49. Lu XC, Xia GG, Lemmon JP, Yang ZG (2010) Advanced materials for sodium-beta alumina batteries: status, challenges and perspectives. J Power Sources 195:2431–2442
50. Weber AZ, Mench MM, Meyers JP, Ross PN, Gostick JT, Liu QH (2011) Redox flow batteries: a review. J Appl Electrochem 41:1137–1164
51. Mellentine JA, Culver WJ, Savinell RF (2011) Simulation and optimization of a flow battery in an area regulation application. J Appl Electrochem 41:1167–1174
52. Aaron D, Tang ZJ, Papandrew AB, Zawodzinski TA (2011) Polarization curve analysis of all-vanadium redox flow batteries. J Appl Electrochem 41:1175–1182
53. Wu XW, Yamamura T, Ohta S, Zhang QX, Lv FC, Liu CM et al (2011) Acceleration of the redox kinetics of VO^{2+}/VO_2^+ and V^{3+}/V^{2+} couples on carbon paper. J Appl Electrochem 41:1183–1190
54. Shinkle AA, Sleightholme AES, Thompson LT, Monroe CW (2011) Electrode kinetics in non-aqueous vanadium acetylacetonate redox flow batteries. J Appl Electrochem 41:1191–1199
55. Kim S, Tighe TB, Schwenzer B, Yan JL, Zhang JL, Liu J et al (2011) Chemical and mechanical degradation of sulfonated poly (sulfone) membranes in vanadium redox flow batteries. J Appl Electrochem 41:1201–1213
56. Zhang JL, Li LY, Nie ZM, Chen BW, Vijayakumar M, Kim S et al (2011) Effects of additives on the stability of electrolytes for all-vanadium redox flow batteries. J Appl Electrochem 41:1215–1221
57. Menictas C, Skyllas-Kazacos M (2011) Performance of vanadium-oxygen redox fuel cell. J Appl Electrochem 41:1223–1232
58. Skyllas-Kazacos M, Milne N (2011) Evaluation of iodide and titanium halide redox couple combinations for common electrolyte redox flow cell systems. J Appl Electrochem 41:1233–1243
59. Zhang R, Weidner JW (2011) Analysis of a gas-phase Br_2-H_2 redox flow battery. J Appl Electrochem 41:1245–1252
60. Kiros Y (1996) Electrocatalytic properties of Co, Pt, and Pt-Co on carbon for the reduction of oxygen in alkaline fuel cells. J Electrochem Soc 143:2152–2157
61. Chrzanowski W, Wieckowski A (1998) Surface structure effects in platinum/ruthenium methanol oxidation electrocatalysis. Langmuir 14:1967–1970
62. Gasteiger HA, Markovic N, Ross PN, Cairns EJ (1994) CO electrooxidation on well-characterized Pt-Ru alloys. J Phys Chem 98:617–625
63. Gasteiger HA, Markovic N, Ross PN, Cairns EJ (1994) Electrooxidation of small organic-molecules on well-characterized Pt-Ru alloys. Electrochim Acta 39:1825–1832
64. Petukhov AV, Akemann W, Friedrich KA, Stimming U (1998) Kinetics of electrooxidation of a CO monolayer at the platinum/electrolyte interface. Surf Sci 402:182–186
65. Friedrich KA, Geyzers KP, Linke U, Stimming U, Stumper J (1996) CO adsorption and oxidation on a Pt(111) electrode modified by ruthenium deposition: an IR spectroscopic study. J Electroanal Chem 402:123–128
66. Alonso-Vante N, Tributsch H, Solorza-Feria O (1995) Kinetics studies of oxygen reduction in acid-medium on novel semiconducting transition-metal chalcogenides. Electrochim Acta 40:567–576
67. Solorza-Feria O, Ellmer K, Giersig M, Alonso-Vante N (1994) Novel low-temperature synthesis of semiconducting transition metal chalcogenide electrocatalyst for multielectron charge transfer: molecular oxygen reduction. Electrochim Acta 39:1647–1653

68. Dong SJ, Qiu QS (1991) Electrodeposition of platinum particles on glassy-carbon modified with cobalt porphyrin and Nafion film and their electrocatalytic reduction of dioxygen. J Electroanal Chem 314:223–239
69. Gupta S, Tryk D, Zecevic SK, Aldred W, Guo D, Savinell RF (1998) Methanol-tolerant electrocatalysts for oxygen reduction in a polymer electrolyte membrane fuel cell. J Appl Electrochem 28:673–682
70. DOE (2011) Phosphoric Acid Fuel Cell Technology. Available at: http://www.fossil.energy.gov/programs/powersystems/fuelcells/fuelscells_phosacid.html. Accessed 31 Oct 2011
71. FCTec (2011) Phosphoric Acid Fuel Cells (PAFC). Available at: http://www.fctec.com/fctec_types_pafc.asp. Accessed 31 Oct 2011
72. Yuh C, Johnsen R, Farooque M, Maru H (1995) Status of carbonate fuel-cell materials. J Power Sources 56:1–10
73. Badwal SPS, Giddey S, Ciacchi FT (2006) Hydrogen and oxygen generation with polymer electrolyte membrane (PEM)-based electrolytic technology. Ionics 12:7–14
74. Armaroli N, Balzani V (2011) The hydrogen issue. ChemSusChem 4:21–36
75. Farrauto R, Hwang S, Shore L, Ruettinger W, Lampert J, Giroux T et al (2003) New material needs for hydrocarbon fuel processing: generating hydrogen for the PEM fuel cell. Annu Rev Mater Res 33:1–27
76. Onsan ZI (2007) Catalytic processes for clean hydrogen production from hydrocarbons. Turk J Chem 31:531–550
77. Palo DR, Dagle RA, Holladay JD (2007) Methanol steam reforming for hydrogen production. Chem Rev 107:3992–4021
78. Ogden JM (1999) Prospects for building a hydrogen energy infrastructure. Annu Rev Energy Environ 24:227–279
79. Navarro RM, Pena MA, Fierro JLG (2007) Hydrogen production reactions from carbon feedstocks: fossils fuels and biomass. Chem Rev 107:3952–3991
80. Longwell JP, Rubin ES, Wilson J (1995) Coal: energy for the future. Prog Energy Combustion Sci 21:269–360
81. Simon P, Gogotsi Y (2008) Materials for electrochemical capacitors. Nat Mater 7:845–854
82. Amatucci GG, Badway F, DuPasquier A (2000) Novel asymmetric hybrid cells and the use of pseudo-reference electrodes in three electrode cell characterization. The Electrochemical Society, Pennington, New Jersey, USA. In: Nazri GA, Thackeray M, Ohzuku T (eds) Intercalation Compounds for Battery Materials, Proceedings, vol 99. pp 344–359
83. Kuhn AT (1971) Industrial electrochemical processes. Elsevier Science Limited, Amsterdam
84. Bommaraju TV, O'Brien TF, Hine F (2005) Handbook of chlor-alkali technology. Springer Science+Business Media, Inc., New York
85. Moussallem I, Jorissen J, Kunz U, Pinnow S, Turek T (2008) Chlor-alkali electrolysis with oxygen depolarized cathodes: history, present status and future prospects. J Appl Electrochem 38:1177–1194
86. Venkatesh S, Tilak BV (1983) Chlor-alkali technology. J Chem Educ 60:276–278
87. Haupin WE (1983) Electrochemistry of the Hall-Heroult process for aluminum smelting. J Chem Educ 60:279–282
88. Ferreira BK (2008) Three-dimensional electrodes for the removal of metals from dilute solutions: a review. Mineral Proc Extr Metall Rev 29:330–371
89. Cooper WC (1985) Reviews of applied electrochemistry 11. Advances and future-prospects in copper electrowinning. J Appl Electrochem 15:789–805
90. Leroy RL (1983) Industrial water electrolysis—present and future. Int J Hydrogen Energy 8:401–417
91. Abe R (2010) Recent progress on photocatalytic and photoelectrochemical water splitting under visible light irradiation. J Photochem Photobiol C Photochem Rev 11:179–209
92. Utley J (1997) Trends in organic electrosynthesis. Chem Soc Rev 26:157–167
93. Sequeira CAC, Santos DMF (2009) Electrochemical routes for industrial synthesis. J Braz Chem Soc 20:387–406
94. Srinivasan V, Arora P, Ramadass P (2006) Report on the electrolytic industries for the year 2004. J Electrochem Soc 153:K1–K14
95. Dukes RR (1970) Report of electrolytic industries for year 1968. J Electrochem Soc 117:C9–C14
96. Electrochemistry Encyclopedia (2002) Industrial Organic Electrosynthesis. Available at: http://electrochem.cwru.edu/encycl/art-o01-org-ind.htm. Accessed 5 Sept 2011
97. Sivula K, Le Formal F, Gratzel M (2011) Solar water splitting: progress using hematite (α-Fe_2O_3) photoelectrodes. ChemSusChem 4:432–449
98. Boddy PJ (1968) Oxygen evolution on semiconducting TiO_2. J Electrochem Soc 115:199
99. Fujishima A, Honda K (1972) Electrochemical photolysis of water at a semiconductor electrode. Nature 238:37–38
100. Murphy AB, Barnes PRF, Randeniya LK, Plumb IC, Grey IE, Horne MD et al (2006) Efficiency of solar water splitting using semiconductor electrodes. Int J Hydrogen Energy 31:1999–2017
101. Gray HB (2009) Powering the planet with solar fuel. Nat Chem 1:7
102. CCI Solar (2011) Center for Chemical Innovation (CCI Solar), California Institute of Technology. Available at: http://ccisolar.caltech.edu. Accessed 28 Sept 2011
103. Barbir F (2005) PEM electrolysis for production of hydrogen from renewable energy sources. Solar Energy 78:661–669
104. Phillips J (1995) Control and Pollution Prevention Options for Ammonia Emissions. Technical Rpt. EPA-456/R-95-002, ViGYAN Incorporated, Research Triangle Park, NC
105. Index Mundi (2011) Ammonia: Estimated World Production, By Country. Available at: http://www.indexmundi.com/en/commodities/minerals/nitrogen/nitrogen_t12.html. Accessed 30 Oct 2011
106. Mansell GE (2005) An improved ammonia inventory for the WRAP Domain. Technical, ENVIRON International Corporation Novato, California, USA.
107. Bouwman AF, Lee DS, Asman WAH, Dentener FJ, VanderHoek KW, Olivier JGJ (1997) A global high-resolution emission inventory for ammonia. Global Biogeochem Cycles 11:561–587
108. Sommer SG, Hutchings NJ (2001) Ammonia emission from field applied manure and its reduction—invited paper. Eur J Agron 15:1–15
109. CEER (2009) Center for Electrochemical Engineering Research (CEER), Ohio University. Available at: http://www.ohio.edu/ceer/. Accessed 1 Aug 2011
110. Botte GG, Vitse F, Cooper M (2009) Electro-catalysts for the oxidation of ammonia in alkaline media and its application to hydrogen production, ammonia fuel cells, ammonia electrochemical sensors, and purification process for ammonia-contained effluents. United States, US 7,485,211
111. Botte GG (2010) Electro-catalysts for the oxidation of ammonia in alkaline media. United States, US 7,803,264
112. Botte GG (2009) Electrochemical method for providing hydrogen using ammonia and ethanol. United States, Patent Pending US 2009/0050489
113. Botte GG (2010) Carbon fiber-electrocatalysts for the oxidation of ammonia, and ethanol in alkaline media and their application to

hydrogen production, fuel cells, and purification processes. United States, Patent Pending WO 2007/047630

114. Vitse F, Cooper M, Botte GG (2005) On the use of ammonia electrolysis for hydrogen production. J Power Sources 142:18–26
115. Cooper M, Botte GG (2006) Hydrogen production from the electro-oxidation of ammonia catalyzed by platinum and rhodium on raney nickel substrate. J Electrochem Soc 153:A1894–A1901
116. Bonnin EP, Biddinger EJ, Botte GG (2008) Effect of catalyst on electrolysis of ammonia effluents. J Power Sources 182:284–290
117. Boggs BK, Botte GG (2009) On-board hydrogen storage and production: an application of ammonia electrolysis. J Power Sources 192:573–581
118. Boggs BK, Botte GG (2010) Optimization of Pt-Ir on carbon fiber paper for the electro-oxidation of ammonia in alkaline media. Electrochim Acta 55:5287–5293
119. Botte GG (2008) Urea electrolysis. United States, Provisional Patent US 61/104,478
120. Botte GG (2009) Electrolytic cells and methods for the production of ammonia and hydrogen. United States, Patent Pending US 2009/0095636
121. Boggs BK, King RL, Botte GG (2009) Urea electrolysis: direct hydrogen production from urine. Chem Commun 4859–4861
122. Daramola DA, Singh D, Botte GG (2010) Dissociation rates of urea in the presence of NiOOH catalyst: a DFT analysis. J Phys Chem A 114:11513–11521
123. King RL, Botte GG (2011) Hydrogen production via urea electrolysis using a gel electrolyte. J Power Sources 196:2773–2778
124. Wang D, Yan W, Botte GG (2011) Exfoliated nickel hydroxide nanosheets for urea electrolysis. Electrochem Commun 13:1135–1138
125. King RL, Botte GG (2011) Investigation of multi-metal catalysts for stable hydrogen production via urea electrolysis. J Power Sources 196:9579–9584
126. Coughlin RW, Farooque M (1979) Hydrogen production from coal, water and electrons. Nature 279:301–303
127. Farooque M, Coughlin RW (1979) Electrochemical gasification of coal (investigation of operating-conditions and variables). Fuel 58:705–712
128. Coughlin RW, Farooque M (1980) Electrochemical gasification of coal—simultaneous production of hydrogen and carbon-dioxide by a single reaction involving coal, water, and electrons. Ind Eng Chem Proc Design Dev 19:211–219
129. Coughlin RW, Farooque M (1980) Consideration of electrodes and electrolytes for electrochemical gasification of coal by anodic-oxidation. J Appl Electrochem 10:729–740
130. Coughlin RW, Farooque M (1982) Thermodynamic, kinetic, and mass balance aspects of coal-depolarized water electrolysis. Ind Eng Chem Proc Design Dev 21:559–564
131. Botte GG (2006) Electrocatalysts and additives for the oxidation of solid fuels and their application to hydrogen production, fuel cells, and water remediation processes. United States, Patent Pending WO 2006/121981
132. Botte GG, Jin X (2010) Electrochemical technique to measure concentration of multivalent cations simultaneously. United States, Patent Pending WO 2007/133534
133. Botte GG (2011) Pretreatment method for the synthesis of carbon nanotubes and carbon nanostructures from coal and carbon chars. United States, US 8,029,759
134. Patil P, De Abreu Y, Botte GG (2006) Electrooxidation of coal slurries on different electrode materials. J Power Sources 158:368–377
135. Sathe N, Botte GG (2006) Assessment of coal and graphite electrolysis on carbon fiber electrodes. J Power Sources 161:513–523
136. De Abreu Y, Patil P, Marquez AI, Botte GG (2007) Characterization of electrooxidized Pittsburgh No. 8 Coal. Fuel 86:573–584
137. Jin X, Botte GG (2007) Feasibility of hydrogen production from coal electrolysis at intermediate temperatures. J Power Sources 171:826–834
138. Jin X, Botte GG (2009) Electrochemical technique to measure Fe (II) and Fe(III) concentrations simultaneously. J Appl Electrochem 39:1709–1717
139. Jin X, Botte GG (2010) Understanding the kinetics of coal electrolysis at intermediate temperatures. J Power Sources 195:4935–4942
140. Lu XC, Lemmon JP, Sprenkle V, Yang ZG (2010) Sodium-beta alumina batteries: status and challenges. JOM 62:31–36
141. Wikipedia (2011) Electric double-layer capacitor. Available at: http://en.wikipedia.org/wiki/File:Supercapacitor_diagram.svg. Accessed 5 July 2011
142. Roeb M, Neises M, Monnerie N, Sattler C, Pitz-Paal R (2011) Technologies and trends in solar power and fuels. Energy Environ Sci 4:2503–2511
143. Muthuvel M, Botte GG (2009) Trends in ammonia electrolysis. In: White RE (ed) Modern aspects of electrochemistry. Springer Science+Business Media, Inc., New York, pp 207–245

Index

A
Accuracy, definition, 247
Acetaldehyde, 94, 273, 279, 312, 313, 321, 335, 342–344, 347, 349, 363–366, 750, 800, 990, 1053, 1165, 1483
Acetamide, 649
Acetic acid, 38, 66, 96, 312, 313, 322, 327, 328, 335, 336, 340, 342–344, 347–351, 366, 369–371, 373, 432–434, 437, 451, 469, 477, 506, 801, 893, 925, 1053, 1077–1079, 1093, 1131, 1134, 1152, 1154–1156, 1259, 1260, 1286, 1295, 1296
  via methanol carbonation, 342, 349
Acetic anhydride, 83, 327, 330, 342, 347, 350–351, 370, 374, 432–434, 800, 801, 1053, 1093, 1096, 1481
  from coal, 10
Acetone, 12, 61, 72, 181, 254, 309, 314, 320, 328–330, 335, 338–340, 344, 351, 352, 363, 364, 367, 370, 375, 377, 379–381, 431, 432, 434–436, 448, 450, 467, 477, 530, 564, 626, 923, 924, 929, 1053, 1107, 1131, 1161, 1163, 1164, 1168, 1175, 1213, 1291, 1365, 1370, 1450, 1481, 1483, 1531
Acetone cyanohydrin, 181, 344, 564, 929, 930
Acetonitrile, 244, 256, 928, 929
Acetosolv process, 1079, 1286
Acetylene, 19, 61, 69, 111, 309, 313, 316, 321, 326–328, 347, 358, 366–367, 446, 450, 626, 634, 800, 824, 1039, 1043, 1051–1053, 1480
Acid dyes, acid black, 455, 469, 479–483, 491, 503, 519–522, 529–530
Acid gases, 279, 312, 723, 807, 893, 894
Acrolein, 66, 94, 335, 344–345, 348, 366, 563, 929, 1163
Acrylamide, 36, 83, 314, 447, 932, 1204, 1213–1215, 1218, 1283
Acrylic acid, 313, 326, 335, 344–345, 366, 370, 378, 563, 564, 801, 1129, 1154
Acrylic adhesives, wax in, 566
Acrylic monomers, in adhesives, 563, 564, 567
  solvency, 564
Acrylic polymer manufacture
  acrylonitrile, 446, 447
  spinning, 448–449
Adhesion
  to dirty metals, 566
  factors affecting, 549
  hydrogen bonding, 549
  interactions, 549
  low energy surfaces, 550
  mechanisms, 549
  surface treatments to improve, 550
Adhesives
  acrylic, 547–568
  bonded construction, 550
  epoxy, 547–568
  primers, surface treatment, 550
  urethane, 547–568
Adipic acid, 330, 335, 345, 373, 376, 436–438, 440, 468, 477, 910, 1154, 1160, 1173, 1291
Adiponitrile, 10, 314, 322, 437, 438, 929, 930, 1530, 1531
Adsorbents, solid phase, 241–243
Aerogels, 257–260, 272, 273
Aerosols, 10, 20, 92, 167, 172, 174, 179, 241, 246, 293, 330, 467, 658, 660, 661, 770, 876, 1053, 1295, 1298, 1304, 1391
Affinity chromatography (AC), 1243–1246
Agrochemicals
  alleopathy,
  aryloxyphenoxypropionic acid, 646
  azoles, 409
  benzoyl ureas, 645, 671, 674
  biological agents, 658
  biological agents carbonates, 653
  chirality, role of chloroacetanilides, 656
  definition, 661
  delivery systems, 665
  diarylether carboxylic acids,
  dinitroanilines, 649
  insect growth regulators, 674
  types, 659, 660
Agrochemicals, future directions
  industry, characteristics
    chemistry, 654–656
    classes, 643
    history, 641
    manufacture, 641
    regulatory aspects, 664
  insect pheromones, 641, 642
  obsolescence, 321, 334, 357
  organochlorine, 643, 644
  organophosphorous, 656
  pyrethroids, 644, 656, 672
  recombinant DNA technology, 652
  role in agriculture, 643
  sulfonyl ureas, 647, 657, 676–679, 681
  triazenes, 645, 647
  types
    algicides,
    bactericides, 642
    fumigants, 642
    fungicides, 646–654
    herbicides, 643–646
    insect growth regulators, 674
    insecticides, 642–643
    molluscicides, 642
    nematicides, 642
    plant growth regulators, 642
    types, rodenticides, 661
  ureas, 657
  usage, 641
Agrochemicals industry, classical chemicals, 643

J.A. Kent (ed.), *Handbook of Industrial Chemistry and Biotechnology*,
DOI 10.1007/978-1-4614-4259-2, © Springer Science+Business Media New York 2012

Air
  ambient, 231, 232, 241–243, 886, 1046, 1363
  containing volatile compounds, 242
  purification, 201, 274–275, 278–279
  separation plant (flow diagram), 800, 895, 1039
Alchemix,
Alkyllithium catalyst for polymerization, 627, 629
Allophanate, in urethane adhesive, 560
Alpha amylase, 1165, 1166, 1423
AlphaScreen, 1234
Alumina (Al_2O_3), 28, 71, 192, 203, 225, 242, 252, 253, 257, 260, 274, 276, 291, 317, 332, 341, 342, 357, 359, 376, 379, 585, 601, 723, 736, 789, 792, 817, 822, 825, 892, 1039, 1206, 1298, 1363, 1501, 1504, 1515, 1516, 1520, 1527, 1528
American Institute of Chemical Engineers, 48–50, 78, 85–87, 99, 101, 113, 128, 181
American Public Health Association, 237
American Society for Testing and Materials (ASTM), 83, 237, 281, 283, 295, 419, 659, 749, 750, 754, 1097, 1261, 1265
American Water Works Association (AWWA), 237
Amines
  alicyclic, 916
  aliphatic, 916
  alkyl, 918, 919
  aromatic as curatives in epoxy adhesives, 555
  butyl, 919
  cyclohexyl, 918, 919
  dimethyl, 94, 366, 917, 918, 1456
  heterocyclic, 516
  methyl diethanolamine, 893, 894, 1050
  t-octyl, 918
Amino acid, 406, 589, 647, 651–653, 677–679, 929, 1142, 1148, 1169, 1172, 1173, 1184–1188, 1191, 1193, 1195, 1204, 1214, 1217, 1218, 1223, 1235, 1268, 1325, 1350, 1384
Amino acid analysis, 1268
Aminoacylases, 1204, 1205, 1207, 1217
Aminoglycoside 3-phosphotransferase, 1233
Ammonia
  anhydrous, 755, 877, 879, 880, 900–902, 966–968, 975, 981, 982, 992, 1033
  aqueous, 342, 854, 881, 901, 902, 920, 1076
  catalysts, 897, 898
  converter, 899
  dissociated, 901
  gas, 508, 910, 969, 985
  liquid, 846, 901, 902, 966
  oxidation, 903–906, 908
  pipeline, 902
  plant, 800, 881, 889, 897–900, 913
  processes, 877, 891, 895, 898, 899, 904, 906, 909, 912, 918, 926, 927
  properties, 880
  quality, 901–902
  toxicity, 901
  uses, 900–901
    wood plasticization, 900
Ammonia synthesis, for fertilizer production, 965
Ammonium bisulfate, wood pulping, 1076
Ammonium carbamate, 899, 913
Ammonium carbonate, 849, 854, 895
Ammonium hydroxide, 736, 848, 875, 901, 1076
Ammonium nitrate
  fertilizer, 910
  properties, 912
  Stamicarbon process, 910–911
Ammonium phosphates, 4, 327, 877, 878, 974–976, 982, 983, 1015, 1096, 1097, 1530
  fertilizers, 4, 877, 969, 974, 1014
Ammonium sulfate fertilizer, 901
Amylase, 12, 13, 1165, 1179, 1183, 1184, 1193–1201, 1203, 1411, 1423, 1424, 1465
Amylopectin, 1183, 1184, 1196, 1197, 1423, 1424
Amylose, 1184, 1196, 1197, 1214, 1423, 1424
Anaerobic adhesives, 563, 564
Analysis/analyses
  definition, 247
  elemental, 239, 245
  field, 235, 245
  qualitative, definition, 247
  quantitative, definition, 247
  remote laboratory, 235
Analytes
  characteristics, 237
  definition, 247
  nonvolatile, 238
  sampling techniques, 237
  semivolatile, 238
  volatile, 237–238
Analytical method
  attributes, definition, 247
  chromatographic, 243–244
  colorimetric, 245
  compromises, 243
  definition, 247
  development, 236
  documentation, 236
  electrochemical, 245, 246
  figures of merit, 236
  Gov't regulations, 236
  immunoassays, 245, 246
  multi-analyte, 234–235
  nonchromatographic, 245–246
  single analyte, 234–235
  standardized, 237
Analytical quality assurance, 236, 237
Anhydrides, as curatives for epoxy adhesives, 557
Aniline, processes, 929–932
Animal-component-free, ACF, 1236
Animal fats
  animal feed uses, 1355
  edible lards and tallows, 138–139
  inedible, 1353
  rendering processes, 1353
  restaurant greases, 1353
  specifications, uses, 1354
Anionic polymers, coupling agent, 578–579
Anode, 84, 85, 207, 286, 700, 726, 1027, 1028, 1032, 1037, 1301, 1499–1503, 1505, 1507–1510, 1513–1522, 1524–1530, 1533, 1535
Anthropogenic substances, 231
Antibiotics, 12, 392, 394, 395, 406, 407, 652, 653, 919, 1131–1133, 1139, 1142, 1143, 1154, 1165, 1175–1178, 1190, 1204, 1214, 1215, 1218, 1221, 1278
  recovery process, 1142
Antidegradants in rubber, 623
Arabinans, arabinofuranosidases, 1265, 1267, 1280, 1281, 1284
Arabinose, 1099, 1100, 1103, 1166, 1201, 1258–1260, 1275, 1281–1284, 1531
Arachidonic acid, 1176
Aramids
  manufacture, 458–459

  meta-aramids, 459–460
  para-aramids, 458
Arithmetic mean (data), 133
Arrhenius equation, 210, 603, 1187
Aspartame, 8, 9, 24, 1171, 1174, 1204, 1213, 1214, 1218, 1219
Association of Official Analytical Chemists (AOAC), 237
Atomic absorption spectrometry (AAS), 245, 876
Atomic emission spectrometry, 245
Automated screening (AS), 1237
Available chlorine, 1035–1037
AZCL polysaccharide hydrolysis, 1283
Azoic dye, 481–482
Azoles, 408

B
Bagasse, 1067, 1166, 1261, 1287, 1403, 1407, 1408, 1411–1412, 1421
Basic dyes, 469, 482–483, 503, 530
Battery/batteries
  alkaline batteries, 1508
  capacity, 1504–1505
  flow batteries, 1514–1515
  high temperature batteries, 1515–1516
  lead acid battery, 1509–1511
  lithium ion battery, 1513–1514
  nickel-cadmium, 1511–1512
  nickel metal hydride, 1512–1513
  zinc-air, 1508–1509
  zinc-chloride, 1507–1508
Beet juice, purification, 1419
Beet sugar, 14, 1403, 1404, 1409, 1410, 1413, 1418, 1420, 1422, 1427, 1428
Beet sugar, juice extraction and processing, 1419
Benzene, chemicals from, 9, 18, 31, 62, 63, 163, 171, 172, 179, 180, 182–184, 211, 216, 217, 274, 275, 310, 312, 313, 316–320, 323, 332–336, 338–340, 345, 346, 352, 358, 374, 376, 380, 436, 437, 441, 464, 505, 508, 509, 511, 513, 526, 530, 627, 653, 672, 700, 707, 718, 721, 750, 755, 770, 799, 807, 894, 910, 919, 930, 932, 1120, 1176, 1299, 1456
Benzopyrene, 238
Benzoic acid, 15, 319, 320, 335, 337–338, 371, 373, 441, 512, 649, 650
Benzyl chloride, 358, 512
Beta amylase, 1193–1196, 1423, 1424
Bias, definition, 247
Biocatalysis
  advantages, 1204
  enzymes, 1184–1204
  non-aqueous, 1212–1213
  products, 1213–1222
Biocatalysts
  discovery, 1222
  engineering, 1190–1192
  immobilization, 1206–1208
  whole cell, 1208–1210
Biodiesel, 92, 191, 192, 217, 1168–1170, 1194, 1202, 1224, 1331, 1332, 1346, 1354, 1373, 1374, 1392–1396
Bioinformatics, 1133, 1135, 1185
Biological organisms, 1251
Biomass
  analysis, 1261–1263, 1265, 1266, 1268–1270, 1273
  biochemical conversion, 1251
  boilers, 1292, 1303
  combustion, 1292, 1294, 1302–1304
  consolidated processing, 1260
  deconstruction, 1273–1292
  lignocellulosic, 1170, 1194, 1201, 1203, 1251, 1252, 1258, 1291
  potential, 1258
  supply, 1258
  thermochemical conversion, 1251–1253, 1292–1305
Biomass gasification
  reaction, 1296–1297
  reactor design, 1297–1298
  syngas
    cleanup, 1299–1300
    high pressure process, 1302
    impurities, 1298–1300
    liquid fuels, 1301–1302
Biomechanical pulping, 1112
Biomolecular engineering, 1134, 1135
Bio-oil, 1105–1107, 1294–1296, 1395
Biopharmaceuticals, 1147, 1152, 1178–1179, 1228
Biopolymers, 35, 593–594, 1097, 1132, 1250–1252, 1284–1286, 1288, 1289
Biopower, 1254, 1255, 1303, 1304
Bioreactor
  aeration, 1239
  batch mode, 1145, 1211
  batch reactors, 1211, 1212, 1239–1240
  biosensors, 1239
  comparison, 1212
  disposable, 1239
  perfusion, 1240, 1242
  plug-flow reactors, 1211–1212
  process control, 1240
  process parameters, 1240
  scale-up, 1242–1243
Biorefinery, 1058, 1097, 1105, 1224, 1253, 1259, 1260, 1284–1286, 1288–1292, 1296
Biotechnology
  approved products, 30, 1112
  bio-bleaching pulp, 1112–1113
  bio-pulping wood, 1112
  emergence, 1135
  for farm productivity, 1330
  genetically modified food, 34, 665
Biotechnology, agriculture/processing
  crop production, 1326
  enzyme use, 1331
  herbicide tolerant crops, 1326–1327
  insect resistant crops, 1326
  light determinate varieties, 1330, 1336, 1337
  transgenic crop concerns, 1327–1329
  transgenic foods labeling, 1328
Biotechnology for farm productivity, regulatory aspects, 1326
Biotherapeutics, 1229
Biphenyl, 239, 497, 652
Bisphenol, 11, 328–330, 362, 367, 551–553, 556, 599, 639
Biuret
  of hexamethylene diisocyanate, 562
  in urethane adhesives, 560
Bleaches, 451, 1035–1037, 1200, 1451, 1467, 1468
  for wood pulp, 1035, 1036
Block copolymers
  from styrene and butadiene, 582
  types, 582
Borol solution, 1025
Bottom-up, 250, 252, 256, 265, 286, 296

Boudouard reaction, 325, 883, 888, 1297
Broad spectrum analysis strategy, 233, 234
Bromates, 1037
Bromine, 15, 22, 36, 67, 84, 94, 150, 151, 336, 469, 601, 603, 633, 824, 1019, 1034–1035, 1159, 1514, 1515
  market distribution, 1035
Butadiene, 35, 36, 44, 275, 311, 312, 316, 318–322, 328, 335, 342, 377, 379, 437, 579, 582, 589, 595, 601, 619–621, 627–634, 636–638, 799, 929
n-Butane, 166, 179, 180, 182–184, 217, 316, 318, 319, 321, 346, 808
Butane, chemicals from, 31, 64, 68, 69, 215, 321, 335, 345, 346, 726, 807, 820, 927, 1051, 1154, 1453
  oxidation, 309
1,4-butanediol, 10, 309, 327–330, 332, 366–367, 377, 553, 1107, 1160
Butanes and butylenes, 1154
Butanol, wood pulping, 1097
1-butene, 319, 321, 361, 367, 592
Butene-2, 10
t-Butyl alcohol, 956
Butyl alcohols, 956, 1008, 1222
Butylenes, 734, 1154
Butylraldehyde, 228
Butyl rubber, 41, 322, 330, 595, 624, 632–633
  from isobutylene monomers, 632
Butyraldehyde
  plasticizer, 228
  polyvinyl alcohol, 222
  production, 228
  rhodium tri carbonyl triphenyl phosphate, 228
C
Calcined dolomites, 1300
Calcium chloride markets,
Calcium hypochlorite, 1035, 1036
Cancer treatment and detection, 844
Cane sugar refining, 1409, 1413, 1425
Capacitor, 62, 278, 614, 1493, 1506–1507, 1522, 1523, 1536
Capillary electrophoresis, 243, 244
$CaPO_4$ precipitation, 1234
Caprolactam, 9, 44, 99, 320, 327, 330, 345, 376, 437, 477, 877, 879, 901, 932, 1218
Carbodiimides, 560, 561, 1206
  in urethane adhesives, 561
Carbofuran, 238, 643
Carbohydrates
  dehydration, 1289–1291
  hydrolysis, 1288
  oxidation, 1289
  pyrolysis, 1288–1289
Carbonates, 107–1126, 1156, 1159, 1301, 1414, 1419, 1424, 1431, 1440–1443, 1460, 1464–1466, 1483, 1500, 1501, 1509, 1510, 1513, 1517, 1518, 1520
Carbon black, 38, 242, 265, 600, 623–624, 629, 630, 886, 889, 1051, 1053, 1123, 1126, 1128
Carbon dioxide
  liquid, 611, 1049, 1050
  production and uses, 1050
  by solvent absorption process, 1049
  supercritical, 581, 1049
Carbon disulfide, 65, 326, 357, 425, 477, 1086
Carbon fibers, 11, 463–464, 592
Carbon monoxide, shift, 892–893
Carbon sequestration, 1254
Carboxylic acids, as curatives in epoxy adhesives, 557
Cardiovascular agents, congestive heart failure
  clopidogrel, 398
  enalapril, 398
  lisinopril, 398
  sumatriptan, 398
Cardiovascular agents, hypertension agents
  amlodipine, 396
  irbesartan,
  losartan, 397
  metoprolol, 397
Casale high efficiency trays (HET), 913
Catacarb process, 893
Catalysis
  general acid, in epoxy adhesions, 555
  isocyanate-hydroxyl reactions, 559, 562
Catalyst carriers
  Al_2O_3, 203
  C, 203
  CeO_2, 203
  dispersion, 203
  $MgCl_2$, 229
  particulate, 207
  pore size, 203
  properties, acidity, 203
  SiO_2, 203
  surface area, 203, 204
  TiO_2, 203
  ZrO_2, 203
Catalyst characterization, 213–214
Catalyst deactivation
  attrition, 224
  carrier sintering, 212
  catalyst component sintering, 212
  coking, 213
  selective poisoning, 212–213
Catalyst dispersion, 203
Catalyst materials
  molecular sieves, 203
  SiO_2-Al_2O_3, 203, 215
Catalyst metals
  Ag, 202
  Co, 202, 213
  Cr, 202, 228
  Cu, 202, 213, 228
  Fe, 202
  Mn, 202
  Mo, 202
  Ni, 202, 213, 228
  Pd, 202
  Pt, 202
  Rh, 202
  Ru, 202
  Ti, 580, 930
  V, 202
Catalyst passivation, 731
Catalyst preparation, 207–208, 229
Catalyst regeneration
  coke, 216
  cracking process, 215–216
  hydro-demetallization, 214–215
  hydro-desulfurization, 214–215
  reforming, 216–217
Catalysts, historical
  metal supported, 736
  sulfur compounds, 213
Catalyst types

    heterogeneous, 208
    homogeneous, 228
    Rh tricarbonyl triphenyl phosphate, 228
Catalytic active sites, 208, 212
Catalytic converter
    carbon monoxide, 207, 217–218
    nitric oxide, 207
    oxidation, 217–218
    three way catalyst, 218–219
    unburned hydrocarbons, 207, 217–218
Catalytic materials, Faujasite zeolite, 215
Catalytic reactions
    aromatization, 311
    cracking, 202, 215
    dealkylation, 272
    dehydrogenation, 317, 377–379
    electrochemical, 1516, 1517
    hydrodemetallization, 214–215
    hydrodesulfurization, 214–215
    hydroformulation, 228
    hydrogenation, 210
    isomerization, 207, 217
    kinetics, 209–211
        rate limiting steps, 211
    methanation, 226
    naphtha reforming, 202, 216–217
    oxidation, 217–218
    polymerization, 202
    preferential oxidation, 228
    rate models, Langmuir-Hinshelwood, 211
    rate models, Mars-van Krevelen, 211
    reduction, 217
    steam reforming, 218
    table, 202
    three way catalysis, 218–219
    water gas shift, 202, 207
Catalytic reactors
    batch slurry, 209
    fixed bed, 209
    fluid bed, 209
Catalytic steam reforming, 207, 884, 1299
Cathode, 84, 286, 1025, 1028, 1037, 1122, 1301, 1499–1503, 1505, 1507–1510, 1513–1522, 1524–1530, 1533, 1535
Caustic soda, 100, 427, 1017, 1018, 1020, 1021, 1023, 1024, 1026–1033, 1035–1037, 1067, 1072, 1434, 1435, 1437, 1438, 1440–1442, 1523, 1524, 1526
Cell adaptation, sequential adaptation
    serum-free adaptation, 1237–1238
    starve and save adaptation, 1238
Cell banking, 1229, 1235
Cell lines, mammalian, for industrial production, 92, 1229–1246
Cell liquor, 1028, 1037
Cellobiases. *See* %-Glucosidases
Cellobiosyl fluorides, 1279
Cell stability, 1235
Cellulase(s), 1095, 1102, 1112, 1113, 1150, 1152, 1162, 1166, 1183–1185, 1187, 1192, 1194, 1200–1201, 1252, 1259–1261, 1273–1281
    activity measurement, 1276
Cellulose
    assays, 1274
    deconstruction, 1288
    dehydration, 1102
    mathematical modeling, 1278–1279
    production, 1200
Cellulose acetate
    secondary acetate, 432–433
    spinning, 433–436
    triacetate, 433–434
Cellulose hydrolysis
    acid, 232, 1100–1102, 1198, 1267, 1288, 1290, 1417, 1423, 1424
    enzymatic, 1099, 1102, 1170, 1198, 1201, 1213, 1252, 1259, 1273–1275, 1279, 1280
    hemicelluloses, 1281
Cell voltage, 1503–1505, 1507, 1508, 1510–1513, 1515, 1522, 1526–1530, 1534, 1535
Cell wall genomics, 1271–1272
Central nervous system agents
    Alzheimer's disease
        donepezil, 406
        memantine, 406
    antidepressants
        fluoxetine, 405
        paroxetine, 405
        serotonin uptake inhibitors, 405
        sertraline, 405
        venlafaxine, 405
    anxiolytics
        alprazolam, 405
        buspirone, 405
    bipolar disorders
        aripiprazole, 405
        gabapentin, 406
        olanzapine, 405
        risperidone, 405
Cephalosporins, 1176–1178, 1214–1217
Chain termination in free radical polymerization, 564
Chain transfer in free radical polymerization, 564
Charcoal, 242, 463, 848, 1000, 1103–1105, 1157, 1206, 1294, 1295, 1476, 1477, 1522
Charged particle transmutation products, 844–845
Chemical additives, 23, 25, 427, 768
Chemical analysis, strategies, 231–232
Chemical catalysis, 1168, 1188, 1225, 1251, 1284–1292
Chemical-defined (CD), 1236
Chemical explosives
    ammonium nitrate and ANFO, 1485–1486
    bulk emulsions, 1489–1490
    chemistry, 1474–1476
    classification, 1478
    explosions, 1474–142
    historical, 1476–1478
    initiation systems, 1491–1494
    manufacture, 1480–1483
        ammonium nitrate and ANFO, 1485–1486
        bulk emulsions, 1486–1488
        DDNP (2-Diazo-4,6-dinitrophenol), 1482
        dynamite, 1484
        HNS (2,2,4,4,6,6-Hexanitrostilbene), 1481
        NG (Nitroglycerin), 1483
        packaged, 1484
        PETN (Pentaerythritol tetranitrate), 1483
        RDX and HMX (cyclic nitramines), 1481
        TATB (1,3,5-Triamino-2,4,6-trinitrobenzene), 1482
        TNT (2,4,6-Trinitrotoluene), 1480–1481
    markets, 1473–1474
    packaged, 1483–1484
    properties (table), 1479
    sensitivity, 1478
    structural characteristics, 1478–1480
Chemical fibers, 3–5
Chemical industry

Chemical industry (*cont.*)
  Arab countries, 7
  biotechnology, 12–14
  characteristics, commercial development, 1, 7, 12, 13
  characteristics, investment trends, 1, 16, 22, 28
  characteristics, obsolescence/research, 5, 10–12
  characteristics, technological orientation, 10–12
  China, 19, 30, 34
  coatings industry, 16
  definition, 13
  economic context, 1–4, 10, 13, 27
  electronics, 26–27
  electronics/computers, 8, 15, 16, 25–27, 36, 38, 40
  energy use, 8, 15
  entrepreneurs, role, 42
  fine chemicals, 39–40
  foreign trade, U.S.A, 42
  France, nationalization, 4–6
  future trends, 34, 41, 45
  globalization, 31, 32, 45
  history, 1–45
  Italy, Spain, 6
  mergers and acquisitions, 4
  NAICS code 325,
  NAICS code 3251,
  NAICS code 3255,
  oil, relationship with, 40–41
  overcapacity in 1970s, 1–2
  public opinion, regulatory activities, 29–30
  R&D, 39
  R&D spending, 11, 12
  restructuring in 1970s, 1980s, 2–4
  specialty chemicals, 15–16, 35–39
  terrorism, 830
  USA, 3
Chemical properties and environmental fate, 174
Chemical reactor design, green, 162, 164
Chemical warfare agents (CWAs), 261, 281
Chemisorption, on catalysts, 202–205, 213
Chinese hamster ovary, CHO cell lines, CHO-S, CHO-K1, CHODG44, CHO-DUXB11, 1229
Chiral chemistry, 31
Chirality, 656
Chiral resolution, 1217, 1218, 1220, 1221
Chloramine, 922, 923
Chlorinated ethanes, 315, 328–330, 360–361
Chlorinated ethylenes, 360
Chlorine
  domestic US demand
    electrolysis cells, 1499
    Eltech diaphragm cell, 1029
    energy efficiency of, 1032
    flow sheet, 1028
    fused salt electrolysis of metal chlorides, 1027
    global production, 1027
    membrane cell, 1028–1033
    mercury cell, 1028, 1029, 1031, 1033
    minimum energy requirement, 1033
    producers in U.S, 1027, 1031
    pulp bleaching, 1027
Chlorine dioxide
  commercial generator, 1036
  pulp bleaching, 185, 1037
Chloroacetanilides,
Chloro-alkali technology
  diaphragm cell, 1028, 1029, 1031, 1033
  membrane cell, 1028–1033
  mercury cell, 1028, 1029, 1031–1033
Chlorobenzenes, 319, 358, 512, 1120, 1524
Chlorocarbons, 260, 279
1-Chloro-2,3-epoxypropane, in epoxy resin manufacture, 551
Chlorohydrin process, 1455
Chloromethanes, 326, 343, 344, 357–359
Chloroprene rubber, Neoprene, 620, 634
Chlorosulfonated polyethylene, 566, 623, 636
Chromatography
  detectors, 244–245
  gas and liquid, 233, 243, 244, 282, 394, 1397, 1421
  protein purification, 1243, 1245
Chromium oxide, 229, 317, 374, 892
Citric acid, 12, 24, 25, 959, 977, 1131, 1150, 1152, 1154, 1156–1158, 1365, 1375, 1377, 1384, 1426
Clathrates, 817, 821, 827
Claus process, 723, 755, 819, 893, 1001, 1003, 1004
Clostridium thermosaccharolyticum, 1260
Coal
  classification by rank and type, 752
  coke, 761, 769
  composition and analysis, 754–756
  consumption, 749, 750, 761, 766
  heat of combustion (table), 755
  liquefaction, 757–758, 771, 779, 783, 789–791, 793, 794, 797–799
  mining, 756–758
  origin and classification, 751–753
  petrochemicals from
  preparation, 758–759
  production, 749, 750, 761
  reserves, 749, 794
  synthetic fuels from, 749, 756, 761
  top producing countries (table), 751
Coal, chemicals from
  coal gasification-acetic anhydride (flow sheet), 801
  Eastman Chemical Company, 801
  Sasol process, 801
Coal combustion
  combustion equipment, 762–766
  pollution controls, 766–767
  processes, 762
Coal gasification
  chemistry, 771–773
  gasifiers, 773–776
  power generation, 776–777
  power generation, integrated gasification combined cycle (IGCC), 777
  process descriptions, 777–781
Coal liquefaction
  direct liquefaction, 786–788
  indirect liquefaction, 794
  indirect liquefaction, Fischer-Tropsch, 794–797
  pyrolysis based processes, effects of process parameters, 784
  pyrolysis based processes, nature and use of products, 785
Coal utilization, environmental concerns, 761–762
Coatings
  application equipment regulations, SCAQMD Clean-Air Act, 1119
  application equipment requirements, 1119
  automotive, 1122–1123
  background and historical development, 1117
  cross-linking technologies, 1123
  electron beam and UV curing, 1122
  environmental emissions, 1118

  - hazardous waste regulations, 1120
  - lacquers, 1129
  - markets, 1117, 1118
  - paint, 1117–1129
  - for plastics, 1123–1124
  - powder coatings, 1122
  - solvent-borne, major constituents, 1121–1123
  - technical trends, 1120–1122
  - volatile organic compounds (VOC) regulations, by state and region (US), 1118
- Coatings industry, 1117, 1128
- Cocoa butter, replacers, 1390
- Coconut oil, 223, 1334, 1336, 1351–1352, 1374, 1382, 1387, 1393, 1397, 1432, 1434
- Coke
  - combustion systems, 216, 769, 888
  - gasification, 881, 888–889
- Combustion, 47, 53, 60, 65–73, 75, 76, 79, 203, 206, 213, 214, 216, 217, 220, 221, 293, 317, 324, 327, 334, 335, 338, 341, 342, 345, 379, 459, 587, 588, 603, 712, 728, 751, 755, 754–756, 761–769, 772–775, 777, 779, 781, 783, 785, 788, 807, 813, 819, 821, 876, 881, 882, 886–888, 890, 905, 923, 925, 926, 931, 1000, 1001, 1003, 1007, 1009, 1011–1014, 1020, 1026, 1034, 1042–1044, 1049, 1053, 1097–1099, 1250, 1251, 1253, 1254, 1267, 1292–1294, 1296–1304, 1355, 1368, 1393, 1394, 1474–1476
  - hazards in chemical plants, 220
- Commercial uses, 273, 279, 285–286, 289, 291, 294–295, 762, 1273
- Component titration (CT), 1237
- Composite materials, emergence, 2
- Congeners, 171, 233, 238, 239, 247
- Conjugated linoleic acids, 1339
- Consolidated biomass processing, 1260
- Continuous monitoring, 209, 235–236
- Control charts, 136, 138–141
- Copolymerization
  - random, 582
  - styrene and butadiene, 582
- Copolymers, block and graft, 582–583
- Corn
  - stover composition database, 1270
  - stover feedstock rapid analysis method, 1269
  - sweeteners, enzymes in, 1423
  - wet-milling, 1198, 1199, 1285
- Corrosion
  - hydrogen embrittlement, 885
  - inhibitor, 550, 892, 894, 909, 919, 1463, 1469
- Cottonseed oil, 1336, 1348, 1349, 1363, 1366, 1374, 1379, 1381, 1385, 1387, 1388, 1391, 1393, 1397
- Cottonseed oil, delintering
  - gossypol, 1379
  - miscella refining, 1363
- Coupling agents, silanes, 11, 549
- Cryogenic gases, boiling points and concentration in air
  - processes for, 1039
- Cryogenic traps, 241, 243
- Cumene, 313, 314, 319, 320, 328, 332–335, 338, 340, 1465
- Curing structural adhesives
  - crosslinking, 560
  - definition, 548–549
  - evaporation, 549
  - hot melt, 549
- Cyanamide, 556, 912, 916, 932
- Cyanocobalamin, 1174–1176
- Cyanuric acid, 915
- Cyanuric chloride, 533, 540, 929
- Cyclohexane, 99, 100, 216, 229, 319, 320, 328, 335, 345, 372, 374, 376, 436–438, 553
- Cyclopentadiene, 632

D

- Degradation, microbiological, 232, 239
- Dehydrogenase, 651, 1155, 1160, 1162, 1164, 1172, 1174, 1184, 1218, 1220, 1260
- Depolymerization, 1166, 1252, 1281–1283, 1294–1296
- Detection limit, 247, 902
- Detergents, additives
  - anticaking agents, 1466
  - antiredeposition polymers, 1465
  - builders, 1464
  - enzymes, 1465
  - optical brighteners, 1465
- Detergents, production
  - agglomeration processing, 1468
  - liquid detergents, 1469
  - spray drying, 1466
  - typical formulation, 1466
- Diaminodiphenyl sulfone, curing epoxy adhesives, 556
- Diammonium phosphate, 974–976, 1015, 1096
- 2-Diazo-4,6-dinitrophenol (DDNP), 1479, 1482
- Dichlorobenzene, 358
- 2,4-Dichlorotriazines (2, 4-D), 238, 650
- Diethanolamine, 893, 920, 1050
- Diglyceride salad and cooking oils, 1370, 1387, 1393
- Dihydrofolate reductase (DHFR), 1229, 1233
- Diisocyanate dimers, in urethane adhesives, 560
- Diisocyanates, 35, 43, 319, 451, 560–562, 596, 608–610, 635, 636, 910, 1088
- Dimensionally stable anodes, 467, 1096
- Dimethyl sulfide, 1074, 1086
- Dimethyl sulfoxide (DMSO), 15, 1059, 1086, 1213, 1235, 1281, 1291
- Dioxin, 9, 30
- Dipole interaction, adhesion, 549
- Diuron, 238, 648
- DNA, 13, 92, 286, 394, 407, 408, 410, 502, 528, 532, 652, 1131, 1133, 1135, 1136, 1142, 1145, 1153, 1178, 1179, 1183, 1185, 1187, 1190, 1193, 1196, 1197, 1205, 1223, 1234, 1235, 1243, 1245, 1259, 1264, 1324, 1325
- Docosahexaenoic acid, 1176, 1335
- Dust explosions, 75, 76, 1128
- Dyeing, 327, 426, 436, 445, 447, 449, 450, 455, 468, 469, 473–475, 477, 478, 480, 481, 483, 486, 487, 494, 496–499, 520, 1017
  - pigment, 499
- Dyeing textile fibers
  - dye-bath preparation, 496–497
  - fiber preparation, 496
  - finishing, 497
  - methods/batch, 497–498
- Dye manufacture
  - anthraquinone
    - acid, 529–530
    - basic, 530
    - disperse, 526–529
    - reactive, 530–533
  - azo
    - disazo, 519–521
    - monoazo, 518–519
    - polyazo, 521–523
  - fluorescent brighteners, 540–543

Dye manufacture (*cont.*)
  nitro, 515
  phthalocyanine, 540
  sulfur, 538–540
  triphenylmethane, 523–524
  vat
    anthraquinoid, 533–537
    indigoid, 537–538
  xanthenes, 524–525
Dyes, classification
  acid, 479–481
  azoic, 481–482
  basic, 482–483
  direct, 483–487
  disperse, 487–491
  reactive, 491
  sulfur, 491–494
  vat, 494–496
Dyes, liquid crystal, 499
Dyes, non-textile uses
  biomedical, 502–503
  hair, 503–504
  ink-jet, 499–500
  laser, 502
  organic photoconductors and infrared absorbent toners, 500–502
  photographic, 504–505
  production and sales, 543–545
  synthetic, history of, 474
Dyestuff industry, 474–476
Dynamic vulcanization, 595, 632, 638
Dynamite, 9, 912, 1477–1480, 1483, 1484, 1489, 1491, 1492

E
Effectiveness factor (catalysts), 204, 1206
Elastomeric fibers, 451–452
Elastomers
  high temperature, 595, 619, 638, 639
  olefin, 632
  silicone, 634–635
  thermoplastic, 582, 583, 586, 594–596, 613, 621, 622, 632, 636, 639
Electrochemistry
  electrolyte, 1500–1501
  Faraday's law, 1504
  Nernst equation, 1502–1503
  oxidation, 1499
  reduction, 1499
  reference electrode, 1501–1502
  separator, 1501
  standard potential, 1502
Electrolysis
  ammonia electrolysis, 1533–1534
  coal electrolysis, 1535
  urea electrolysis, 520–521
  water electrolysis, 87, 1498, 1504, 1510, 1522, 1529, 1532–1535
Electrolyzer, 277, 1030, 1031, 1046, 1504, 1524, 1526, 1527, 1529, 1530, 1533–1535
Elemental analyses, 231, 239, 245–246, 1098, 1296
Emergency planning, 90, 92, 97, 98, 108, 123–125
Emergency preparedness programs
  checklist, 109, 110, 114, 117, 118, 120–123
  community involvement, 121–122
  elements, 107, 111, 120, 123
  emergency response plan, 107, 109, 113, 114, 117
  hazard identification/mitigation, 110, 112–113
  laws and regulations, 124–125
  managing, 107–128
  need for, 108
  personnel training, 113–114, 120
  preventing and predicting, 107–110
  release detection, 112–113
Endocellulase measurement, 1279
Endoenzymes, 1280
Endoglucanases, 1102, 1261, 1273, 1276–1279
Endoxylanases, 1275, 1280–1282
Energy
  crops, 1169–1170, 1254, 1292, 1304
  density, 1105, 1168, 1292, 1296, 1505, 1507–1515, 1522, 1523
  storage, 291, 1170, 1200, 1342, 1497–1536
  use, 159, 164, 182, 192, 193, 291, 838, 913, 1251, 1254, 1257
Environmental effects, 70, 294, 296, 1046
Environmental Fate and Risk Assessment Tool (EFRAT), 181–183
Environmental fate of chemicals
  heuristics for predicting, 176
  transformation in the environment
    abiotic processes, 172–174
    biotic processes, 170–172
  translocation, 166–170
Environmental impact assessment
  combined tiers assessment, 182–186
  EFRAT, 181–183
  HYSYS process simulator, 181–182
  maleic anhydride example, 179, 182, 184
  methyl methacrylate synthesis, 180, 181
  Tier I-Tier III, 176–182
Environmental risk assessment, 156, 176
Environmental risk minimization, 156, 182
Enzymatic degradation, 1273, 1275, 1330
Enzymatic hydrolysis, 1099, 1102, 1170, 1198, 1201, 1213, 1252, 1259, 1273–1275, 1279, 1280
Enzyme databases, 1187, 1274
Enzyme-linked immunosorbant assay, 1234
Enzymes
  accessory, 1280
  binding, 1186, 1188, 1189, 1207, 1214
  debranching, 1197, 1199, 1274, 1275, 1280, 1281, 1283, 1284
  depolymerizing, 1281, 1282, 1284
  immobilized
    glucose isomerase, 1200, 1206, 1425
    kinetics of, 1188–1189, 1206
    penicillin acylase, 1207
    production, 1206, 1207, 1217, 1223
  industrial, 1143, 1145, 1183–1225, 1250
  polysaccharide hydrolysis, 1283, 1288
  recovery process, 1136, 1142, 1143, 1145, 1210, 1220
Epichlorohydrin, 313, 314, 361, 362, 551–554, 558, 599, 919
  in epoxy resin manufacture, 551–554, 919
Epoxide, adhesive, 551, 557
Epoxy resins
  in acrylic adhesives, 548, 566
  commercial, 551–554
  cycloaliphatic, 553
  reactive diluents, 552–554, 558
    monofunctional, 552, 554, 558
    polyfunctional, 552, 553, 557
  specialty, 554
  transesterification, 557, 564
Escherichia coli, 13, 273, 276, 278, 1103, 1133, 1134, 1160, 1164, 1167–1171, 1173, 1174, 1190, 1196, 1202, 1216, 1218, 1220, 1230–1232, 1243, 1259, 1260

Esterases, 1189, 1201–1202, 1205, 1220, 1280, 1281, 1283, 1284
Ethane, products from, 7, 718, 807, 812, 814, 815, 820, 823, 1051
Ethanol
  cellulose hydrolysis, 1259
  feedstock
    carbohydrate content, 1268, 1271, 1272
    composition, 1269–1272, 1294, 1297, 1298
    effect on ethanol price, 1271, 1272
    proximate and ultimate analyses, 1292
  fermentation, 1049, 1050, 1131, 1165, 1166, 1173, 1258–1259
  metabolic yield, 1260
  process yield, 1260
  recovery, 1152, 1166
  volumetric productivity, 1210, 1215
  wood pulping, 1069
Ethanolamines, 43, 313, 315, 558, 819, 919–921, 1346
Ethyl benzene, 211, 275, 627
Ethyl chloride, 342, 343, 1165
Ethylenediamine, 17, 360, 452, 919, 921, 922, 1478, 1479
Ethylene-propylene elastomers, 23
Ethylenes
  dichloride, 275, 312, 313, 315, 328, 358–360, 368, 1024, 1034
  glycols, 7, 313, 315, 328–330, 368, 377, 441, 442, 448, 477, 589, 800, 817, 1093, 1154, 1286, 1384, 1462, 1465, 1484, 1531
  oligomers, 324, 334, 353, 369
  oxide, 2, 7, 18, 27, 67, 71, 95, 201, 202, 309, 312, 313, 315, 328–330, 335, 340–342, 361, 368, 369, 441, 442, 446, 532, 920, 1163, 1451, 1455, 1461, 1463, 1500, 1501
Ethylene/vinyl acetate copolymers, 586
Exoenzymes, 1280
Exoglucanases, 1102, 1273, 1276, 1279, 1280
  measurement, 1274, 1279, 1280
Experimentation, contributors to process variations, 142–144
  industrial, 145, 146
Explosions, in chemical plants, 9, 47, 101
  physical, 79, 80
Explosives. *See* Chemical explosives
Extractives, 1058–1061, 1063, 1073, 1086, 1105, 1108, 1111, 1265, 1267, 1268
Extremophiles, 1133, 1223

F
Fast pyrolysis, 1105, 1295–1296
Fats and oils
  analysis methods, 1332, 1338, 1362, 1363
  chemistry, nomenclature, definitions, 1333–1338
  chemurgy era achievements, 1332, 1391–1392
  crystallization principles, crystal types and energy levels, 1340–137
    triglycerides and polymorphism, 1340–1341
  extraction equipment, 1360
  fat solids determination
    plastic range, 1373
    temperature-solids profiles, 1373
Fats and oils, food products
  antioxidants, 1383–1384
  diglyceride salad/frying, 1370, 1374, 1375, 1381, 1385–1386
  emulsifiers/surfactants, HLB, 1385, 1452
  frying oils, 1385–1386
  industrial uses
    fatty acid methyl esters, biodiesel, 1392–1396
    historical, 1390
    industrial crops fatty acids, 1392, 1394
    oleochemical derivatives, 1394
    soy oil and lecithin, uses, 1392, 1393
    total fats and oil uses in U.S, 1393
    vegetable oil ester fuels, 1393
  nutrition labeling, 1381–1383
    trans fats labeling, 1381–1383
  oilseed extraction, whole seeds composition
    basic processes, 1349
    flow sheet, 1356, 1357
    seed preparation, 1356–1360
    solvent extraction, 1360
  processing profiled solids products
    base stocks preparation, 1388
    cocoa butter, 1390–1391
    margarines and spreads, 1387–1390
    shortenings, 1390
  processing profiles of solids modified products
    interesterified cocoa butter, 1390–1391
    margarines, spreads, 1387–1390
    rearranged lards, 1339, 1353
    shortenings, 1390
  refining and processing
    alkali-neutralization, 1365–1367
    bleaching, 1368
    deodorization and physical refining, 1376–1378
    flow sheets, 1363, 1364, 1367, 1375, 1389
    hydrogenation, 1371–1373
    interesterification, chemical, 1373–1376
    interesterification, enzymatic, 1373–1376
    phosphatides degumming, 1363–1365
    silica gel adsorption, 1367
    thermal fractionation, 1370–1371
    winterization, 1370, 1393
Fats and oils processing industry
  evolution, 1329, 1330
  fats/oils consumption patterns, 1027
  regulatory agencies, 1040, 1080, 1136
  world production, trade, 1452–1453
Fat-soluble vitamins, 1331, 1342, 1344, 1355
Fatty acids
  biohydrogenation, 1339
  composition, edible fats/oils, 1335, 1336
  composition, industrial crops, 1337
  conjugated linoleic acids, 1339
    from wood, 1339
  heat effects on saturation, 1335, 1344
  interesterification, 1340, 1373–1376
  names, properties, sources, 1333–1335, 1346
  oxidation, 1338–1339
  polymorphism, 1340–1341
  trans fatty acids, 1340, 1375, 1382
Fatty alcohols
  methyl ester process, 1453–1455
  oxo process for, 1454
  surfactants from, 1462
  Ziegler process, 1455
Fault tree, 86, 87
Fermentation
  formulation process, 1142, 1148–1153
  medium sterilization, 1136–1137
  organic acids recovery
  process development, 1144, 1152
  products, from sugars, 1142–1144
  regulatory aspects, 1136, 1141, 1152–1153
  sulfite waste liquor, 1165
Fermenting organism, 1252
Fermentor
  instrumentation and control, 1141–1142
  oxygen transfer, 1140, 1141

Fermentor (*cont.*)
  scale-up, 1140–1141
  types, 1138–1140
Fertilizers
  availability to plants, 959
  bulk blends, 985–986
  compound, 980–981
  controlled release, 989–993
  food production and, 959–994
  grade, 959
  industry, historical, 963–964
  from mineral phosphates, processing, 970–976
  miscellaneous low-volume
    fluid, 986–989
    granular processing, 981–985
    mixed, 986–989
    nongranular, 981
  natural organics, 966, 970
  nitrogen, 965–969
  phosphate, 969–976
  physical quality of, 993
  potassium salts, 978–980
  production, 968
  raw materials for production of, 964–965
  from synthetic ammonia, processing and handling, 966–969
  world consumption of, 963, 964
  world production of, 974, 981
Fiber variants, 466–469
  chemical and physical, 466–469
Fine chemicals, 14–15, 30, 32, 35, 36, 38–41, 44, 175, 190, 1110, 1204, 1221, 1224, 1265
Fischer-Tropsch, 10, 27, 312, 326, 705, 777, 782, 794–797, 800, 815, 816, 821, 822, 1105, 1106, 1252, 1254, 1300, 1301
  for coal liquefaction, 783, 786
Fish oils, 1323, 1334, 1336, 1342, 1347, 1352–1355
Fission activation products, 842–843
Fission, control of nuclear, 840
  nuclear fuel, 840
Fission neutron transmutation products, 838, 843–844
Flash pyrolysis, 325, 1105, 1294
Flavors, fragrances, 19–21, 33, 37, 1086, 1110, 1204
Fluid bed granulation, 913, 1149
Fluorescence-activated cell sorting (FACS), 1234, 1235
Fluorine from phosphoric acid manufacture, 955
Fluorocarbon elastomers, 639
Fluorosilicones, 634–635
Fluor solvent process, 893
Food safety, 290, 665, 1332, 1333, 1352–1353, 1386
Formacell process, 1079, 1286
Formaldehyde, 9, 27, 84, 95, 201, 202, 275, 321, 326–330, 335, 342, 344, 346–347, 352, 363–367, 370, 377, 379–381, 436, 450, 552, 561, 571, 596–599, 628, 738, 821, 824, 915, 921, 929, 991, 1087, 1088, 1092–1078, 1096, 1107, 1111, 1121, 1481, 1483, 1531
Formalin, 95
Formamide, 925, 928
Freeport Sulfur Company, 1003
Free radical, definition, 563–567, 577, 626, 636, 786, 1078
Fuel cell
  alkaline fuel cell, 1301, 1517–1518
  direct methanol fuel cell, 1301, 1517, 1519
  molten carbonate fuel cell, 1301, 1517, 1520
  oxide fuel cell, 1301, 1500, 1517, 1520–1521
  phosphoric acid fuel cell, 1301, 1517, 1519–1520
  proton exchange membrane fuel cell, 1501, 1516–1519
Fuel processing, 1521–1522
Fugene, 1234
Functional SBR, 629–630
Furfural, 208, 437, 726, 740, 1102, 1107, 1259, 1268, 1285, 1286, 1289, 1290, 1412
Fusion
  nuclear fuel, 840
  Tokamak reactor, 841, 842

G
G418, 1233
Galactomannan, 1282, 1283
Galactose, 1059, 1099, 1100, 1166, 1201, 1258, 1259, 1275, 1282–1284, 1426
Gas chromatography, 233, 243, 282, 1279, 1397, 1421
Gasifier reactor design, 1297–1298
Gas-to-liquids technology, 815–816
Gastrointestinal and genitourinary agents
  antisecretory
    esomeprazole, 401
    lansoprazole, 401
    omeprazole, 401
    ranitidine, 401
  benign prostate hyperplasia
    alfuzosin, 402
    doxazosin, 402
    tamsulosin, 402
  erectile dysfunction
    sildenafil, 403
    tadalafil, 403
    vardenafil, 403
Gaussian probability density, 135
Gel filtration (GF), 1146, 1243–1245
Gene-shuffling, 1191–1192
Genetically engineered organisms, 1164, 1196, 1252
Genetic screening, 1271–1272
Getter gauze, 227, 906
Giammarco-Vetrocoke, 893, 894
Glass fibers, 11, 71, 461–463, 469, 470, 586, 587, 589, 590, 592, 598, 601, 616, 617, 741, 976, 1012, 1155
Glass transition temperature, 445, 565, 574–576, 582–584, 611, 619, 620, 622–623, 631, 633, 635, 638
Glucoamylase, 1165, 1166, 1180, 1184, 1185, 1188, 1192, 1194, 1197–1200, 1423, 1424
Glucomannan, 1059, 1060, 1201, 1275, 1282, 1283
Gluconic acid, 1131, 1152, 1154, 1156, 1158, 1159, 1161, 1162, 1531
Glucose
  from cellulose, 1258, 1289
  isomerase, 1180, 1184, 1197, 1200, 1206, 1212–1214, 1423–1425
  oxidation, 1289
  syrups, 1197, 1198, 1213, 1214, 1416, 1423–1425
%-Glucosidases, 1102, 1276, 1280
%-Glucosides, 1280
Glucuronidases, 1280, 1281
Glutamic acid, 13, 1170–1172, 1175
Glutamine synthetase (GS, CHO-GS, NS0-GS), 649, 1229, 1233
Glycerin, 12, 1338, 1391, 1432, 1434, 1436–1438, 1440, 1443, 1444, 1446, 1447, 1452–1454, 1483, 1484
Glycidyl ethers, in adhesives, 551–553, 555
Glycol ethers, 313–315, 328–330, 353, 369, 372, 818, 1165
Glycosyl hydrolase, 1264, 1274, 1282–1284
Glycosyltransferase, 1184, 1187, 1205, 1220, 1222, 1264
Glyphosate, 11, 238, 649, 665, 684, 1326, 1327
Gold nanoparticles, 253–254
Good manufacturing practice (GMP), 1239

Graphite fibers, 463–464
Green chemistry
  concepts, approaches, tools, 157, 160
  definition, 155–157
  twelve principles, 157–159
Green engineering
  concepts, approaches, tools, 157, 160
  definition, 155–157
  principles, 159–160
Greenhouse gas, 164, 165, 179, 182, 191, 192, 194, 217, 227, 275, 593, 666, 700, 742, 761, 777, 792, 830, 1097, 1253–1256, 1303, 1304

H

Haber, L.F., 226, 876, 877, 895, 912, 966, 1301
Haldor Topsøe, 28, 885, 890, 899, 900
Handling strength (polymers), 555
Hargreaves process, 1022
Hazard and operability studies (HAZOP), 64, 85–87, 93, 103, 110
Hazardous materials storage, 50, 53, 56, 57, 89, 100, 108, 112, 114, 122–124, 157
Hazardous, toxic, reactive chemicals, 176
Helium
  sources, 1049
  storage and transport, 1049
  uses, 1049
Hemi-cellulases, 1102, 1112, 1183, 1184, 1201, 1273, 1274, 1278, 1280, 1281, 1284
Hemicellulose
  debranching enzymes, 1274, 1275, 1281
  depolymerization enzymes, 126–1283
  hydrolysis, 1102, 1275, 1281
Heptenes, 713
Herbicides, 11, 30, 31, 173, 238, 642, 643, 646, 647, 650, 656, 675, 678, 680–684, 875, 919, 920, 929, 932, 987, 992, 1037, 1108, 1111, 1120, 1214, 1221, 1264, 1326–1328, 1330, 1405, 1412, 1427, 1428
Hevea (rubber) tree, 621, 1168
Hexamine, 921–922
Hexanitrostilbene (HNS), 1478, 1479, 1481–1482
High fructose corn syrup (HFCS), 1195, 1197, 1200, 1204, 1206, 1207, 1212–1214, 1221, 1422, 1423, 1425, 1426
High molecular weight polyethylene
  gel spinning, 456
  production, 456–457
  routes to high performance, 456
Holocellulose, 1058, 1100
Host cell lines, 1229–1233
Host cell protein (HCP), 1243, 1245
hph phosphotransferase gene, 1233
Hydantoinase, 1217
Hydrazine
  processes
    Bayer Ketazine, 923–924
    Hoffman (urea), 923
    Peroxide, 924
    Raschig, 922–923
    Raschig/Olin, 923
  properties, 922, 923
Hydrocarbons, chlorinated, 30, 272, 342, 435
Hydrochloric acid, 95, 343, 344, 401, 743, 881, 930, 956, 980, 999, 1011, 1017, 1032–1034, 1036, 1048, 1100, 1198, 1397, 1398, 1424
Hydrogenation, catalytic vegetable oils, 223–224
Hydrogen bonding, 237, 272, 439, 549, 555, 574, 576, 588, 600, 609, 623, 1044, 1050, 1066, 1081, 1186, 1250
Hydrogen, by-product, 879, 882, 889
Hydrogen cyanide
  Andrussow process, 925–926
  by-product, 925, 926, 928, 929
  processes, 926–929
  properties, 925
  safety, 925
Hydrogen, forms, ortho
  in fuel cells, 1046, 1048
  high temperature processes, 1044
  from off-gas purification, 1044, 1045
  processes, 1044–1048
    coal-based, 1046, 1047
    electrochemical, 1046
    partial, 1043
    partial oxidation, 1043
    steam reforming
  properties, 1044, 1048
  shipping, 1048
  from steam-methane reformation, 1044
  uses, 1045, 1048
Hydrogen from chlorine cells, 1025
Hydrogen sulfide, 84, 85, 95, 112, 113, 279, 702, 723, 767, 770, 772, 773, 801, 807–809, 819, 821, 846, 893, 894, 997, 998, 1000, 1001, 1003, 1011, 1074, 1071, 1521
Hydrogen Texaco (HyTex), 889
5-Hydroxymethylfurfural, 1107, 1286
Hygromycin, 1233
Hysteresis in polymers, 619, 630
HYSYS simulation software, 181, 182

I

(S)-Ibuprofen, 1220
IEC. *See* Ion exchange chromatography (IEC)
IGCC. *See* Integrated gasification combined cycle (IGCC)
Imine, 374, 561, 563, 916, 918
Immobilized enzymes, 1200, 1206–1208, 1210, 1211, 1217, 1223, 1412, 1425. *See also* Enzymes
  production, 1206, 1207
Immunoassays, 245, 246, 288
Impurities, 9, 26, 63, 162, 214, 215, 224, 228, 256, 286, 323, 336, 340, 346, 375, 376, 415, 416, 424, 426, 438, 448, 550, 554, 564, 606, 624, 631, 657, 662, 664, 723, 726, 735, 758, 761, 768, 772, 780, 794, 848, 852, 853, 895, 926, 939, 949, 956, 975, 976, 982, 987, 991, 1018, 1021, 1029, 1032, 1040, 1045, 1050, 1053, 1079, 1108, 1137, 1144, 1147, 1157, 1176, 1207, 1212, 1243, 1245, 1298, 1299, 1301, 1341, 1354, 1363, 1368, 1397, 1399, 1406, 1408, 1411, 1413, 1414, 1419, 1420, 1425, 1440, 1443, 1444, 1448, 1482, 1513, 1527
Industrial electrosynthesis
  chlor-alkali process, 1522–1524
  electrolysis, 1531–1535
  electrorefining, 1528–1529
  electrowinning, 1528–1529
  Hall-Heroult process, 1527–1528
  organic electrosynthesis, 1529–1531
Industrial enzymes
  applications, 1183–1185, 1190, 1192–1193
  markets, 1183–1185, 1192
  production, 1183–1185, 1192, 1193
Industrial gases
  from air
    adsorption separation, 1039, 1040
    cryogenic separation, 1039
    membrane separation, 1039–1040

Industrial gases (*cont.*)
  argon, 1043
  business consolidation, 35–39, 41
  cryogenic gases, 1039–1045, 1048–1050
  hydrogen, 1043–1048
  nitrogen, 1042
  overview, 1039–1042
  oxygen, 1043
  properties (table), 1044, 1048
  from wood, 1057–1113
Inedible fats, 1396
Infectious disease agents
  antibacterials
    azithromycin, 406, 407
    ciprofloxacin, 407
    daptomycin, 407
    linezolid, 407
  antifungals
    fluconazole, 408
    itraconazole, 408
  antineoplastics
    antivirals, atazanavir, 409, 410
    bortezomib, 410
    cetuximab, 410
    gefitinib, 410
    imitanib, 410
    irinotecan, 410
    paclitaxel, 410
    tamoxifen, 410
  antivirals
    efavirenz, 409
    indinavir, 409
    ribavirin, 409
Inflammation and osteoporosis agents
  arthritis
    celecoxib, 404
    rofecoxib, 404
    valdecoxib, 404
  osteoporosis
    alendronate, 404
    raloxifene, 404
Initiator, free radical polymerization, 565, 577, 578
Integrated gasification combined cycle (IGCC), 767–769, 771, 777, 781, 782, 800–802, 894, 1300–1301
Inter-agency Working Group on Nanotechnology (IWGN), 251
Interface (adhesion), definition, 547–548
Internal ribosomal entry site (IRES), 1233
International Standards Organization, 153, 187–189, 237, 754, 857, 1329
Interphase, definition, 550
Ion exchange chromatography (IEC), 243, 244, 246, 1196, 1243, 1244
Ion-exchange membrane, 1028, 1526
Ionic liquids, 165, 1223, 1224, 1291
Ionomers, 586, 631, 636–638
Isobutanes/isobutylene, 19, 101, 134, 180, 181, 313, 318, 319, 321, 335, 341, 352, 367, 564, 578, 632, 633, 721, 722, 732, 734, 798, 808, 918
Isocyanate dimers, in urethane adhesives, 560, 562
Isocyanates, blocked
  commercial, 562
  reactions, 562
  in urethane adhesives, 562–563
Isophthalic acid, 319, 320, 336–338, 445, 446, 468, 589, 1362
Isoprene, 322, 579, 595, 619–621, 625, 626, 629, 632, 633, 637, 1095, 1154, 1166–1168, 1342–1344
Isopropanol, 192, 193, 313–315, 328, 338, 367–368, 377, 379, 381, 1168, 1176
Isotope enrichment processes, 845–847
Isotopes, uses of, 844
Itaconic acid, 1154, 1157, 1158
IUBMB enzyme classifications, 1274
IUPAC Filter Paper Assay, 1276–1278

K
KBR Reforming Exchanger System (KRES), 885, 899
Ketazine, 922–924
Keto-gulonic acid, 1154
Kinetics, catalytic reactions
  batch slurry, 209
  fixed bed, 209, 210
Kjeldahl method, 1267
Klebsiella oxytoca, 1259
Kraft process, 1072–1075, 1077, 1078, 1086, 1285, 1287–1288
  wood pulping, 1069

L
Lacquers, 16, 1117–1129
Lactic acid, 12, 466, 470, 594, 929, 1103, 1113, 1152, 1154, 1160–1161, 1172, 1221, 1260, 1406
Land use, 179, 959
L-aspartic acid, 1214, 1218, 1219
Latex, from Hevea tree, 624
Latex, viscosity, 625, 629
Levulinic acid, 1100, 1107, 1154, 1268, 1285, 1286, 1288–1291
Lewis acid, in adhesive cure, 549, 554, 557, 558
Lewis base, in adhesives cure, 554, 557, 558, 567
Life-cycle assessments
  goal and scope, 188–189
  inventory, 189–190
  in practice, 191–194
  principles, 187
Life-cycle inventory, components, 187, 188
Lifestyle illnesses, 288–289
Lignin
  analysis, 1267
  deconstruction, 1291–1292
  enzymatic degradation, 1273
  oxidation processes, 1291
  production, 1285
  structure, 1060, 1073, 1291
Lipases, 1112, 1184, 1185, 1187, 1192, 1201–1202, 1205, 1207, 1212, 1220, 1223, 1224, 1331, 1333, 1342, 1346, 1349, 1376, 1387, 1391, 1447
Lipids, 181, 224, 594, 647, 653, 1142, 1170, 1201, 1202, 1234, 1235, 1272, 1323, 1324, 1331, 1332, 1334, 1341, 1342
Lipofectamine, 1234
Liquefied gas storage, 40
Liquefied natural gas (LNG)
  flow diagram, 1051
  sources and uses, 1051
Liquid chromatography, 243, 244, 394, 1397
Liquid hydrogen storage, 1048
Lithium hypochlorite, 1036
Lithium metal catalyst, 626
L-lysine, 13, 1131, 1217, 1218
Lumber, 1057, 1059, 1086–1089, 1094, 1098
Lysine, 1154, 1170–1173, 1350

M

Magnesium hydroxide, 601, 1067
MAIC (Six Sigma disciplines), 153
Maleic anhydride, 10, 17, 25, 31, 179, 182, 184, 322, 323, 328, 335, 338, 345–346, 377, 455, 598, 1465, 1466
  production, environmental impact, 183, 184
Manganese, forms, 232, 336
Mannan, 1265, 1275, 1282
  depolymerization, 1282
Mannheim process, 980, 1022
Mannose, 1059, 1099, 1100, 1166, 1201, 1258, 1259, 1275, 1283
Mass spectrometry, 245, 246, 1185, 1187
Master cell bank (MCB), 1235
Matrix attached regions (MARs), 1233
MDI (4,4-methylenebis (phenyisocyanate), 931
Media, cell culture, 1235, 1239
Median (data), 133, 134, 287
Melamine, 7, 44, 330, 596, 597, 599, 915–917, 1087, 1088, 1121–1123
Membrane cell, 653, 675, 1029–1033, 1210, 1457, 1524, 1526–1527
  voltage distribution, 1526, 1527
Mercury cell, 1029, 1031–1033, 1524–1525
Mercury, forms, 767, 1025, 1028, 1031, 1033, 1035, 1480, 1524
Metabolic agents
  diabetes agents
    glimepiride, 400
    metformin, 400
    rosiglitazone, 400
  hyperlipidemia agents
    atorvastatin, 399
    pravastatin, 399
    rosuvastatin, 399
    simvastatin, 399
    vytorin, 399
  obesity agents
    orlistat, 401
    sibutramine, 401
Metabolic pathway engineering, 1167, 1259–1260
Metagenome, 1190
Metallocene elastomers, 638
Methane, coal bed, 809–811
Methane conversion
  ammoxidation, 824–825
  direct to chemical products, 822–823
  indirect via nonsyngas intermediates, 822
  indirect via syngas, 822
  oxidative coupling, 823–824
  partial oxidation, 824
  pyrolysis or cracking, 824
  reductive nitrilization, 825
Methane conversion processes, 821–825
Methane hydrate, description, 821
  natural gas source, 821
Methanol from wood, 1103, 1106
Methanol reformer hydrogen (MRH), 890
Methionine, 13, 651, 929, 930, 1170–1172, 1195, 1197, 1204, 1214, 1217, 1233, 1234
Methotrexate, 1233, 1234
Methylamine, processes
  production, 919
  properties, 918
Methyl chloroform, 328, 357, 1524
Methyl diethanolamine, 893, 894, 1050
Methylene dianiline, curing epoxy adhesives, 556
Methyltaurine, 919
Methyl tertiary butyl ether (MTBE), 3, 14, 191, 237, 320, 322, 323, 327, 821
Metolochlor, 1221
Microbe
  diversity, 1132–1133
  engineering, 1134–1135
  kinetics, 1137–1138
  screen and selection, 1133–1134
  strain development, 1135–1136
Microdenier fibers, 445, 465, 476
Miscellaneous agents, 412–414
Miscellaneous pharmaceutical agents, 412–414
Mixture experiments, 149–152
Modacrylics, 420, 421, 449–450
Molasses
  beet, 1156, 1175, 1420–1421, 1425, 1426
  black strap, 1156, 1165, 1410, 1425
  cane, 1156, 1411, 1421, 1426
  composition, 1425, 1426
Molecular sieves, 203, 242, 368, 734, 785, 817, 824, 875, 895, 909, 1039, 1152, 1455
Molecular technology, 250
Monoammonium phosphate (MPA), 976, 982, 989
Monochlorobenzene, 319, 320, 358
Monoethanolamine, 773, 892–894, 898, 919–921, 1050
Monoliths
  advantages, 207, 209
  ceramic, 207, 208, 218, 219, 221
  cordierite, 220
  metal, 207, 208, 218
  washcoat, 207, 218, 220
Morpholine, 651, 689, 919, 920, 1086
Multipurpose gasification (MPG) process, 888

N

Nafion, 1501, 1516, 1518, 1526, 1530
Nanofibers, 289, 290, 465–466, 1522
Nanomaterials, 203, 251–253, 265–266, 271–274, 279, 280, 284, 286–296, 1208
Nanoscale, 249, 251, 252, 257, 260, 261, 265–267, 273–276, 279–282, 285–287, 289, 291, 294, 572, 593, 602
Nanoscale regime, 251
NanoScale zero valent iron (NZVI), 276, 278–280
Nanotechnology
  applications, 249–262
  definition, 249–250
  dielectrics,
  gold nanoparticles, 253–254
  IWGN, 251
  manufacturing technology, 250
  metal oxide nanoparticles, 256–257
  metal-supported catalysts
  silver nanoparticles, 254–255
Naphthalene derivatives, 505
(S)-Naproxen, 1214, 1220
Natural gas
  biological origins, 825–826
  characteristics, 807–808
  composition of raw gas, 807, 808
  evolution of industry, 809
  future sourcing, 827
  H_2S and Cl_2 removal, 819
  hydrates, formation and prevention, 816, 817, 820, 821
  industry, 809–816, 820
  international commodity, 815
  liquefied (LNG), 815
  liquids, 812–813

Natural gas (*cont.*)
  lower crust and mantle domains, 826–827
  methane conversion, 821–825
  non-biological origins, 826
  petrochemical products from, 812, 813
  production, 812–813
  reserves, U.S., 813
  transmission and sale, 816–819
  well through processing, 810
Naval stores, 1086, 1104, 1105, 1108–1111
Nelson method, 885
Neo gene, 1233
Nermst equation, in oxygen sensor,
Nitration, 15, 39, 319, 320, 505–508, 910, 930, 1008, 1480, 1481, 1483
Nitric acid
  abatement, 906, 909
  concentrated, 903, 906, 908
  destruction by catalysis of NOx (DCN), 906, 908, 909
  high efficiency absorption (HEA), 909
  processes, 903–910
  properties, 903
  red fuming, 903, 909, 910
  stabilizers, 909
  uses, 910
Nitric phosphate, 969, 970, 974
Nitriding, 900
Nitrilase, 1205, 1220
Nitrile rubber, 44. 566, 595, 633–634
Nitrobenzene, 84, 210, 238, 319, 320, 507, 508, 726, 910, 929–931, 1120, 1531
Nitrogen
  characteristics, 875
  consumption, 877–879
  fertilizer, 4, 877–880, 900, 910, 965–970, 980, 990
  industrial, 879
  oxides, 95, 182, 242, 245, 278, 723, 723, 761, 766, 875–876, 879, 901, 903, 905, 909, 965, 1008, 1009, 1118, 1294
  properties, 876
Nitrogen trichlorid,
Nitroglycerine, 238
Nitrous oxide, 875, 876, 912, 1039, 1053
  from ammonium nitrate, 876, 912, 1053
N-methyl-2-pyrrolidone (NMP), 458, 893, 894
N,N, dimethylformamide, 255, 919
NO_x emissions, 221, 742, 766, 767, 876, 879, 886, 909, 1289, 1303, 1304
NS0 cell lines, 1229
Nuclear fission, 830, 832, 840, 841
  products, 830, 831, 840
Nuclear fuel reprocessing, 861
Nuclear industry
  alternative energy supply, 836–838
  definitions, 829
  isotopes, 830, 840, 845–847
  nuclides production, 842
  outlook, 834–836
  products, 843–844
  safety, 834–836
  status and outlook, 830–834
  waste products, 848
Nuclear material production, 829, 836
Nuclear material transportation, 829, 863
Nuclear medicine, 844, 872
Nuclear processes, 829, 838–839, 843, 845
Nuclear proliferation, 835
Nuclear radiation, exposure, 835
Nuclear reactors
  fast reactor (LMFBR), 864
  heavy water reactor (CANDU), 864, 865
  light water reactors, (PWR, BWR), 864, 865
  types, 864
Nuclear spent fuel reprocessing, 861
Nuclear tracers, 871–872
Nuclear waste
  airborne, 861
  liquid, 860–861
  solid, 861–862
  storage and disposal, 862
Nylon
  drawing, 439–440
  melt spinning, 438–4397
  new developments, 440
  Nylon 6, 437–438
  Nylon 66, 436–437

O
Octane number (gasoline), 216, 217, 729, 738
Octenes, 586
Oils, from wood, 1057–1113
Olefins, 3, 11, 42, 44, 101, 134, 166, 309–319, 321–324, 328–329, 331, 332, 334, 342, 344, 347, 348, 352, 356, 367–370, 377–379, 558, 579, 595, 639, 712, 714, 719, 722, 724, 730, 734, 738, 744, 795, 798, 800, 801, 823, 918, 1033, 1301, 1451, 1454–1456, 1461
Olive oil, 223, 1336, 1337, 1346, 1351, 1379, 1380, 1385, 1432
Olivine, 977, 1298, 1300
Organic foods, 1328–1330
Organoboranes, initiator in acrylic adhesives, 566–567
Organosolv process, 1079, 1286
Organosolv pulping, wood, 1077–1079
Oriented strand board (OSB), 1087–1089
Oxirane ring, in adhesives, 551
Oxo chemicals, 800
Oxonium ion, 555
Oxo process, 355, 1454
  for fatty alcohols, 1454
Oxygen sensor (automobiles), 22, 61, 189, 201, 222, 274, 291, 738, 763751, 1091, 1090, 1391
Paclitaxel, 392, 410, 1111
Paint industry, 16–17
Paint pigments
  aluminum, 1128–1129
  inorganic, 1126–1127
  organic, 1127–1128
  pearlescent, 1128
Palladium, 70, 279, 343, 352, 823, 843, 892, 905, 906
Palm oil, 192, 1332, 1334, 1336, 1344, 1346–1348, 1352, 1354, 1364, 1368, 1370, 1375, 1377, 1379, 1382, 1388, 1393, 1396, 1398, 1432, 1434
Paper
  consumption, 1066, 1069
  converting, 185, 1084–1085
  finishing, 1084–1085
  polymer modified, 550, 585
  properties, control of, 1082, 1084
  recycling, 1081
Paper making, pollution control, 30, 180, 183, 184, 202, 231, 766–767, 963, 1058, 1066, 1067, 1073, 1079, 1081, 1082, 1084, 1085, 1090, 1097, 1099, 1112, 1119, 1201, 1344, 1447, 1450, 1466
n-Paraffins, 316, 323, 328, 334, 377, 379, 721, 732, 733, 735, 752, 1455

Para-xylene, 318
Particleboard, 1087, 1088, 1090, 1098, 1122
Particulate matter in air, 231, 243
PCR reaction, 1205, 1223
Peanut oil, 1336, 1347, 1374, 1393
Pectin, 24, 1060, 1184, 1194, 1195, 1201, 1250, 1268, 1273, 1280, 1283, 1284, 1384
Pectinases, 1183, 1184, 1201, 1273, 1274, 1283–1284
Penicillins, 1176–1178, 1190, 1215–1217
Pentaerythritol tetranitrate (PETN), 1475–1477, 1479–1481, 1483, 1491, 1492
Peptization, 634
Permeabilized cells, 1210
Permeation, 281–282, 325, 326, 620, 1041, 1152, 1339, 1501
PET. *See* Polyethylene terephthalate (PET)
PETN. *See* Pentaerythritol tetranitrate (PETN)
Pharmaceutical agents
  antineoplastics, 410–411
  cardiovascular agents, 396–398
  central nervous system agents, 405–406
  gastrointestinal/genitourinary agents, 401–402
  infectious diseases, 406–411
  inflammation and osteoporosis, 404–405
Pharmaceutical industry characteristics
  chemical process R&D, 415–416
  clinical trials, 396
  drug delivery, 395
  libraries, discovery, 413
  libraries, optimization, 414
  libraries, targeted, 413–414
  medicinal chemistry, 392–393
  patents, 395
  pharmacodynamics, 394
  pharmacokinetics, 394–395
  research strategies, 393–394
  toxicity, 394–395
Pharmaceutical industry, development of, 391–417
Phenol hydrogenation, 1291
Phenols, curative in epoxy adhesives, 552, 554, 558
m-Phenylenediamine, curing epoxy adhesives, 556, 558
Pheromones, insect, 658
Phosphate minerals, 939, 943, 970–977
Phosphate, naturally occurring, 939, 940
Phosphate ores, 901, 940–942, 944, 970, 977
  mining and beneficiation, 941–945
Phosphate rock
  acidulated, 977
  chemical processing, 945
  direct application, 976
  treatment and end products, 940
  world production, 945, 974
Phosphates
  finished fertilizers, routes into, 970, 971
  industrial, 946–947
  thermal phosphate fertilizers, 977
Phosphatides, 1331, 1344–1346, 1355–1356, 1363–1368, 1379, 1385, 1387, 1443, 1444
Phosphoric acid
  annual production, 947
  dihydrate process, 945, 947–952
  environmental aspects of production, 957
  fertilizer grade, processing, 970
  hemihydrate process, 947, 948, 950–952, 955
  thermal process, 945–946
  unit operations, 952–954
  wet process
    animal feed supplements, 955–956, 977
    by-products, 955–957, 974
    typical analysis, 972
Phosphorus, elemental, production, 232, 430, 560, 602–603, 945, 946, 959
  forms, 232, 430, 560, 602–603, 945, 946, 959
Photochemicals, 25–26, 172, 179, 191–194, 237, 360–361, 1125–1126
Phthalic anhydride, 37–38, 319, 320, 335, 338, 339, 345, 373, 441, 514, 524, 526, 598
Pine oil, 19, 1086, 1108–1111
Plantation white sugar, 1403, 1412
Plant growth regulators, 642
Plasmids, 1193, 1233, 1234
Plasticizers, 38, 228, 320, 330, 354, 365, 373, 376, 461, 462, 550, 553, 575, 587, 601, 611, 613, 623, 624, 636, 639, 910, 919, 946, 1355, 1391, 1392
Plastic processes, 605
Plastic processing
  autoclave molding, 617
  blow molding, 587, 589, 600, 605, 611, 612
  blown film, 585, 586, 606
  bulk molding compound, 598, 616
  calendaring, 613
  centrifugal casting, 617
  cold press molding, 616
  compression molding, 597, 598, 605, 615, 616, 1091
  conformal coating, 614
  continuous laminating, 617
  encapsulation, 614
  extrusion
    coating, 606
    coextrusion, 606
    devolatilization, 605, 606, 785
    die forming, 606
    hose and tubing, 606
    land fracture, 604
    melt conveying, 605, 606
    melt fracture, 604
    melting, 587, 605, 606
    mixing, 605–606
    profile, 606
    solids conveying, 605
    twin-screw, 606
  filament-winding, 598, 602, 617
  flow properties, measurement, 603–604
  foam reservoir molding, 617
  hand-layup molding, 598, 616, 617
  impregnation, 615
  injection molding, 596, 600, 604, 605, 607–611, 615
  injection molding, molding cycle, 607–608, 615
  liquid casting processes, 614–615
  low-density foams, 610–611
  low-density foams, polyurethane, 610
  matched die molding, 616
  melt viscosity, 600, 604, 605
  perform molding, 609
  pipe extrusion, 587, 605, 606, 611, 617
  polyethylene foam, 611
  polystyrene foam, 611
  polyurethane flexible foam, 608
  polyurethane rigid foam, 610, 931
  polyvinyl chloride foam, 611
  potting, 614
  powder coating
    electrostatic fluid bed, 612, 613

Plastic processing (*cont.*)
    electrostatic spray, 612, 613
    fluid bed, 612–613
  prepreg molding, 616
  pressure-bag molding, 617
  pultrusion, 598, 602, 617
  reaction injection molding (RIM), 596, 597, 608–609
  reinforced plastics, 572, 615–616
  resin transfer molding, 616
  rheology, 603–604
  rotational molding, 612
  rubber plug molding, 617
  sheet extrusion, 377
  sheet molding compound, 598, 616
  sprayup molding, 598, 616
  structural foam, 609–610
  syntactic foam, 615
  thermoforming, 587, 600, 605, 612, 1091
  transfer molding, 615, 616
  vacuum-bag molding, 617
  vinyl plastisol, 612–614
  wire coating, 593, 606
Plastics
  abrasion resistant, 571, 585, 590, 593, 595, 602, 606
  antioxidants, 587, 600
  barrier, 593
  barrier properties, 602
  biostabilizers, 601
  building and construction, 593, 595
  chemical resistance, 572, 593–596, 598, 600, 614, 615
  coefficient of thermal expansion, 600, 602
  color, 572, 601, 612
  colorants, 601
  commodity thermoplastics, 585–590, 592
  continuous fiber reinforcement, 602
  cost, 572, 584, 590, 594, 601–603, 607, 612, 613, 615, 616
  coupling agents, 601
  creep-resistance, 590, 596, 600–602, 615
  critical properties, 601–602
  dielectric breakdown, 602
  elasticity, 604
  electrical conductivity, 602
  electrical insulation, 572, 593, 614
  engineering, 590–593, 596, 597, 615
  epoxy resins
    markets, 598–599, 615
    properties, 598–599
  filled and reinforced, 572
  fillers, 572, 587, 593, 601, 602, 609
  fire performance, 602–603
  flame-retardance, 572, 587
  flame-retardants, 601–603, 610
  flotation, 572, 596, 610, 611
  foams, 572, 609
  health and environment, 603
  heat-resistance, 588, 593, 596, 597, 599, 602, 615
  impact cushioning, 572
  impact strength, 571, 592, 593, 600, 615
  lubricants, 587, 601, 604, 605, 608
  lubricity, 571, 593
  maximum use temperature, 592, 600
  melamine-formaldehyde, 596, 599
    properties, 596, 599
  modulus, 571
  one-step conversion, 601, 608
  opacity, 572, 601
  organic peroxides, 601
  packaging, 572, 585–589, 593, 594, 596, 600, 603, 606, 610–612
  permeability, 572, 593, 594, 600
  phenol-formaldehyde, 571, 596, 597
  phenolic
    engineering properties, 589, 590, 592, 597, 615
    properties, 597
  plasticizers, 575, 587, 601, 611, 613
  polymer blends, 601
  processability, 571, 572, 585, 586, 588, 590, 593–595, 600, 601, 604
  product design, 571, 602
  recycling, 605
  reinforcing fibers, 572, 592–593, 601, 602, 609, 615–617
  softness, 572, 586, 610
  solvent resistance, 602
  stabilizers, 581, 587, 593, 600, 611, 613
  surface, 613
  thermal conductivity, 602
  thermal insulation, 571, 610, 611
  thermoset
    markets, 572, 594, 596, 601
    properties, 572–574, 585, 594, 597, 598, 600, 615
    thermoset polyester, 598
  transparency, 572, 586, 588, 589, 592–593, 600, 606, 611
  ultraviolet light stabilizers, 600
  unsaturated polyesters, 598, 601, 615
  urea-formaldehyde, 596–598
    properties, 596, 598
  water resistance, 572
  weathering, 572, 587, 589, 602
  weather-resistance, 593
  welding, 550
Platinum (Pt), 27, 28, 202, 207, 209, 211–213, 216–219, 221, 222, 224, 226, 227, 275, 276, 279, 285, 288, 317, 318, 376–379, 825, 904–906, 908, 926, 927, 1008, 1455, 1499, 1502, 1516–1520, 1525
Plutonium processing, 833, 835, 838, 841, 856, 859, 862
Plywood, 572, 597, 1057, 1087–1090, 1096
Pollution prevention
  heuristics, 164–166, 171, 174, 176, 177
  heuristics for chemical processes, 160–165
  separations devices, 162, 163, 165–166
Pollution Prevention Act of 1990, 155–156, 161
Polyamide polyamine (synthesis), 556, 559
Polybenzimidazole (PBI), 421, 459, 460
Polybutadiene, impact modifier, 631, 633
Polybutenes, 23, 322
Polychlorinated biphenyl (PCBs), 171, 174, 260, 277, 278
Polyesters
  drawing, 442–444, 454, 465, 467, 468
  filament yarn process, 444–445
  heat setting, 443–445
  manufacture, 441–447, 451, 452, 454, 459, 465–470
  new developments, 445–446
  polyethylene glycol, wood stabilization, 448, 468
  staple process, 444
  textured yarns, 443–445
Polyethylene
  chlorinated, 636
  high-density, 7, 44, 101, 452, 579, 585
  linear low-density, 3, 586
  low-density, 229, 455, 578, 585
  production, (Cr catalyst), 101, 377
  production, (Ziegler catalyst), 328, 329
  ultra-high molecular weight, 456, 585, 593
  ultra low density, 586

  very low density, 586
Polyethylene glycol, wood stabilization, 1091
Polyethyleneimine (PEI), 1206–1207, 1234
Polyethylene rubbers, 636–638
Polyethylenes, 2, 3, 6, 7, 11, 31, 35, 41, 43, 44, 101–103, 137, 141, 201, 228–229, 312, 313, 322, 330, 448, 452, 453, 455–457, 465, 468, 566, 571, 573, 574, 576, 578, 579, 585, 586, 590, 593, 601, 611, 612, 623, 632, 636–638, 659, 889, 902, 1084, 1091, 1443, 1444, 1501
Polyethylene terephthalate (PET), 11, 35, 131, 136, 193, 320, 377, 441, 442, 445, 446, 470, 585, 589, 592, 593, 603, 611, 1124
Polyisobutylenes, 23, 575
Polyisoprene using free radical catalyst, 626
Polymer chains
  branching, 578, 585, 586, 619, 621, 627
  cross linking, 560, 566, 620, 622, 634, 636
  crystallinity, 442, 476, 576, 586, 623
  curatives to crosslink, 555, 557, 623
Polymerization
  addition, 577
  anionic, 578
  bulk, 580
  cationic, 577, 578
  chain growth, 554, 555, 557
  chain-reaction, 577
  coordination, 579
  emulsion, 581, 613, 1461
  free radical, 565, 577–579, 586, 601, 638
  metallocene catalysts, 579, 580, 638
  photo initiated, 566, 1122
  solution, 458, 565, 580–581, 588
  step growth, 551, 554, 555, 559, 567, 577
  step-reaction, 576–577
  suspension, 581, 611, 613, 638
  synthetic rubber, 581, 620, 630, 632, 1166, 1168
  Ziegler-Natta catalysts, 575, 579, 638
Polymerization catalysts
  alkyllithium, 627, 629
  free radical, 579
  neodymium, 630
  Ziegler-Natta, 229, 455, 575, 579, 580, 626, 630, 632, 638
Polymers
  amorphous, 574, 575, 584, 608
  atactic, 452, 453, 575, 579, 586, 630, 638
  chain structure, 447, 456, 573–575, 630
  chain termination, 564
  chemical structure, 574, 576
  cold-drawing, 439, 584, 585
  creep, 445, 457, 461, 547, 583, 587, 590, 596, 597, 600–602, 639
  cross-linking, 450, 464, 560, 566, 577, 589, 596, 599–602, 610, 611, 616, 620, 622, 624, 625, 634–639, 1092, 1121, 1123–1126, 1206, 1207, 1214, 1251, 1273, 1338, 1384, 1484
  crystallinity, 440, 442, 461, 476, 576, 578, 582, 585, 586, 592, 593, 600, 623, 1058
  crystallization kinetics, 576, 623
  fatigue, 288, 468, 558, 584, 585
  glass transition temperature, 445, 565, 574–576, 582–584, 611, 619, 622–623, 631, 633, 635, 638
  hydrogen-bonding, 439, 549, 555, 574, 576, 588, 600, 623, 1058, 1064
  isotactic, 453, 579, 586
  melting point, 438, 442, 457, 461, 576, 600, 625
  molecular flexibility/rigidity, 600
  molecular weight, 445, 454, 548–550, 554, 565, 635, 1455
  orientation, 439, 442, 445, 446, 456, 457, 468, 550, 562, 587, 600, 604, 1064, 1123, 1279
  polarity, 574, 575, 588, 593, 600, 602, 608
  SEBS, 595, 637
  semicrystalline, 574, 576, 595, 632
  sequence distribution, 621
  stress relaxation, 583, 584, 639
  stress-strain behavior, 584
  structure-polymers, syndiotactic, 469, 573, 579, 600, 602, 619–622, 629, 1262
  yield point, 584
Polypropylene
  diethyl aluminum fluoride, 229
  production, 229
  Ziegler-Natta catalyst, 539, 638
Polypropylene fibers
  atactic, 452, 453, 575, 579, 586, 638
  isotactic, 452, 579, 586, 638
  production, 4, 452
  syndiotactic, 452, 453, 574, 575, 579
  synthesis, 453
Polysaccharides, 470, 1058, 1073, 1086, 1097, 1099–1103, 1107, 1108, 1112, 1142, 1148, 1183, 1201, 1218, 1250–1251, 1273–1275, 1280–1284, 1288, 1406, 1408, 1411, 1414, 1420, 1421, 1423, 1425, 1426
  hydrolysis, 1283, 1288
Polystyrene
  expanded, 588
  high-impact, 588, 589, 631, 637
  styrene/acrylonitrile copolymers, 588–589, 593, 629, 631
Polytetrafluoroethylene (PTFE)
  expanded polymer (ePTFE), 462, 470, 471
  fiber manufacture, 461
  polymer manufacture, 460–461
  properties, 461–462
Polyurethane, 33–34, 41–43, 241, 320, 330, 452, 466, 478, 548, 560, 562, 563, 567, 571, 595, 596, 601, 608–610, 613, 614, 617, 635–637, 910, 931, 1047, 1091, 1107, 1121, 1163, 1355
  flexible foam, 596, 610, 611
  markets, 596
  reaction injection molding (RIM), 596, 597, 601, 608–610, 636, 931
  rigid foam, 596, 610
  rubber, 635, 636
Polyvinyl chloride
  elastomer, 43
  flexible, 587, 600, 601, 611
  plastisol/organosol, 587
  rigid, 587, 600, 601
  thermal stabilizers, 600
Population growth, world, projections, 421, 838, 961, 965, 1329, 1347, 1348
Potassium, minerals for fertilizer, 380, 765, 844, 939, 959, 964, 969, 978–981, 985, 987, 1019, 1156, 1299
Potassium nitrate fertilizer, 910, 979, 980, 984, 1476, 1478
Potassium phosphates, fertilizer, 980
Potassium salts, fertilizer, extraction and processing, 441, 893, 964, 978–980, 1022, 1177
Potassium sulfate fertilizer, 979–980, 984, 986, 992
Pressure swing adsorption (PSA), 227–228, 326, 875, 885, 886, 889–892, 1039, 1040, 1042, 1043, 1045, 1050
Prill, 911–913, 1477, 1485, 1486, 1489
Primary alcohols, 370, 374, 1331, 1454, 1455
Printing inks, 1051, 1086, 1111, 1117–1129, 1158, 1331, 1355, 1391, 1392
Printing processes, 1124, 1125
Printing textiles, 498–499
PRISM separators, 892
Process monitoring tools

Process monitoring tools (*cont.*)
  control charts, 136, 138, 140, 141
  control limits, 137, 138, 140
  R-charts, 137, 138, 140
  s-charts, 137–139
  x-bar charts, 137, 138, 140
Process safety management (PSM), 48, 55, 65, 89–97, 99, 103, 108, 110, 125
Process variation, contributors to, 139, 142–144, 153, 314, 340, 345, 783, 905, 1019, 1367
Promoter
  CHEF1, 1233, 1234
  CHO-EF1, 1233
  cytomegalo virus, CMV, 1233
  SV40, 1233
Propanediol, 1154, 1190
Propane, products from, 73, 75, 77, 215, 311, 313–314, 316, 352, 724, 752, 770, 807–809, 812–815, 820, 821, 823, 825, 927, 929, 1051, 1224, 1362, 1453
Propionaldehyde, 279, 312, 313, 344, 363, 365, 366
Propylene, chemicals from, 15
Propylene glycols, 314, 328–300, 368, 598, 801, 1151, 1204
Propylene oligomers, 313
Propylene oxide, 3, 10, 309, 313, 314, 322, 328–330, 335, 340–341, 357, 361, 362, 369, 379, 555, 635, 801, 1455–1456, 1461, 1462, 1529
Proteases, 393, 400, 409, 413, 414, 417, 1145, 1147, 1148, 1150, 1151, 1179, 1183–1189, 1191–1197, 1203–1205, 1212, 1218–1219, 1223, 1224, 1235, 1240, 1241, 1465
Protective suit, 280
Protein
  analysis, 393
  crystallization, 625, 626, 1146, 1147, 1172
  dry formulations, 1148–1150
  expression, 393
  fibers, 463, 477, 480
Protein-free (PF), 1236
Proteins and peptides, recovery, 244, 1142, 1148, 1151
Proximol, 890
PSA.*See* Pressure swing adsorption (PSA)
PSM.*See* Process safety management (PSM)
Pulmonary agents, 403–404
  asthma and allergic rhinitis, fexofenadine, 403–404
Pulp and paper
  bleaching, 30, 1037, 1085
  environmental concerns, 879, 1525
  papermaking, 1066–1068, 1079, 1085, 1112, 1201
  pulping
    biomechanical, 1112
    chemical, 1072, 1080, 1288
    mechanical, 1067, 1071–1072
    solvent, 1077, 1078
PUREX process, 857–859, 861
Purification processing (in cell culture processes), 1240
Purisol process, 893, 894
Puromycin, 1233
Pyrites, 174, 756, 758, 789, 997, 998, 1000, 1002, 1005, 1008, 1009, 1011, 1012, 1015

R
Racemase, 1187, 1217, 1218
Radiation sources, 831, 833, 835, 870
Radioactive decay, 826, 829, 830, 838, 839, 857, 870
Radioactive waste management, 829, 848, 859–860
Radioactivity, measurement, 231
Radioisotope applications, 831, 870
Rapeseed/canola, development, 1347, 1348, 1380
Rapeseed oil and canola oil, 1202, 1334, 1336, 1337, 1358–1359, 1371, 1374, 1380, 1387, 1393
Rayon
  chemical manufacture, 424–426
  cuprammonium cellulose, 428
  high absorbency rayon, 430
  high wet modulus rayon, 430
  Lyocell, 430–431, 1086
  spun yarn, 419, 429, 468
  textile operations, 428–430, 463
  wet spinning, 426–428, 448, 452, 477
R charts, 137–140
RDX and HMX, 1481
Reactive adhesives diluents, 552–554, 558
Rectisol, 801, 893, 894
Redox reaction, in free radical polymerization, 447
Rhamnogalacturonan, 1283, 1284
Rhodium, 28, 228, 349–351, 354–356, 843, 904, 905, 926
Riboflavin, 1174, 1175
Right to Know Act, 92, 124
Risk and hazards evaluation, HAZOP, 64, 85–87, 93, 103, 110, 156
Risk management plan (RMP), 93, 97–98, 125, 128
RKN process, 885
RMP.*See* Risk management plan (RMP)
Rodenticides, 642, 696
Rosin, 19, 35, 628, 634, 1068, 1082, 1086, 1108–1111, 1408
Rubber
  elasticity, 622
  elasticity, entropy, 622
  elastic shear modulus, 622
  fillers, 619, 623–625, 630, 632, 635
  hydrogenated nitrile rubber, 633–634
  properties, 620, 622–623, 626, 633, 634, 639
  reinforcing agents, 623–624, 635, 637
  styrene-butadiene, 619–621, 626–629, 631, 636
Rubber-plastic alloys, 638
Ruthenium (Ru), 331, 857, 927, 1284

S
Saccharification and fermentation, 1097, 1099–1103, 1259
Saccharification of wood, 1102
Saccharomyces sp., 1260
Saccharum, 1404
Safety considerations
  boiling liquid explosions, 77
  catastrophic failure, 47, 97–98
  chemical process accidents, 47, 48, 99
  combustion hazards, 65–73
  damage estimates, 77, 80
  dust explosions, 75–81
  explosions, 47, 50, 53, 54, 58, 61, 64, 68, 69, 72–81, 86, 96, 99, 101–103
  government regulations, 47, 92–93
  highly hazardous chemicals, list, 92–94, 96–97
  inherently safer plants, 58
  management systems, 48, 51, 55–61, 64, 65, 91, 92, 99–101, 103
  piping systems, 48, 53–54, 57, 63, 97
  process safety management (PSM), 48, 55, 90–97, 103
  risk/hazard evaluation, 48, 55, 65, 83, 85–87
  static electricity, 48, 61–64, 71
Safety measures, 292, 998, 1149
Salt cake, 860, 1067
Salt, vacuum pan evaporation, 1018, 1019

- Samples
  - discrete, 232, 235–236
  - processing, 234, 240, 242, 243, 246
  - vapor phase, 237, 241–242
  - volatile components, 241
- Sampling
  - ground water, 241, 738
  - procedures, 239
  - strategies, 232
  - techniques, 232, 237–242
- Sasol process for synthetic fuels from coal, 816
- Scale-up, 252, 255, 289, 415, 794–795, 886, 1136, 1140–1142, 1233, 1240, 1242–1243, 1245–1246, 1295, 1524
- Scatter plot, 135
- s-charts, 137–139
- SCR.*See* Selective catalytic reduction (SCR)
- sec-Butanol, 322, 367–368, 377, 379, 381
- Selective catalytic reduction (SCR), 221, 222, 723, 755, 755, 909, 1304
- Selexol process, 893, 894
- Self-cleaning, 266, 276, 280, 283–285, 291, 1352
- Semiconductor nanoparticles, 255–256
- Sequential transfection, 1234
- Serum, cell culture, 1235, 1243
- Serum-free (SF), 1235–1238
- Sh Ble gene, 1233
- Shear force, shear stress, 603, 1149, 1240–1242
- Silicone elastomers, 634–635
- Single superphosphate fertilizer, 939, 970–972
- Six Sigma programs, 153
  - MAIC, 153
- Smart textiles, 285
- Soap
  - chemistry, 1431–1432, 1455, 1462
  - fatty acids for, 1431–1451
  - functional properties, 1433–1435
  - manufacture
    - aesthetics, 1443
    - fat refining, 1444
    - fat splitting, 1444–1448
    - fatty acid distillation, 1448–1450
    - fatty acid neutralization, 1432, 1440
    - neutral fat saponification, 1432, 1434
    - raw materials, 1432–1437, 1440–1442, 1451–1455, 1463, 1466, 1468, 1469
    - soap drying, 1438–1439
    - soap finishing, 1439, 1443
  - pellets, 1438, 1439, 1443
- Soda ash
  - dry-mining process, 1018–1020
  - production and reserves, 1020
  - Solvay ammonia soda process, 1019, 1020
  - synthetic ash process, 1431
  - uses in the US, 1017
- Sodium
  - bicarbonate, 1021, 1022, 1240, 1241
  - chlorate, 849, 850, 1017, 1036, 1037, 1067
  - chlorite,
  - hypochlorite, 1035–1037
  - hyposulfite, 1025
  - phosphates, 946, 1019
  - silicates, 219, 1011, 1025–1026, 1208, 1367, 1466
  - sulfides, 1023–1024
  - thiosulfate, 1024
- Sodium bisulfite, 508, 625, 1012, 1025
  - wood pulping, 1069
- Sodium chloride
  - consumption by usage, 1017, 1019
  - extraction by solution mining, 979, 1017, 1019
- Sodium sulfate
  - kraft wood pulping, 1069
  - major reserves, 1022
  - production and consumption, 1023
- Sodium sulfite, 26, 1012, 1013, 1024, 1025, 1112
  - sulfite wood pulping, 1024, 1112
- Soil remediation, 278
- Solar salt production, 1017–1019, 1034
- Solvated metal atom dispersion, 252–253
- Solvent separation process.*See* Organosolv process
- Somogyi method, 1278
- Sour gas, shift, 892
- Soybean
  - development, 192, 665, 1325, 1331–1332, 1366
  - transgenic varieties, 31, 665, 1325–1330, 1337
- Soybean oil, 191, 223, 553, 1124, 1161, 1325, 1332, 1336, 1346–1348, 1363, 1364, 1366, 1367, 1372–1374, 1377, 1379, 1383, 1385–1388, 1391–1394, 1396
- Specialty chemicals, 15–16, 32, 33, 35–38, 40, 43, 202, 551–553, 562, 567, 929, 1068, 1152, 1180, 1296, 1355, 1451
- Spent media analysis (SMA), 1237
- Stamicarbon, 913
- Standard deviation (data), 133–138, 141, 144, 145, 147, 1264
- Starch
  - in cane juice, 1403, 1407, 1408, 1411, 1425
  - conversion, 1152, 1423
  - enzyme hydrolysis, 1423, 1424
  - hydrolysis, 1416, 1425
  - hydrolyzates, 1423, 1424
  - sweeteners from, 1403, 1407, 1408, 1411, 1414, 1416, 1417, 1421, 1423–1425, 1427
- Statistical methods
  - box plot, 132, 133
  - business process improvement, 131, 153–154
  - chemical applications, 131, 149–153
  - contributions to process variations, 143
  - control charts, 138–141
  - dot and box plots, 132
  - factors, influence of, 195
  - factors *vs.* response, 144–149
  - Gaussian probability, 135
  - hierarchical data structure, 143, 144
  - histograms, various, 132, 133
  - industrial experimentation, 141–149
  - lattice design, 150
  - MAIC, 153
  - mean, 134–144, 147
  - median, 133, 134
  - mixture experiments, 149–152
  - model building, 131, 152–153
  - quadratic response surface, 148
  - run chart, 131, 132
  - scatter plot, 135
  - simplex centroid design, 150, 151
  - simplex coordinate system, 149, 150
  - six sigma, 153
  - standard deviation, 133–138, 141, 144, 145, 147
  - stem and leaf plots, 133, 136
- Sterols, 651, 652, 689, 690, 1344
- Stover 5c, 1270
- Styrene, 3, 7, 10, 35, 36, 42, 44, 60, 63, 82, 211, 275, 313, 320, 322, 328–330, 335, 340–341, 377, 379–380, 447, 578, 582, 584, 588, 589, 593, 595, 598, 601, 620, 627–629, 631, 636, 637, 801, 1092, 1415, 1465, 1466

Styrene (*cont.*)
  inhibitor, 60, 82, 330, 344
Styrene-butadiene rubber, 35, 619, 620, 626–633, 636, 1110
Subtilisin, 1179, 1183, 1185, 1194, 1195, 1224
Succinic acid, 23, 1107, 1154, 1158, 1160, 1286, 1531
Sucrose
  derivatives, 1423
  sucralose, 1423, 1426
  sucrochemistry, 1422–1423
  sucrose acetate isobutyrate, 1423
Sugar
  brown, 1404, 1412, 1416, 1417, 1425, 1427
  consumption (cane and beet), 1409, 1412, 1422
  direct consumption, 1409, 1412
  genetic engineering, 1427–1428
  history, 1403–1404
  bagasse, 1411–1412, 1421
  organic, 1412
  powdered, 1416, 1417
  raw cane sugar, 1404, 1425, 1427
  regulations and trade, 1412, 1427
  specialty, 1416–1417, 1427
  from wood, 1405, 1413
Sugar act, 1427
Sugar beet, sugar production, 1403, 1404, 1409, 1410, 1413, 1418, 1420, 1422, 1427
Sugarcane
  breeding programs, 1405
  history, 1403–1404
  processing, 1406–1410, 1418, 1420, 1425, 1427
Sugarcane processing
  centrifugation, 1410, 1424
  crystallization, 1410, 1415–1416, 1420, 1421
  diffusion, 1407, 1408
  evaporation, 1409, 1420
  juice extraction, 1406–1407
  juice purification, 1408, 1419–1420
  milling, 1406–1408, 1411, 1417
  sulfitation, 1408–1409, 1412, 1419
Sulfar fibers, 464, 465
Sulfinol, 892, 894
Sulfite process, wood pulping, 1075–1077, 1086
Sulfolane, 893, 894
Sulfur
  block, 999
  bright, 997
  Car-Sul, 1000, 1003
  claus process, 1001, 1003, 1004
  consumption, 1005–1006
  corrosion, 998, 999, 1003, 1010, 1012, 1014
  crushed bulk, 999
  dark, 997, 1003
  deposits, 997, 1000, 1002
  dust, 998, 999
  ellethorpe, 1000
  forming
    air prilled, 999
    granular, 998, 999
    pelletized, 999
    wet prilled, 999
  Frasch mining, 997, 1001, 1002, 1006
  from H_2S, 1012–1013
  liming, 1000
  production, 997, 998, 1002–1003, 1005
  properties, 997
  solidification, 998–1000
  transportation and storage, 997–999
Sulfuric acid
  industry development, 1007–1009
  manufacture, contact process, 1008, 1010
  production and consumption, 1015
  properties, 1007
  uses, 1007
Sulfur industry, development, 999–1002
Sulfur melting, 1009
Sumitomo-BF PSA process, 892
Sunflower seed, 1337, 1342, 1344, 1347, 1348, 1350, 1358, 1359, 1364, 1370, 1380, 1432
  development, 1342, 1347, 1348, 1350
Supercritical fluid chromatography, 244
Supercritical fluids, 756, 885, 887, 889, 1049, 1223–1224
Superphosphoric acid, 945, 950–951, 954–956, 980, 987
Surface active agents, 17–19, 427, 663, 1452
Surface (in adhesion), definition, 550
Surfactant production, intermediates
  alkylbenzene, 1451, 1452, 1456–1458, 1460
  alkylphenol, 1451, 1452, 1456, 1461
  ethylene and propylene oxides, 1455–1456, 1461, 1462
  fatty alcohols, 1444, 1451, 1453–1455, 1461–1463, 1470
  fatty amines, 1456, 1461
  olefins, 1451, 1454–1456
Surfactants
  amphoteric, 1463
  anionic
    alkyl and ether sulfates, 1460–1461
    alkylaryl sulfonates, 1457–1460
    alpha olefin sulfonate, 1461
    methyl ester sulfonates, 1461
    sulfonation, 1458–1460
  cationic, 1463–1464
  consumption, 1456–1457
  ethoxylated ethyl esters
  features, 1451
  intermediates, 1453–1456
  nonionic, 1461–1463
  production, raw materials, 1452–1456
Sustainable chemistry, 156
Sweeteners
  carbohydrate based, 1426
  relative sweetness, 1426
Synthesis gas
  purified, 375, 801, 877, 890–895, 1043
  sulfur removal, 892
Synthetic detergents, 17, 1431–1470
Synthetic organic chemicals production volume, 326

T
Tall oil, 1073, 1074, 1086, 1105, 1108–1111, 1337, 1344, 1456
Talodura process, 1412
Talofloc process, 1414
Tannin, 1088, 1108, 1111–1112, 1267
Terephthalic acid (TPA), 3, 10, 13, 275, 309, 319, 320, 327, 335–338, 342, 373, 377, 441, 442, 445, 447, 589, 1154
Terpenes, 19, 1109, 1168, 1169
Tetrachloroethylene, 172, 1120
Tetracyclines, 653, 1176, 1178
Textile fibers
  fiber consumption, 421–424
  history, 420–421, 457
  hydrophobic/hydrophilic, 478

  manufacture
    acrylics, 446–448
    cellulose acetate, 432–434
    elastomeric, 451–452
    microdenier, 445, 465, 478
    nylon, 421, 436–438
  physical properties, 420, 446, 459
  polyesters, 441–446
  polyolefins, 452–457
  rayon, 419–431, 434–436, 448, 449, 452, 460, 463, 464, 467, 468, 470
  tensile properties, 460, 468
Thermodynamic decomposition potential, 69, 190, 205, 821, 905, 1148, 1151, 1294, 1301, 1502, 1503
Thermolysin, 1195, 1214, 1218–1219
Thermoplastic elastomer
  polyetheramide, 594
  polyetherester, 451, 596, 608
  polyurethane, 595, 596, 636
  styrene-diene, 585, 588, 589, 601, 621, 629, 631, 636
  thermoplastic olefin, 632
  thermoplastic vulcanizates, 595
Thermoplastics
  engineering, 585, 592, 597, 615
  major markets, 633
  specialty, 585, 590–593
Thermosetting plastics, 572, 598, 608, 614, 615
Thiols, curative, 413, 555, 557
  in epoxy adhesives, 555, 557
Tin tetrachloride, 630
Tocopherols, tocotrienols, 25, 1344, 1346, 1377, 1384
Toluene derivatives, 310–311
Toluene diisocyanate (TDI), 35, 319, 560, 561, 596, 609, 635, 910
Top-down, 86, 250, 256, 265, 286
Total elements determination, 232–233, 243
Toxicity studies, 654
TPA.*See* Terephthalic acid (TPA)
Trans fats nutritional labeling, 1381–1383
Transfectin, 1234
Transmutation products, 843–845
Triacylglycerols, triglycerides, 1169, 1331, 1333, 1335, 1338–1342, 1363, 1370–1372, 1374, 1376, 1381–1384, 1388, 1389, 1393, 1397, 1398
Triaminotrinitrobenzene (TATB), 1478, 1479, 1482
Trichlorobenzene, 1482
Trichoderma reesei, 1186–1187, 1194, 1197, 1275, 1277, 1278
Trinitrotoluene (TNT), 27, 77, 79–81, 101, 1475–1481, 1492
Triple superphosphate fertilizer, 970, 972–974
Trommsdorff effect, free radical polymerization, 565
Tryptophan, 1170–1174, 1179, 1217
Turpentine, 19, 1074, 1086, 1104, 1105, 1108–1111, 1296, 1342–1343
Two-stage acid hydrolysis, 1267

U
Ubiquitous chromatin opening elements (UCOE), 1234
Uranium fuel cycle, 848–859
Uranium processing, 855–859
Urea
  consumption, 878
  dimethyl, 919
  fertilizer
    granular, 985, 991
    prill, 912, 913, 999
    processes, 913, 916, 922, 923
  properties, 913
Ureaform, fertilizer, 991
Uretidinediones, in urethane adhesives, 560

V
Vacuum collapse, 82
Vanillin, 19, 1086, 1292
  from wood, 1086, 1292
Vapor phase chromatography, 233
Vectors, disease, 92, 1164, 1229
Vegetable oils, catalytic hydrogenation, 223–224
Vinyl acetate, 312, 313, 327, 342, 347, 447, 450, 586, 587, 601, 800, 1081, 1129
Vinyl chloride, 2, 7, 9, 41, 313, 326–328, 342, 356–360, 450, 587, 601, 611, 638, 1027, 1053, 1120, 1121, 1129
Vinyls, 2, 42, 449–450, 614
Viscoelasticity, 583–584
Viscometric assays, 1279
Vitamins, fat-soluble, 1331, 1342, 1344, 1355
Volatile organic compounds (VOC), in coatings, 165, 193, 201, 213, 274, 279, 750, 769, 1118–1124, 1126, 1360
  regulations, 1118–1120
Vulcanization, 23, 594, 595, 622, 632, 638

W
Waste management, 97, 98, 161, 193, 292, 829, 848, 859–860
Waste minimization, hierarchical rules, 161–165
Water
  boiler feed, 886, 893, 1013
  drinking, 232, 237, 242, 276–278, 955, 1035
  ground, 171, 173, 181, 232, 237, 241, 266, 276, 277, 280, 294, 750, 863, 879, 957, 1488
  waste, 85, 162, 164–166, 178, 193, 237, 243, 262, 293, 347, 491, 755, 780, 785, 901, 1024, 1035, 1139, 1203, 1299, 1480, 1528, 1531, 1533–1535
Water Environment Federation, 237
Water gas shift reaction, 207, 324, 325, 349, 375, 772, 795, 883, 888–889, 892, 1251, 1252, 1297, 1521
Water glass, 1025
Water remediation, 277–278
Waxes, 429, 566, 624, 639, 659, 700, 752, 795–797, 991, 992, 1081, 1084, 1110, 1267, 1323–1398, 1408, 1437, 1442, 1455, 1479, 1484
Wax, in acrylic adhesive, 566
Wet chemical analysis, 1273
Wet scrubbing, 894, 1010, 1299
Wood
  Bergius Rheinau process, 1100
  biotechnology chemicals, 1093, 1097, 1104
  chemical composition, 1058–1061
  combustion, 1097–1099
  composites, 1058, 1089, 1092, 1096
  destructive distillation, 1103–1105, 1108
  energy, fuels, chemicals, 1097
  fire retardants, 1096–1097
  gas and oil from, 1086, 1103–1106, 1110
  hydrolysis, 1099, 1101
  Madison process, 1101
  medicinal, 1086, 1108, 1111
  modified wood, 1091–1093, 1095
  naval store chemicals, 1105
  polymer composites, 1086–1087, 1090–1092
  preservation, 1095–1096

Wood (*cont.*)
  pulp and paper, 1066–1068
  pulping, processes/yields, 1066, 1101, 1111, 1112
  saccharification, 1058, 1097, 1099–1103, 1113
  structure, 1057, 1058, 1061–1085, 1092, 1112
  thermal decomposition, 1092, 1097, 1103–1104
  uses, summary, 1057–1059, 1066, 1072, 1075–1078, 1083, 1085, 1086, 1088, 1090–1092, 1094, 1097, 1098, 1104, 1105, 1108, 1110, 1111
Wood pulp bleaching, environmental impact factors, 184, 185, 1036
Working cell bank (WCB), 1235

X
Xanthan gum, 35, 1162
x-bar charts, 137
Xylan, 1059, 1107, 1201, 1265, 1267, 1270, 1275, 1278, 1280–1282
Xylanases, 1080, 1112, 1192, 1194, 1201, 1281, 1282
Xylene derivatives, 310–311
Xyloglucanases, 1282
Xylose, 1059, 1099, 1100, 1103, 1154, 1166, 1170, 1180, 1184, 1187, 1195, 1200, 1201, 1258–1260, 1275, 1278, 1280–1283, 1285, 1286, 1290

Z
Zeocin, 1233
Zeolites, 10, 27, 202, 203, 213, 215, 216, 271, 274, 314, 316, 318, 332–334, 352, 733–736, 797–798, 824, 918, 920, 930, 1012, 1026, 1152, 1291, 1425, 1464–1466, 1470
Ziegler alcohol reaction chemistry, 1455
Ziegler-Natta, 455, 579, 580, 626, 630
  catalyst, 575, 579, 626, 630, 632, 638
Zirconium processing, 845
  production, 847–848
Zymomonas mobilis, 1170, 1259